Designing Invisible Handcuffs

Formal Investigations in Institutions and Organizations
for Multi-agent Systems

SIKS Dissertation Series No. 2007-16

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

Printed by Gildeprint Drukkerijen B.V., Enschede
ISBN 978-90-393-4619-8

Designing Invisible Handcuffs

Formal Investigations in Institutions and Organizations
for Multi-agent Systems

Ontwerpen van Onzichtbare Handboeien

Formele Onderzoeken naar Instituties en Organizaties voor
Multi-agent Systemen

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. W.H. Gispen,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op maandag 17 september 2007 des ochtends te 10.30 uur

door

Davide Grossi

geboren op 20 april 1979, te Udine

Promotor:	Prof. dr. J.-J.Ch. Meyer
Co-promotor:	Dr. F.P.M. Dignum

Ai miei genitori

Contents

Acknowledgments

"Right now it is only a notion, but I think I can get the money to make it into a concept then later turn it into an idea."

W. Allen, "Annie Hall", 1977

Before starting my PhD at the Intelligent Systems group I had already spent a couple of months there, working on my master thesis to be defended at the University of Pisa, where I graduated. At that time many notions about norms, institutions, organizations and agents were loitering around my mind. It then happened, somehow, that my current co-promotor —and I will always be grateful to him for that— found some money, which was turned into my 4 years contract. Now, 4 years later, I present in this thesis what has grown out of those notions. Some of them became concepts, some of them ideas. In the end, in one way or another, they have made it to become theories of some kind[1].

Still, the title remained quite a notion-like one, as well as the illustration on the cover, which is inspired by works of Banksy, a graffiti artist active in London (`www.banksy.co.uk`). Institutions and organizations lay boundaries upon agents' activities. But such boundaries do not work like cages, rather like norms. And norms, we know, can be violated. They do not work like physical handcuffs, rather like some sort of subtle, invisible ones.

The speculations about title and cover could go on, of course, virtually for many pages. However, my plan is now to sincerely thank a number of people. First of all, I want to thank my supervisors John-Jules Meyer and Frank Dignum. They have both been excellent supervisors, always ready to hear with interest what I had to tell. I thank John-Jules especially for always being so supportive and attentive, for his cheerfulness and for his sincere scientific curiosity. Meeting with him is always such a great pleasure. I thank Frank especially for giving me the great freedom I enjoyed in my researches, for really helping me in making my research aims concrete and feasible and for always stimulating me with new problems and issues. I also thank Johan van Benthem, Cristiano Castelfranchi, Andrew Jones, Jan

[1]Many thanks to Riotmaker Productions (`www.riotmaker.net`) for providing me with the correct reference of the opening quote of this preface, and for much of the music I played in the background while writing this thesis.

van Eijck and Rineke Verbrugge for accepting to be in the reading commission of this thesis. It has been an honor for me to have such a distinguished commission.

Science is a collective enterprise. I want to thank all my co-authors: Frank Dignum, John-Jules Meyer, Mehdi Dastani, Virginia Dignum, Huib Aldewereld, Javier-Vàzquez-Salceda and Lambèr Royakkers. Besides all my co-authors, many other people have actively took part in the elaboration of the contents of this thesis. Chapter 2 have substantially benefited of the suggestions of the anonymous reviewers of CLIMA'04 and CLIMA'05, CRR'05 and of Gerard Vreeswijk. Chapter 4 is, in my view, a typical example of what can be achieved by constantly exchanging ideas with researchers working on the same issues. I sincerely thank them all. The chapter grew out of discussions with Antonino Rotolo, Guido Governatori, Guido Boella, Leon van der Torre[2], Cristiano Castelfranchi, Frank Hindriks, and, in particular, with Andrew Jones. By discussing with him my first contribution to the formal analysis of counts-as presented at ICAIL'05, I've become clearly aware of what still was lacking. Many of the insights presented in Chapter 4 grew out of the attempt to find answers to the questions raised in that discussion. The anonymous reviewers of the Journal of Logic and Computation and of DEON'06 have also significantly contributed to the present form of the chapter. I am very grateful to them as well. Chapters 5 and 6 have benefited of the comments of the anonymous reviewers of AAMAS'05, AAMAS'06, COIN'06, AAMAS'07, Juan Antonio Rodríguez-Aguilar and, in particular, Philippe Pasquier. I am very grateful to them all. I finally want to thank Rineke Verbrugge for having the patience to provide a detailed commentary to a first version of this thesis, and Tom van der Weide, Joost Westra, John-Jules Meyer and Frank Dignum for helping with the Dutch summary.

Special thanks go to the whole Intelligent Systems Group. Doing research there has been a scientifically stimulating and humanly enriching experience. In particular I want to thank Jurriaan van Diggelen, Geert Jonker, Henk-Jan Lebbink, Paul Harrenstein, Huib Aldewereld, Javier-Vàzquez-Salceda, Birna van Riemsdijk, Cees Pierik, Mehdi Dastani, Henri Prakken, Jan Broersen, Martin Caminada, Susan van der Braak, Joris Hulstijn, John-Jules Meyer, Frank and Virginia Dignum, Tom van der Weide, Joost Westra and Paolo Turrini. With them I shared rooms, daily lunches, coffee breaks, dinners, Sinterklaas, snow-board vacations, picknicks in Vondelpark, skate evenings, screenings of Dutch and Italian cinematographic masterworks, pizzas (Hawaii), Hawaii, conferences, late nights in Amsterdam and a lot of "gezelligheid". It has been a great time! And many thanks go also to our Balie, Wilke Schram, Richard Starmans, and the guys at the Systeembeheer, for always being so helpful and cooperative.

I thank the whole group of good friends I've built up in these 4 years spent in Utrecht. First of all the indispensable little Italian enclave in Utrecht, i.e., Adriano Galati, Paolo Turrini, Marco and Yildith, Matteo, Carlo and Claartje. You are all different examples of what I love of my country. Then I thank Pierre-Olivier Guerra and Andreas Olofstam for all the delicious dinners, the Duvel sessions and the nice

[2]I want to say, in passing, that the discussions with Leon were definetely the most pleasant. Always over a glass of nice wine or Belgian beer, after great BBQs in the relaxing atmosphere of his garden in Diemen Zuid. Thanks for that!

times in Malmö, the group of my first glorious Dutch course in Autumn 2003, Karin de Wit, Leila Kushan, Stella Vassilaki, Asia Korecka and the group of the "Kroeselaan mansion", the great people of UFO (Utregse Frisbee Organisatie), Caroline Schmidt, and Magdalena Koralewska. I also want to thank the housemates I've lived with in W.A. Vultostraat 7 in all these years, namely, Javier-Vàzquez-Salceda, Petar Todorov, Henri Ervasti and the newly arrived Marcin Zielinski, since I can say the atmosphere has always been nice and relaxed.

Strangely enough, if you can move abroad and still feel home it is also thanks to the people you've left behind. I'm lucky to have a number of long-lasting friendships which are completely indifferent to geographic distance and which, in these years, have shown to be alive probably more than ever. I thank first of all the the whole Udine/Friuli group and, in particular, Paolo Ermano ("il Bauli", also known to me as "Mastello"), Nicola Trifiletti ("il Nic"), Marco Tozzi ("il Tozzi"), my "kumce" Gorica and her family Drago, Vesna and Goran, and my Virgil in the "Udine by night" Giulio Cerno. I then want to thank the whole Pisa group and, in particular, Alessandro Torza ("il Torza") e il Pensiero Puro, Federico Torza with "Borrelli e i Tupacamaru", Roberto Colozza ('r "Murena" but not "il Colozza" since he's from Rome), Francesco Capurro ("il Capu") und die Übergesellschaft, il Balordo (that's just it), Marta Belardinelli, Carlo Proietti, Marco Bresciani ("il Brescia"), and also Loredana Santoro, who does not belong to either of these groups but who also is one of these people making me lucky.

My final thanks go to my whole family and in particular to my mother Maria, my father Danilo, my brother Marco and my grandmother Alice. If it is true what I said about my old friendships, that is even more true about them. I know they will always be there for me, supporting my choices and being proud of me. In their honor, so to say, I conclude by quoting a small excerpt in dialect taken from a popular tale from where I come from, the region of Friuli. I think it nicely summarizes, in simple popular words, what doing research is all about:

"Ogni mês si fâs la lune
Ogni dì si impare une[3]"

which in English means, "Every month the moon grows / Every day you learn something new."

Utrecht, 15.07.2007

[3]Unknown author, *The master of all masters*, popular tale from Friuli, Italy.

Chapter 1

Introduction

"Personne n'a jamais vu un État. Ni à l'oeil nu ni au microscope, ni en photo ni d'avion."

"Nobody has ever seen a state. Not with his bare eyes, nor using a microscope, nor in a picture, nor from a plane."

R. Debray, "L' État Séducteur", 1993, p.65

As the opening quote suggests, there is something peculiar about social notions which makes them difficult to grasp, namely their radical invisibility. After all, it is not for nothing that the notorious Smithian "hand" is "invisible[1]" ([Smith, 1776], Book 4 Chapter 2). The following quote[2] insists on this aspect:

> "A foreigner visiting Oxford or Cambridge for the first time is shown a number of colleges, libraries, playing fields, museums, scientific departments and administrative offices. He then asks 'But where is the University? I have seen where the members of the Colleges live, where the Registrar works, where the scientists experiment and the rest. But I have not yet seen the University in which reside and work the members of your University.' It has then to be explained to him that the University is not another collateral institution, some ulterior counterpart to the colleges, laboratories and offices which he has seen. The University is just the way in which all that he has already seen is organized. When they are seen and when their coordination is understood, the University has been seen" ([Ryle, 1949], 17-18).

[1] "[. . .] every individual [. . .] neither intends to promote the public interest, nor knows how much he is promoting it. [He] intends only his own gain, and he is in this, as in many other cases, led by *an invisible hand* to promote an end which was no part of his intention" ([Smith, 1776], p.400).

[2] The quote has been suggested to us by [Hindriks, 2006].

In human societies, institutions and organizations (e.g., universities, companies, legal systems, etc.) have developed as means for regulating interactions between human agents in order to guarantee these interactions to enjoy certain desirable properties. In a way, they set invisible boundaries —"handcuffs"— to the outstretch of the "invisible hand". Purpose of this work is to investigate what institutions and organizations are, and to provide tools for giving formal foundations to this enterprise. This will provide the means for a formal analysis of such objects as institutions and organizations making them, so to say, visible. The present work is, as such, predominantly a study in logic applied to the theory of institutions and organizations, where with 'applied logic' we mean what is made precise in the following excerpt:

> "Logic is not applied to philosophical problems the way an engineer might apply some technique for computing stress in a bridge. [...] Logic offers a technical language with relatively precise meanings as an enhancement of philosophical discourse, and an aid to precise *communication*. This is as useful as having mathematical language around in other disciplines: perhaps not just the Book of Nature, but also the Book of Ideas is written in mathematical language" ([van Benthem, 2006], p.3).

Broadly speaking, the aim of this work is therefore to put a number of techniques in place which can make our discourse on institutions and organizations on the one hand theoretically more rigorous, and on the other hand more effective when it comes to its use in the study and in the design of concrete institutions and organizations in a computational setting. In order to pursue this aim a number of precise questions need to find an answer. This introductory section is devoted to their systematic exposition.

1.1 Research Questions

In Multi-agent Systems (MAS) software agents interact which enjoy some degree of autonomy ([Wooldridge and Jennings, 1995]), exactly like human agents in human societies. As a consequence, in MAS the same problem arises of guaranteeing the designed system to exhibit some desired global properties without hampering agents' autonomy. On this ground, the opportunity of a 'technology transfer' from the field of organizational and social theory to distributed Artificial Intelligence first ([Fox, 1988]), and then to MAS ([Dignum, 2003; Vázquez-Salceda, 2004]) has often been advocated. In recent years, some research in MAS is even aiming at the explicit incorporation of entities such as organizations and institutions in computer systems, as testified by several contributions in the AAMAS conference series ([Gini et al., 2002; Rosenschein et al., 2003; Jennings et al., 2004; Dignum et al., 2005; Weiss and Stone, 2006]), the ALFEBIITE project ([Pitt, 1999]), the COIN workshop series ([Boissier et al., 2006; Dignum et al., 2007]) and the Normative Multi-agent Systems seminar ([Boella et al., 2007]).

Notwithstanding this interest for institutions and organizations in the literature on MAS, not much foundational research have been done on trying to answer the key questions "What is an institution?" and "What is an organization?" which, leaving ontological ambitions aside, boil down to the more low-profile questions "How can an institution be formally represented?" and "How can an organization be formally represented?".

1.1.1 Normative systems

To import the notion of institution in the design of MAS means, in our view, to assume a normative system perspective on the to-be-designed system, that is, to think of it in normative terms:

> "[. . .] law [and] computer systems [. . .] may be viewed as instances of *normative systems*. We use the term to refer to any set of interacting agents whose behavior can usefully be regarded as governed by norms" ([Jones and Sergot, 1993], p.276).

The normative system perspective on institutions is, as such, nothing original and it is already a quite acknowledged position within the community working on electronic institutions, or e-Institutions ([Vázquez-Salceda, 2004]). What has not been sufficiently investigated and understood with formal methods is, in our view, the question: what does it amount to, for a MAS, to be put under a set of norms[3]? Or in other words: what does it mean for a designer of an electronic institution (e-Institution) to state a set of norms, to 'make the rules' for the system? And how could a norm-aware agent understand those norms, or in other words, what is the meaning of those norms? That these questions need to be answered formally is a matter of usability of norm formulations in e-Institutions. The content of a norm needs to be communicable for a designer to check the resulting system, and for an agent to be able to actually use those norms in its deliberations:

> "If it were not possible to *communicate* general standards of conduct, which multitudes of individuals could understand, without further direction, as requiring from them certain conduct when occasion arose, nothing that we now recognize as law could exist" ([Hart, 1961], p.121).

The development of MAS needs formally shaped laws, so to say. So, what does it precisely mean to state a set of norms? We advance a precise thesis on this issue, which is inspired by seminal work in social theory:

> "Now, as the original manner of producing physical entities is creation, there is hardly a better way to describe the production of moral entities than by the word *'imposition'* [impositio]. For moral entities do not arise from the intrinsic substantial principles of things but *are superadded to things already existent and physically complete*" ([Pufendorf, 1688], pp. 100-101).

[3]In this work the terms 'norm' and 'rule' will be used as synonyms.

By ignoring for a second the philosophical jargon of the Seventeenth century we can easily extract an illuminating message from Pufendorf's words: what institutions do is to impose properties on already existing entities. That is to say, institutions provide descriptions of entities by making use of conceptualizations that are not proper of the common descriptions of those entities. For example, that cars have wheels is a common factual property, whereas the fact that cars count as vehicles in some technical legal sense is a property that law imposes on the concept "car". To say it with [Searle, 1995], the fact that cars have wheels is a *brute fact*, while the fact that cars are vehicles is an *institutional fact*. Now, if a fact corresponds to the enjoying of a certain property, or description, by some elements of a given domain, then what institutions do is to build institutional descriptions of the elements of the domain upon brute ones. For instance: "objects enjoying such and such properties (e.g. having wheels) count as vehicles. Our first research question is formulated as follows.

Research Question 1. *How can we formally represent institutions once they are understood as the imposition of institutional descriptions over brute ones?*

This question is answered in two steps. Chapter 2 provides the formal tools needed, and Chapter 5 shows that those tools are actually suitable for capturing a number of aspects of institutions which are of particular relevance for MAS.

1.1.2 Abstract and concrete norms

Different normative systems can obviously issue different and inconsistent norms: what is permitted in one institution might be forbidden by a different one. Besides this kind of 'incompatibility' relations between normative systems, a more interesting one arises when an abstract normative system, e.g., the constitution of a land, is made more concrete by another normative system, e.g., a piece of local legislation. The issuing of norms, as it appears in human institutions and eminently in law, has the characteristic of stating norms in such a form that allows them to regulate a wide range of situations and, normally, to be stable for a long period of time. This is achieved by stating norms by means of terms which are vague in a precise sense, that is, which are amenable of further interpretation within certain precise boundaries. In legal theory such terms are said to be "open-textured" ([Hart, 1961]). In law, by interpreting the open-textured terms occurring in legal texts, abstract norms are refined to more concrete ones, and such an interpretation-based refinement of norms is achieved by several levels of legislation and rule-giving from the constitutional one up to the most concrete one of judicial interpretation.

As first argued in [Dignum, 2002], this pervasive feature of human rule-giving is of extreme relevance also for the design of agent institutions. The point is easily explained by a simple example. Suppose a designer wants to develop a MAS handling citizens' data for organ transplantation purposes[4]. The system may for instance be put under the norm "*it is forbidden to discriminate recipients on the basis*

[4]This has actually been attempted. See [Vázquez-Salceda, 2004].

of nationality". The straightforward representation of the meaning of this norm would boil down to something like "transitions `discriminate(x,y,nationality)` should not occur in the system". However, at a system design level, it is very unlikely that the agents operating the MAS will explicitly have an action such as `discriminate(x,y,nationality)` available among their executables. They would rather have a number of concrete protocols some of which, if executed in some precise situations, could give rise to the discrimination of some recipient on the basis of his/her nationality. The level on which norms are specified is more abstract than the level at which the system is specified.

The point is for the designer to understand what it means for the agents in the system to discriminate on the ground of nationality in terms of their executables, and of the protocols that could be built on them. From an institutional standpoint norms need, in order to be incorporated in the system itself, to be somehow translated to a level in which their impact on the system can be described directly. Such translation is a process of norm-refinement which generates more concrete norms, and thus more concrete institutions. We get thus to the second research question.

Research Question 2. *How can we formally account for the relation between more abstract and more concrete institutions, by also giving a formal account of open-texture in institutions?*

Chapter 2 tries to answer this question and Chapter 3 puts the answer in perspective with the broader field of context theory. The issue of abstract and concrete norms comes back in Chapter 5 where it is more tightly addressed from the point of view of MAS.

1.1.3 Regulative and constitutive rules

According to many studies in legal and social theory, normative systems of high complexity, such as law, consist of regulative as well as non-regulative components ([Rawls, 1955; Hart, 1961; Alchourrón and Bulygin, 1971; Jones and Sergot, 1992]). That is, they do not only regulate existing forms of behavior, but they actually specify and create such forms:

> "As a start, we might say that regulative rules regulate antecedently or independently existing forms of behavior [...]. But constitutive rules do not merely regulate, they create or define new forms of behavior" ([Searle, 1969], p.33).

In other words, regulative rules concern what ought to be the case (e.g., "vehicles are not admitted into public parks"), while constitutive rules concern what counts as what in a given institution (e.g., "cars count as vehicles"). The paradigmatic syntax of constitutive rules has been taken to be, since [Searle, 1969] and [Searle, 1995], the form of "counts-as" statements:

> [...] "institutions" are systems of constitutive rules. Every institutional fact is underlain by a (system of) rule(s) of the form "X counts as Y in context C" ([Searle, 1969], pp.51-52).

Counts-as statement are stated relatively to a context because what counts-as what differs from institution to institution. It is our main thesis to conceive of counts-as statements as the basic bricks constituting the 'imposition' of institutional properties over brute ones. This stresses, along with what is upheld in [Searle, 1969, 1995], that no institution is given without constitutive rules and therefore without "counts-as".

Now, since the publication of the first paper in deontic logic [von Wright, 1951] much research has been devoted to the study of the formal aspects of regulative rules, giving rise to a considerable amount of literature[5], but attention from researchers in logic to constitutive rules is instead only one decade old. The first contribution in this direction is [Jones and Sergot, 1996], and —to our knowledge— no more literature has been devoted to the formal analysis of constitutive rules until [Governatori et al., 2002; Gelati et al., 2004] and [Boella and van der Torre, 2003; Boella and Van der Torre, 2004; Boella and van der Torre, 2005], which were motivated by the recent developments in MAS heading towards the introduction of institutional notions in agent systems design.

Our work conceives of counts-as statements as statements constituting the conceptualizations that institutions impose on the to-be-regulated domain. We will show that this perspective yields a much more comprehensive formal account of constitutive rules than the ones available in the literature up to this moment. We get thus to our third research question.

Research Question 3. *What is the logic of counts-as statements, and thus of constitutive rules? And how is this logic related to the logic of regulative rules?*

The first part of this question is answered by Chapter 4. The second part of the question is not answered in one fell swoop, so to say, but rather by a discussion which will start in the final part of Chapter 2 and end in the final part of Chapter 5 by reinterpreting the results of Chapter 4 on constitutive rules. We can already give a sneak preview: regulative rules will be seen as those rules of a given institution which 'constitute' the notion of violation according to that institution.

1.1.4 Organizations

Besides institutions, organizations have also obtained increasing attention from the MAS community in the last years. In fact, many methodologies for MAS are based on organizational structures as their cornerstones. The organizational structure of a MAS concerns the agent roles and their relations by means of which the overall behavior of the system is specified. These abstract types of structures are traditionally studied in the branch of sociology called *mathematical sociology* ([Fararo, 1997; Sørensen, 1978]).

Again, despite the importance attributed to this notion, little foundational work has been done in MAS trying to make the notion of organization more precise. It is a fact that what was the case up to the sixties in literature on sociology and

[5]See the series of international workshops in deontic logic and computer science (DEON) started with [Meyer and Wieringa, 1993].

organization theory, is now being the case in the literature about organizations in MAS.

> "The word 'structure' is found extensively in the literature of the social sciences. 'Social structure' and such related concepts such as 'kinship structure', 'authority structure', 'communication structure', and 'sociometric structure' are commonplace. [. . .] But despite the widespread use of structural concepts in the social sciences, it is fair to say that the formal analysis of structure has been relatively underdeveloped in these fields. The technical terminology employed in describing structures is meager; few concepts are defined rigorously. As a consequence, the social scientific description of structural properties tends to be couched in ambiguous terminology, and detailed studies of structure, as such, are rather rare" ([Harary et al., 1965], p.1).

The literature on MAS —with very few exceptions such as in particular [Hannoun et al., 2000; Hübner et al., 2002]— addresses this type of structures only in an informal way by means of diagrams and charts like in [Decker and Lesser, 1995; Horling and Lesser, 2004]. In such informal studies, many issues remain hidden behind such pictures. We believe that, in order to successfully import the notion of organization for the design of MAS, "the space of organizational options must be mapped, and their relative benefits and costs understood" [Horling and Lesser, 2004], and to provide such a "map" a rigorous analysis of organizational structure plays a crucial role. If we want the notion of structure to be of any practical use for MAS, pictures are plainly not enough, since they fail to address two fundamental aspects of organizational structures: 1) The formal (graph-theoretical) properties of the links between the roles in the structure; 2) The 'meaning' of those links, that is to say, the effects they have on the activities of the agents operating the organization. These considerations lead us to the fourth research question.

Research Question 4. *How can we formally represent organizational structure and its effect on agents' activities?*

We provide an answer to this question in Chapter 6.

1.1.5 Institutions vs. organizations

Once all the previous question have been answered, it becomes also possible to precisely understand in what institutions and organizations differ. On the one hand they are both, broadly speaking, coordination means but on the other hand they work with different concepts: institutions with norms, organizations with structures. This will become particularly evident in discussing the notion of role. Roles will be studied, from an institutional point of view, as sets of norms and, from an organizational point of view, as positions in a structure. Aim of this work is also, on the ground of the formal analysis proposed, to point at the differences and similarities of these two perspectives. This would clarify, for instance, when it would be better for a designer to think in institutional terms rather than in organizational ones and vice versa. These considerations lead to the last research question.

Research Question 5. *What are the similarities and the differences between the notion of institution and the notion of organization?*

An answer to this question is provided in Chapter 6.

1.2 Overview of the Thesis

The thesis consists of two parts. The first part is more theoretical and concerns the study of a number of formal tools which directly answer, or which are necessary to answer, some of the research questions raised in the previous section. Familiarity with normal propositional modal logic and description logic[6] is presupposed. In discussing the solutions proposed we will constantly relate them to proposals advanced in other research areas, trying to stress the general nature of the formal phenomena addressed. The second part is instead more focused on issues directly relevant for MAS design and shows how the formalisms proposed suit the representation of several key features of institutions and organizations which are relevant in that light. Related work will be discussed in every chapter.

1.2.1 Part I: "Terminologies, Contexts and Counts-as"

Part I consists of Chapters 2, 3 and 4 and develops the formal theory of contextual terminologies and counts-as statements.

To get more to the detail, Chapter 2 takes Pufendorf's intuition seriously and views institutions as the imposition of terminologies. From a formal point of view this translates to viewing the normative content of institutions as representable via description logic subsumption statements. However, every institution imposes a different terminology and this calls for the representation of subsumption statements in a contextual form. The chapter presents a formal framework to deal with contextual subsumptions.

Chapter 3 provides an in-depth analysis of the notion of context introduced by the framework of contextual terminologies. It can be understood in a broad sense as a contribution in context theory, where the notion of context investigated is the one emerging by conceiving institutions as what defines the context of reference for their own norms.

Concluding Part I, Chapter 4 pushes the theory of contextual terminologies further by taking it as a ground for a formal analysis of counts-as statements. The key thesis will be the following: institutions state terminologies and counts-as statements are statements talking about these terminologies, but they talk about terminologies in at least three different ways: a) by expressing what logically follows from a terminology; b) by expressing what a terminology adds with respect to

[6] For an introduction to modal logic and to description logic we refer the reader to [Blackburn et al., 2001] and, respectively, [Baader et al., 2002].

common-sense terminologies holding in general, i.e., what the terminology constitutes as new; c) by expressing the subsumption statements actually used as axioms for defining the terminology, i.e., the constitutive rules of the institution. The analysis of these three senses of counts-as will be accomplished in modal logic.

1.2.2 Part II: "Institutions and Organizations in MAS"

Part II consists of Chapters 5 and 6. It aims at showing how the theory developed in Part I bears interesting implications for understanding institutions and organizations as design metaphors for the development of MAS.

Chapter 5 shows how the formalization in description logic of institutions as terminologies can, in a very natural way, represent a number of institutional notions constantly occurring in the literature of MAS such as, in particular, the notion of role, the notion of infrastructure and the related problem of norm implementation. Roles will be studied as sets of norms: norms regulating how a role can be enacted and 'deacted', and norms conferring an institutional status to the agents enacting the role. Infrastructures will be thought of as what specifies the transitions that can take place between the agents interacting according to a given institution. This issue directly relates to the problem of the implementation of norms and the specification of appropriate sanctioning procedures guarding norm compliance.

Finally, Chapter 6 addresses the notion of organization. The study of this notion will be first of all linked to the one of institution. In fact, many structural dimensions within an organization refer to notions of a clearly institutional nature (e.g., power, control, etc.). What characterizes the notion of organization, and in particular its inherent graph-theoretical nature, will then be addressed yielding, as a result, a clear conceptual map of the two notions and of how they are related.

1.2.3 Important things left aside ... for the time being

A recurrent way to define institutions in the economic literature is to conceive of them as "the rules of the game" ([North, 1990]). While the work presented here provides an in-depth analysis of the "rules" side of institutions and organizations, it does not touch upon the "game" part of the story.

Game-theoretical aspects of institutions become relevant when tackling the problem of designing concrete institutions, that is, when choosing the set of rules a designer wants to be in force in a society. This problem can be roughly rephrased by the following questions: What games does the designer want the rules to define and how would those games be played? In other words, what are the 'good' rules given the designer's purposes and a certain society?

This opens up the issue of the 'game-theoretical meaning' of rules, issue that falls within the scope of many game-theoretical disciplines and research areas such as mechanism design and implementation theory ([Jackson, 2001]), and the formal verification of mechanisms ([Pauly and Wooldridge, 2003; Pauly, 2005a,b]). Even though tackling such issue is unavoidable for the development of a formal theory of institutions and organizations comprehending institutional and organizational design, the work presented here is concerned only with how to formally represent institutions (and organizations) in terms of rules, and to study them form a logical perspective. It offers thus an answer to the questions raised at the beginning of the introduction: “How can an institution be formally represented?” and “How can an organization be formally represented?”. On the contrary, it does not try to answer questions such as: "How can we formally represent the problem of assessing what institution (or organization) can best ensure desired global properties to hold in some relevant situations?". It is worth saying, however, that the present work does offer the ground for a viable bridge between the notion of rule as studied here, and the one of game as studied for instance in [van Benthem, 2002, 2005]. Such a bridge is the notion of labeled transition system, in terms of which institutions and organizations are analyzed in Part II of the thesis.

Part I

Terminologies, Contexts and Counts-as

Chapter 2

Contextual Terminologies

"From the point of view of its scientific universality, ethics does not talk about the single concrete case in which I have, as agent, to take a decision hic et nunc. It rather looked for [. . .] universal criteria which could be used, on the one hand, to read the ethic good or bad in the single cases [. . .], and on the other hand, to positively determine whether a practical decision is ethically correct or not. The analysis of each single case [...] can imply considerable difficulties; however, the fundamental thing is that everything should in the end depend on a simple subsumption."

E. Husserl, "Vorlesungen über Ethik und Wertlehre", 41

The opening quote from [Husserl, 1988] introduces the playground of the chapter. Ethics, just like any normative system, provides abstract criteria for the classification of good and bad behavior. Ethical agents are agents which comply with those criteria, by being able to apply them to concrete cases. Such application is grounded on logical subsumption.

Now, if we conceive of ethical systems in abstracto as normative systems, and of ethical agents —like the "I" in the quote— just as agents which comply with some given normative system, then the reach of the quote becomes evident for our purposes. If an electronic agent has to autonomously comply with what is stated by the norms of an eInstitution by correctly interpreting the concrete state it is acting in, or if the designer of an eInstitution (or of more eInstitutions) has to understand how to interpret a concrete system's states in terms of those norms, then the logic of this interpretation of concrete states in terms of abstract criteria needs to be formally understood.

The aim of the chapter is to provide a formal framework in which to represent and reason about "universal criteria" of different normative systems, and their interaction in the "analysis of the single cases" agents might be confronted with. In line with the above quote, this aim will be pursued in terms of one key semantic notion, the logical notion of "subsumption" which will be extended to a notion of contextual subsumption. The results presented in this chapter build on the work described in [Grossi et al., 2005c,b, 2006b,e] and provide part of our answer to the first research question.

The chapter proceeds as follows. Since we are interested in subsumptions, Section 2.1 starts by introducing the logics that are most concerned with that notion, namely description logics. In the same section two scenarios are also introduced which exemplify in detail the issues we are going to address, which concerns the representation of different and possibly inconsistent sets of subsumptions, otherwise called terminologies. The point is that different normative systems can categorize the same concepts in different ways. This is a pervasive phenomenon in legal systems and it is related to the so-called open-texture of legal concepts. On the basis of contributions from the legal theory literature, Section 2.1 introduces this notion and the related ones of 'core' and 'penumbra' of the meaning of a concept. The formal framework for representing contextual terminologies is then exposed in Section 2.2 and it is used to formalize the two scenarios in Section 2.3. On the ground of these results, section 2.4 provides a formal analysis of the two notions of core and penumbra and a formal characterization of the notion of open-texture. However, normative systems are not just terminologies, but they are terminologies which are supposed to provide guidelines for agents' behavior. Section 2.5 tackles this issue, which will run as a red thread through the whole thesis (see the third research question), and discusses how to fit deontic concepts in the picture of contextual terminologies. In Section 2.6 the results presented are put in perspective with related work on the notions of vagueness and defeasibility. Finally, in Section 2.7 we recapitulate the main results of the chapter.

2.1 Preliminaries

2.1.1 Terminological logics

Consider two universal sentences: the major premise of the famous example of Barbara syllogism "all men are mortal[1]", and its birth-related counterpart "all men have a mother". Their first-order logic formalizations run as follows:

$$\forall x(\mathtt{man}(x) \rightarrow \mathtt{mortal}(x)) \tag{2.1}$$

$$\forall x(\mathtt{man}(x) \rightarrow \exists y(\mathrm{has}(x, y) \wedge \mathtt{mother}(y))) \tag{2.2}$$

In standard logical analysis, sentences such as "all men are mortal" are split into subsentences in order to represent the interdependency between their truth value and the truth values of their subsentences, thus representing what their truth conditions are: for all substitution instances d of x, if `man` is true of d then `mortal` is true of d. As to "all men have a mother": for all substitution instances d of x, there exists

[1]If all men are mortal and all Greeks are men, then all Greeks are mortal ([Łukasiewicz, 1951]).

a substitution instance d' of y such that if $\texttt{man}$ is true of d, then has is true of (d, d') and $\texttt{mother}$ is true of d'.

Terminological logic —the first reference is, as far as we know, [Quine, 1960]— stresses a different aspect in the logical analysis of sentences, namely their descriptive or conceptual content. In the case of "all men are mortal" and "all men have a mother" a terminological logic analysis of the sentences would emphasize the structure of the asserted property:

$$\forall x(\neg\texttt{man} \sqcup \texttt{mortal})(x) \tag{2.3}$$

$$\forall x(\neg\texttt{man} \sqcup \exists\text{has}.\texttt{mother})(x) \tag{2.4}$$

that is: for all substitution instances d of x, $\neg\texttt{man} \sqcup \texttt{mortal}$ is true of d; for all substitution instances d of x, $\neg\texttt{man} \sqcup \exists\text{has}.\texttt{mother}$ is true of d. What happens in terminological analysis is that sentential complexity is hidden in the complexity of the predicated properties by means of "predicate functors", as they are called in [Quine, 1960, 1971].

The semantics of these functors is the one suggested by the symbolization: $\neg$ as complementation, $\sqcup$ as union, and the "crop functor[2]" $\exists$ as the operation that, applied to a binary relation, yields the set of elements d such that there exists at least one individual d' in such a relation with d. Once monadic predicates are interpreted as sets, and dyadic predicates as binary relations, it is then easy to see that Formulae 2.1 and 2.3 are equivalent as well as 2.2 and 2.4. By making a step further:

$$\neg\texttt{man} \sqcup \texttt{mortal} \equiv \top \tag{2.5}$$

$$\neg\texttt{man} \sqcup \exists\text{has}.\texttt{mother} \equiv \top \tag{2.6}$$

that is: $\neg\texttt{man} \sqcup \texttt{mortal}$ is true of all d; $\neg\texttt{man} \sqcup \exists\text{has}.\texttt{mother}$ is true of all d.

2.1.2 Description Logics

Formulae 2.5 and 2.6 are what we recognize today as Description Logic (DL) statements ([Baader et al., 2002]). It happened that in the last twenty-five years terminological logics have found a revival in the form of DL. Description logics are a whole spectrum of knowledge-representation languages which handle concept description expressions, which are endowed with the model-theoretic semantics we sketched above, and which usually enjoy attractive computational complexity. We will be using DL essentially as a terminological logic rather than a computational formalism, although we will briefly touch upon complexity issues in Section 5.4.

In DL elementary descriptions are atomic concepts (monadic predicates) and atomic roles (dyadic predicates) from which complex concept descriptions such as $\neg\texttt{man} \sqcup \texttt{mortal}$ and $\texttt{man} \sqcup \exists\text{has}.\texttt{mother}$ can be built. As an example we provide here the syntax and semantics of the probably most known DL, i.e., $\mathcal{ALC}$, where $\mathcal{AL}$ stands for Attributive Language and $\mathcal{C}$ for complement, indicating that negation of arbitrary concepts, and not only of atomic ones, is allowed. Given a set $\mathbf{A}$ of atomic

[2] The name comes from [Quine, 1960].

concepts A and a set $\mathbf{R}$ of atomic roles R, the set Γ of $\mathcal{ALC}$ concept descriptions γ is defined by the following BNF:

$$\gamma ::= A \mid \bot \mid \top \mid \neg\gamma \mid \gamma_1 \sqcap \gamma_2 \mid \forall R.\gamma.$$

The operator $\forall$ is the dual of the terminological logic cropping operator $\exists$. In DL they are called, respectively, universal and existential restriction operators. Expressions $\forall R.\gamma$ denote the set of elements d such that all elements d' that are in a relation R with them are instances of γ.

An $\mathcal{ALC}$ model is a structure:

$$m = \langle \Delta_m, \mathcal{I}_m \rangle$$

where:

- Δ_m is the non-empty domain of the model;
- $\mathcal{I}_m$ is a function $\mathcal{I}_m : \mathbf{A} \cup \mathbf{R} \longrightarrow \mathcal{P}(\Delta_m) \cup \mathcal{P}(\Delta_m \times \Delta_m)$, such that to every element of $\mathbf{A}$ and $\mathbf{R}$ an element of $\mathcal{P}(\Delta_m)$ and, respectively, of $\mathcal{P}(\Delta_m \times \Delta_m)$ is associated. This interpretation of atomic concepts and roles of $\mathcal{L}_i$ on Δ_m is then inductively extended as follows:

$$\begin{array}{rcl}
\mathcal{I}_m(\top) & = & \Delta_m \\
\mathcal{I}_m(\bot) & = & \emptyset \\
\mathcal{I}_m(\neg\gamma) & = & \Delta_m \setminus \mathcal{I}_m(\gamma) \\
\mathcal{I}_m(\gamma_1 \sqcap \gamma_2) & = & \mathcal{I}_m(\gamma_1) \cap \mathcal{I}_m(\gamma_2) \\
\mathcal{I}_m(\exists R.\gamma) & = & \{d \in \Delta_m \mid \exists d', (d, d') \in I_m(R) \Rightarrow d' \in I_m(\gamma)\} \\
\mathcal{I}_m(\forall R.\gamma) & = & \{d \in \Delta_m \mid \forall d', (d, d') \in I_m(R) \Rightarrow d' \in I_m(\gamma)\}
\end{array}$$

We denote the interpretation and its inductively defined extension both with $\mathcal{I}_m$. An $\mathcal{ALC}$ model m assigns a denotation to each atomic concept (for instance the set of elements of Δ_m that instantiate the concept `man`) and to each atomic role (for instance the set of pairs on Δ_m which are in a relation such that the first element is said to "have" the second element of the pair). Accordingly, meaning is given to each complex concept (e.g., $\neg$`man` $\sqcup$ `mortal` or $\exists$has.`mother`).

A model m is then said to be a model of a concept inclusion statement $\gamma_1 \sqsubseteq \gamma_2$ iff $\mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$. A concept definition $\gamma_1 \equiv \gamma_2$ corresponds in the obvious way to the two assertions $\gamma_1 \sqsubseteq \gamma_2$ and $\gamma_2 \sqsubseteq \gamma_1$.

Inclusion statements and definitions are, in a way, all what DLs are about since they express logical relationships between concept descriptions. A set of inclusion statements is usually called a *taxonomical box* (TBox) or *terminology*. In fact, DLs have been developed in Artificial Intelligence for the representation of terminologies or ontologies, i.e., sets of properties which are held as invariant on a domain (see [Baader et al., 2002]). And this is exactly the feature of the "universal criteria" stated by normative systems. Universal criteria are represented as subsumption statements, and sets of such criteria as terminologies.

It might be worth reminding that, from a logical point of view, many DLs can be viewed as notational variants of modal logics, or of computationally attractive fragments of first-order logic. DL $\mathcal{ALC}$ corresponds to the multi-modal system $\mathbf{K}_n$ ([Schild, 1991]) or, equivalently, to the binary fragment of first order logic admitting equality and fomulae with only one free variable ([Blackburn et al., 2001]).

2.1.3 Open-texture

Reasoning with normative (e.g., legal) concepts often means reasoning with underspecified or, more technically, *open-textured* conceptual apparatuses. In [Hart, 1961], open-textured concepts are conceived of as concepts whose application to certain borderline cases becomes indeterminate. In other words, subsumption under open-textured concepts is less transparent and that's why, to get back to the quote opening the chapter, "the analysis of each single case [...] can imply considerable difficulties". These open-textured related difficulties of terminological reasoning constitute the starting point of the chapter. We quote here an excerpt from [Hart, 1958] neatly exposing this type of problems.

> "[Suppose a] legal rule forbids you to take a vehicle into the public park. Plainly this forbids an automobile, but what about bicycles, roller skates, toy automobiles? What about airplanes? Are these, as we say, to be called "vehicles" for the purpose of the rule or not? If we are to communicate with each other at all, and if, as in the most elementary form of law, we are to express our intentions that a certain type of behavior be regulated by rules, then the general words we use like "vehicle" in the case I consider must have some standard instance in which no doubts are felt about its application. There must be a *core* of settled meaning, but there will be, as well, a *penumbra* of debatable cases in which words are neither obviously applicable nor obviously ruled out. [...] We may call the problems which arise outside the hard core of standard instances or settled meaning "problems of the penumbra"; they are always with us whether in relation to such trivial things as the regulation of the use of the public park or in relation to the multidimensional generalities of a constitution" ([Hart, 1958], p.607).

Given a general rule not allowing vehicles within public parks, there might be a municipality allowing bicycles in its parks and, in contrast, another one not allowing them in. What can be subsumed under the concept vehicle according to the first municipality? And what can be subsumed under the concept vehicle according to the second one?

The chapter introduces a framework in which to represent ontologies of different contexts and to reason about them both in isolation, i.e., within the contexts (intra-contextual reasoning), and in interaction, i.e., between contexts (inter-contextual reasoning). At the intra-contextual level a typical question would be of the form: given a set of subsumption relations holding in context C, does a given subsumption "A is a subconcept of B" also hold in context C? At an inter-contextual level, contexts are considered to be in some relation with one another. As an example, consider for instance the context D of a regional legislation as a concrete version of the context D of a national legislation. In this case, a typical question would be: given a set of subsumption relations holding in context C, and given that context D is more concrete than context C, does a given subsumption "A is a subconcept of B" hold in context D?

With such a framework in place it would be possible to represent the ontological aspect of the regulating activity of institutions in a formal way, and the terminologies of different institutions could then be formally represented and reasoned about.

2.1.4 Scenarios

We now depict two scenarios in order to state, in clear terms, the kind of reasoning patterns we are aiming to capture formally. They exemplify quite typical forms of contextual conceptualizations occurring in the normative domain. The first scenario deals with a rule establishing sufficient conditions for a person to be liable of violating the regulation concerning access to public parks in three different municipalities. The second scenario deals with the refinement of a definition of "vehicle" from the abstract context of a general regulation to more concrete contexts of municipal regulations. From a logical point of view, they display description logic forms of reasoning at the level of TBoxes.

Example 2.1. *(The public park scenario: "liability in parks") In the regulation governing access to public parks in region Reg it is stated that vehicles are not allowed within public parks and that: "persons using vehicles within public parks are liable for violating the regulation". In this regulation no mention is made of subconcepts of the concept vehicle, e.g., cars, bicycles, which may help in identifying an instance of vehicle, nor is it stated what it actually means to drive a vehicle: does the fact that I am wheeling my bicycle imply that I am driving it? In municipal regulations subordinated to this regional one, and therefore inheriting its global directives, specific subconcepts are instead handled. In municipality M1 and M2 the following rule holds: "persons driving bicycles within parks are liable of violating the regulation". In M3 instead, it does not hold that to drive a bicycle constitutes a violation. On the other hand, in all M1, M2 and M3 it holds that cars are not allowed in public parks. Moreover, in M2 it does not hold that "persons wheeling bicycles into public parks are liable for violating the regulation" despite liability arises in case bicycles happen to be driven. In M1 and M3 instead, to wheel a bicycle is considered a way of driving it.*

In this scenario the concept of `vehicle` gets various interpretations. Instances of `car` (w.r.t. the terminologies presupposed by M1, M2 and M3) are always instances of `vehicle`, while only in some contexts instances of `bicycle` are also instances of `vehicle`. What also gets various interpretations is the relation driving: somehow driving in M2 has a different meaning than in M1 and M3. Table 1 displays how liability is interpreted in three completely different ways by the contexts at issue, although in all contexts it holds that persons driving vehicles are to be considered liable. Note that context Reg cannot provide any qualification for actions such as

	DRIVE VEHICLE	DRIVE CAR	DRIVE BICYCLE	WHEEL BICYCLE
Reg	liable	*not classifiable*	*not classifiable*	*not classifiable*
M1	liable	liable	liable	liable
M2	liable	liable	liable	not liable
M3	liable	liable	not liable	not liable

Table 2.1: Liability in the public park scenario

driving or wheeling a bicycle simply because its language cannot express those notions.

Example 2.2. *(The public park scenario: "teenagers on skateboards") Consider again a regulation governing access to public parks in region Reg. Also in this regulation no mention is made of subconcepts of the concept vehicle. Nevertheless, a partial definition, specifying necessary conditions for something to be a vehicle, is stated: "vehicles are conveyances which transport persons or objects". In municipal regulations subordinated to this regional one subconcepts are instead introduced. This is done inheriting the definition stated at the Reg level and refining it either incrementing the number of necessary conditions for something to be considered a vehicle or stating sufficient ones. In municipality M1 the definition of vehicle is refined stating that "vehicles are self-propelled", that is, by adding necessary conditions for something to be considered a vehicle; and by stating that "self-propelled conveyances which transport persons or objects are vehicles", that is, by stating corresponding sufficient conditions. As a consequence, the concept vehicle is defined as "self-propelled conveyance which transport persons or objects are vehicles". In municipality M2, instead, the definition of vehicle is simply closed without any refinement: "conveyances which transport persons or objects are vehicles". Besides, in both M1 and M2, it holds that"skateboards are conveyances which are not self-propelled" and "teenagers are persons". These rules determine a different behavior of M1 and M2 with respect to concepts such as "skateboards transporting teenagers". With respect to this concept the following rule holds in M1: "skateboards transporting teenagers are not vehicles". In M2 instead, it holds that: "skateboards transporting teenagers are vehicles".*

The second scenario displays some other aspects of contextual conceptualizations. The concept of `vehicle` again gets various interpretations and it is first specified in its necessary conditions by context Reg and then completely defined in the two concrete contexts M1 and M2. The abstract regulation states that all vehicles are conveyances transporting persons or objects, leaving thus open the possibility for some of such conveyances not to be vehicles. This is the case of skateboards in M1 since M1 refines the abstract rule establishing more necessary conditions (being self-propelled) for conveyances to be classified as vehicles. Context M2 instead, simply closes the abstract rule through establishing that being a conveyance transporting persons or objects is sufficient for being a vehicle. Because of this, the two contexts M1 and M2 validate terminologies diverging on the conceptualization of the complex concept "skateboards transporting teenagers".

These two scenarios exemplify interesting nuances typical of complex context-dependent conceptualizations. They both represent instances of a typical form of contextual reasoning called "categorization" [Benerecetti et al., 2000], or "perspective" [Akman and Surav., 1996], that is, the form of reasoning according to which a same set of entities is conceptualized in many different ways. We will constantly refer back to them in the remainder of the work.

2.1.5 Contextualizing terminologies: ingredients

First, the framework we are looking for should support reasoning about the validity of terminological axioms with respect to contexts, thus giving a semantics to expressions of the type: "the concept `bicycle` is a subconcept of the concept `vehicle` in context M1".

The framework should be able to express the fact that concepts may be unclassifiable within specific contexts, that is, that given concepts can not be said to be subconcepts of any other concept: in the context Reg of the regional regulation, whether a person wheeling a bicycle within a public park is to be considered liable of violating the regulation can not be properly assessed since the concept at issue is not part of the language of the context Reg (see Table 2.1). In some sense, it corresponds to a subsumption which is evaluated with respect to the wrong context. Therefore, we want the framework to be able to express whether a concept gets meaning within a context: "concept `bicycle` is meaningful with respect to context M1". Completely analogous expressions should be available in order to handle a contextualization of role hierarchies such as: "role wheel (wheeling) is a subrole of drive (driving) in context M2" and "role wheel is meaningful in context M2". The key aspect is that context always come with a language (i.e., a logical and a descriptive alphabet), which sets the boundaries of what can be expressed in that context:

> "A general treatment of contexts may indeed wish to exempt contexts from the obligation to interpret every assertion" ([Shoham, 1991], p.400).

Secondly, the framework should provide a representation of context interplay. In particular, it should represent specific relations between contexts. Literature on context theory, such as [McCarthy, 1986; Benerecetti et al., 2000], considers an account of the logical relations holding between different contexts an essential ingredient for a fully-fledged theory of context. For instance, Examples 2.1 and 2.2 consider contexts (M1, M2 and M3) which are all specializations of a more abstract one (Reg) and which inherit the subsumptions holding in the abstract context, e.g. the rule according to which vehicles are not allowed in public parks. This suggests the consideration, for instance, of a generality relation between contexts ('context c_1 is at least as general as context c_2') and of an abstractness relation between contexts ('context c_1 is at most as abstract as context c_2).

In order to study context interplay also operations on contexts will be considered. In particular, we will introduce: a *contextual union* operator and a *contextual focus* operator. The first one joins two contexts yielding a more general one. The second one takes a context and yields another context which is specified on a sublanguage of the first one but which agrees with the first one on what can be expressed on that sublanguage. In other words, the operator prunes the information contained in the context to which it is applied, focusing only on what is expressible in the chosen sublanguage and abstracting from the rest. Finally, also *maximum* and *minimum* contexts will be introduced: these will represent the most general, and respectively, the most specific contexts on a language.

2.1.6 Context in logic: the *local model semantics*

Local model semantics (LMS, [Ghidini and Giunchiglia, 2001]) is a semantic framework for contextual reasoning based on first-order logic[3]. Its key intuition is to conceive of contexts as sets of first-order models, possibly on different languages. It is, to our knowledge, the only logic-based formal approach to contextual reasoning which models contexts as semantic entities. By means of viewing contexts as sets of models it is possible to represent 'local' reasoning. A formula ϕ holds in context i ($i \models \phi$) if and only if for all models m in the set of models c_i associated to i it holds that $m \models \phi$.

Therefore, reasoning within contexts —intra-contextual reasoning— amounts therefore, in LMS, to classical reasoning within a designated set of models. LMS provides also an account of inter-contextual reasoning, that is, of the semantic conditions for inference patterns to be sound, which link what holds in different contexts: 'if it is proven that formula ϕ_1 holds in context i then it can be inferred that formula ϕ_2 holds in context j',

$$\frac{i : \phi_1}{j : \phi_2}.$$

These inference rules are called *bridge rules* and, to say it with the authors:

> "[they] force contexts to agree up to a certain extent" ([Ghidini and Giunchiglia, 2001], p. 242).

The characterization of context interplay is therefore given in terms of contexts 'agreeing' about some information. At the semantic level, LMS treats this intuition by means of relations between sets of models, called *compatibility relations*. A compatibility relation CR is a set of sequences of sets of models of length n, i.e., a n-ary relation on contexts, with n at most countable. As a consequence, a context c_i is always interpreted as an element in a sequence. A CR satisfies a contextual assertion $i : \phi$ ($CR \models i : \phi$) if and only if for all sequences in CR the i^{th} element in the sequence satisfies ϕ.

[3]It might be interesting to note that this framework lies at the ground of the contextual version of OWL, the ontology web language, called C-OWL ([Bouquet et al., 2004]).

By imposing appropriate constraints on the definition of the compatibility relation it becomes possible to make the desired set of bridge rules sound. For instance, we can impose that the first context, denoted by 1, in each sequence is a superset of all the other contexts in the sequence. This guarantees that from ϕ holding in context 1, it can be inferred that ϕ holds also in any context i:

$$\frac{1 : \phi}{i : \phi}.$$

Although assuming the fundamental idea of LMS, we will aim at characterizing context interplay in a quite different way, as explained in the next section.

2.1.7 Contextualizing terminologies: recipe

The framework presented in the following sections is grounded on a precise answer to the two modeling questions we have to answer:

1. How to represent context in a formal way?
2. How to represent context interplay in a formal way?

The first question is tackled using the idea grounding the local models semantics: "contexts = sets of first-order models". However, since we are predominantly interested in concept subsumption, the underlying formal semantic framework will be the one of DL. This allows us to concentrate on simpler models than the ones considered in the general setting of local models semantics.

With respect to the second question, our line consists in pushing the intuition of LMS further in order to provide model-theoretical definitions of operations and relations on contexts. In this will reside the main difference with respect to the LMS. Instead of characterizing context interplay by means of compatibility relations, we will characterize it via operations that can be performed on contexts. This because, as stated in Section 2.1.5, we want to account for some intuitive context operations such as union and focus. This facilitates the view, which will be systematically investigated in Chapter 3, of contexts as elements of algebraic structures with precise formal properties. The relation between this view on context and the LMS's one will be thoroughly investigated in Section 3.2.6.

2.2 Formal Framework

Our proposal consists in mixing the semantics of DL with the idea, borrowed from LMS, of modeling contexts as sets of models. The resulting semantic framework will be able to represent and reason about sets of concept subsumptions, i.e., terminologies, in a contextual setting.

2.2.1 Language

The language $\mathcal{L}^{CT}$ (*language for contextual terminologies*) we are defining can be thought of as a meta-language for TBoxes defined on $\mathcal{AL}$ description logic languages. It consists of assertions about contexts, concepts and roles.

Assertions are built with the following symbols: one context relation symbol $\preccurlyeq$ (context . is at most as general as context .), two meaningfulness relation symbols "$\mathtt{Dnt}^c(. : .)$" (in context . concept . has a denotation) and "$\mathtt{Dnt}^r(. : .)$" (in context . role . has a denotation), and finally two contextual subsumption relation symbols "$. : . \sqsubseteq^c .$" (within context . concept . is a subconcept of concept .) and "$. : . \sqsubseteq^r .$" (within context . role . is a subrole of role .) for, respectively, concept and role subsumption[4].

Concept and role descriptions are built according to a global language $\mathcal{L}$ containing a non-empty at most countable set $\mathbf{A}$ of atomic concepts and an at most countable set $\mathbf{R}$ of atomic roles. Role descriptions can be built only via the unary operator $\overline{R}$ (role complement). Concept descriptions can be built via the zeroary operators $\bot$ (bottom concept) and $\top$ (top concept), the unary operator $\neg$ (complement), the binary operator $\sqcap$ (intersection) and the universal restriction operator $\forall$ which applies to role-concept pairs. This type of $\mathcal{AL}$ language is an $\mathcal{ALC}$ language extended with role complement and role hierarchies and will suffice to represent the scenarios introduced in Section 2.1. In the DL notation it is known as $\mathcal{ALCH}^{(\neg)}$.

Context descriptions are built from a non-empty at most countable set $\mathbf{c}$ of context identifiers and a number of context operators. These are indexed with the language $\mathcal{L}_i$ of $\mathcal{L}$ to which the operation they denote pertains. Each $\mathcal{L}_i$ contains therefore a non-empty at most countable set $\mathbf{A}_i \subseteq \mathbf{A}$ of atomic concepts and an at most countable set $\mathbf{R}_i \subseteq \mathbf{R}$ of atomic roles. The number n of sublanguages considered is assumed to be finite. Concept constructors are: two families of zeroary operators $\{\bot_i\}_{0 \leq i \leq n}$ (minimum contexts) and $\{\top_i\}_{0 \leq i \leq n}$ (maximum contexts), one family of unary operators $\{\mathrm{fcs}_i\}_{0 \leq i \leq n}$ (contextual focus operator), one family of binary operators $\{\curlyvee_i\}_{0 \leq i \leq n}$ (contexts disjunction operator).

The set Ξ of context descriptions (ξ) is defined through the following BNF:

$$\xi ::= c \mid \bot_i \mid \top_i \mid \mathrm{fcs}_i\ \xi \mid \xi_1 \curlyvee_i \xi_2.$$

Concept descriptions and role descriptions are defined in the standard DL way. The set P of roles descriptions (ρ) is defined through the following BNF:

$$\rho ::= R \mid \overline{\rho}.$$

The set Γ of concept descriptions (γ) is defined through the following BNF:

$$\gamma ::= A \mid \bot \mid \top \mid \neg\gamma \mid \gamma_1 \sqcap \gamma_2 \mid \forall\rho.\gamma.$$

[4]We use superscripts in the presentation of the language only in order to stress the distinction between meaningfulness of concepts or roles, and subsumptions of concepts or roles. In what follows, however, superscripts will be dropped since no confusion will arise.

Context descriptions:	$c \mid \bot_i \mid \top_i \mid \mathrm{fcs}_i\ \xi \mid \xi_1 \curlyvee_i \xi_2$
Role descriptions:	$R \mid \overline{\rho}$
Concept descriptions:	$A \mid \bot \mid \top \mid \neg\gamma \mid \gamma_1 \sqcap \gamma_2 \mid \forall\rho.\gamma$
Assertions:	$\mathtt{Dnt}^c(\xi : \gamma) \mid \mathtt{Dnt}^r(\xi : \rho) \mid \xi : \gamma_1 \sqsubseteq^c \gamma_2 \mid \xi : \rho_1 \sqsubseteq^r \rho_2 \mid \xi_1 \preccurlyeq \xi_2$

Table 2.2: Syntax of $\mathcal{L}^{CT}$

Concept union and existential restriction are defined respectively as: $\gamma_1 \sqcup \gamma_2 =_{\mathrm{def}} \neg(\neg\gamma_1 \sqcap \neg\gamma_2)$ and $\exists\rho.\gamma =_{\mathrm{def}} \neg(\forall\rho.\neg\gamma)$.
Finally, the set Σ of assertions (σ) is defined through the following BNF:

$$\sigma ::= \mathtt{Dnt}^c(\xi : \gamma) \mid \mathtt{Dnt}^r(\xi : \rho) \mid \xi : \gamma_1 \sqsubseteq^c \gamma_2 \mid \xi : \rho_1 \sqsubseteq^r \rho_2 \mid \xi_1 \preccurlyeq \xi_2.$$

Notice that no connectives for assertion compositions are enabled in $\mathcal{L}^{CT}$, even though boolean connectives can be obviously introduced. Like in DL, we are just interested in the validity of sets of atomic assertions: $\{\sigma_1, \ldots, \sigma_n\}$. In the case of contextual terminologies, these assertions are not subsumptions but contextual subsumptions, definiteness (Dnt) assertions of concepts and roles, and context generality relations ($\preccurlyeq$). In fact, sets of assertions σ can be thought of as sets of meta theoretical assertions about sets of standard DL subsumptions. The syntax just presented is recapitulated in Table 2.2.

2.2.2 Semantics

As exposed in the previous section, an $\mathcal{L}^{CT}$ language consists of four classes of expressions: Ξ (context constructs), P and Γ (role and concept descriptions), Σ (assertions). Semantics of P and Γ will be the standard description logic semantics of roles and concepts, on which our framework is based. Semantics for Ξ will be given in terms of model theoretic operations on sets of description logic models, and at that stage the semantics of assertions Σ will be defined via an appropriate satisfaction relation. The structures obtained, which we call *contextual terminology models* or ct-models, provide a formal semantics for $\mathcal{L}^{CT}$ languages.

The first step is then to provide the definition of a description logic model for $\mathcal{ALCH}^{(\neg)}$ languages. This extends the semantics of $\mathcal{ALC}$ (see Section 2.1.2) with a clause for the interpretation of role complement.

Definition 2.1. (Models for $\mathcal{L}_i$'s)
A model m for a language $\mathcal{L}_i$ is defined as follows:

$$m = \langle \Delta_m, \mathcal{I}_m \rangle$$

where:

- Δ_m *is the (non empty) domain of the model;*

- *$\mathcal{I}_m$ is a function $\mathcal{I}_m : \mathbf{A_i} \cup \mathbf{R_i} \longrightarrow \mathcal{P}(\Delta_m) \cup \mathcal{P}(\Delta_m \times \Delta_m)$, such that to every element of $\mathbf{A_i}$ and $\mathbf{R_i}$ an element of $\mathcal{P}(\Delta_m)$ and, respectively, of $\mathcal{P}(\Delta_m \times \Delta_m)$ is associated. This interpretation of atomic concepts and roles of $\mathcal{L}_i$ on Δ_m is then inductively extended as follows:*

$$
\begin{array}{rcl}
\mathcal{I}_m(\top) & = & \Delta_m \\
\mathcal{I}_m(\bot) & = & \emptyset \\
\mathcal{I}_m(\neg\gamma) & = & \Delta_m \setminus \mathcal{I}_m(\gamma) \\
\mathcal{I}_m(\gamma_1 \sqcap \gamma_2) & = & \mathcal{I}_m(\gamma_1) \cap \mathcal{I}_m(\gamma_2) \\
\mathcal{I}_m(\forall\rho.\gamma) & = & \{d \in \Delta_m \mid \forall d', (d, d') \in I_m(\rho) \Rightarrow d' \in I_m(\gamma)\} \\
\mathcal{I}_m(\overline{\rho}) & = & \Delta_m \times \Delta_m \setminus \mathcal{I}_m(\rho).
\end{array}
$$

We refer to the inductive extension of $\mathcal{I}_m$ still as $\mathcal{I}_m$. The Role complement operator is interpreted as the complementation on $\Delta_m \times \Delta_m$.

Language $\mathcal{ALCH}^{(\neg)}$ admits role hierarchies, that is, the expression of role inclusion statements. An $\mathcal{ALCH}^{(\neg)}$ model m is said to be a model of a role inclusion statement $\rho_1 \sqsubseteq \rho_2$ iff $\mathcal{I}_m(\rho_1) \subseteq \mathcal{I}_m(\rho_2)$. Role definition is defined in the obvious way.

2.2.3 Models for $\mathcal{L}^{CT}$

We can now define the structures that will work as *contextual terminology models* (ct-models) for language $\mathcal{L}^{CT}$.

Definition 2.2. (ct-models)
A ct-model $\mathbb{M}$ is a structure:

$$\mathbb{M} = \left\langle \{\mathbf{M_i^\Delta}\}_{\mathbf{0 \leq i \leq n}}, \mathbb{I} \right\rangle$$

where:

- *$\{\mathbf{M}_i^\Delta\}_{0 \leq i \leq n}$ is the family of all sets of models $\mathbf{M}_i^\Delta$ for languages $\mathcal{L}_i$ on a given domain Δ. That is, for all models m: $m \in \mathbf{M}_i^\Delta$ if and only if m is a model for $\mathcal{L}_i$ on Δ.*

- *$\mathbb{I}$ is a function $\mathbb{I} : \mathbf{c} \longrightarrow \mathcal{P}(\mathbf{M}_0^\Delta) \cup \ldots \cup \mathcal{P}(\mathbf{M}_n^\Delta)$.*

In other words, function $\mathbb{I}$ associates to each atomic context identifier in $\mathbf{c}$ a subset of the set of all models for some language $\mathcal{L}_i$ on domain Δ: $\mathbb{I}(c) = M$ with $M \subseteq \mathbf{M}_i$ for some i s.t. $0 \leq i \leq n$.

Notice that the domain of all models m is unique: $\forall m', m'' \in \bigcup_{0 \leq i \leq n} \mathbf{M}_i^\Delta$, $\Delta_{m'} = \Delta_{m''} = \Delta$. As explained in Section 2.1 we are interested in modeling different conceptualizations of a same set of entities. Notice besides that function $\mathbb{I}$ can be viewed as labeling sets of models on some language i via atomic context identifiers. In a way $\mathbb{I}$ fixes, for each atomic context identifier, the language on which the context denoted by the identifier is specified. We could say that it is $\mathbb{I}$ itself which fixes a specific language index for each atomic context identifier c.

2.2.4 Context focus

We model focus as a specific operation on sets of models which provides the semantic counterpart for the *contextual focus* operator introduced in $\mathcal{L}^{CT}$. Intuitively, focusing a context ξ on a language $\mathcal{L}_i$ yields a context consisting in that part of ξ which can be expressed in $\mathcal{L}_i$.

Let us first recall the notion of *domain restriction* ($\rceil$) of a function f w.r.t. a subset C of the domain of f. A domain restriction of a function f is the function $C\rceil f$ having C as domain and s.t. for each element of C, f and $C\rceil f$ return the same image: $C\rceil f = \{(x, f(x)) | x \in C\}$. The restriction $\rceil_i m$ of a model m to sublanguage $\mathcal{L}_i$ is defined as follows:

$$\rceil_i m = \langle \Delta_m, \mathbf{A}_i \cup \mathbf{R}_i \rceil \mathcal{I}_m \rangle$$

We can now introduce the operation of context reduction.

Definition 2.3. *(Context reduction:* $\mathtt{red}_i$*)*
Let M be a set of $\mathcal{ALCH}^{(\neg)}$ models on a domain Δ, then:

$$\mathtt{red}_i(M) = \{m \in \mathbf{M}_i^{\Delta} \mid \exists m' \in M : \rceil_i m' = m\}.$$

The following can be proven.

Proposition 2.1. *(Properties of context focus)*
Let M, M', M'' be sets of $\mathcal{ALCH}^{(\neg)}$ models on a domain Δ. Operation $\mathtt{red}_i$ is:

- *Idempotent:* $\mathtt{red}_i(\mathtt{red}_i(M)) = \mathtt{red}_i(M)$.
- *Normal:* $\mathtt{red}_i(\emptyset) = \emptyset$.
- *Additive:* $\mathtt{red}_i(M' \cup M'') = \mathtt{red}_i(M') \cup \mathtt{red}_i(M'')$.
- *Monotonic:* $M' \subseteq M'' \Rightarrow \mathtt{red}_i(M') \subseteq \mathtt{red}_i(M'')$.

Proof. [Idempotency, Normality] Idempotency and normality follow directly from Definition 2.3. [Additivity] Additivity is easily proven showing the following: $\mathtt{red}_i(M' \cup M'')$ is equal to $\{m \mid m = \langle \Delta_{m'}, \mathbf{A}_i \rceil \mathcal{I}_{m'} \rangle \ \&\ m' \in M' \cup M''\}$, which is in turn equal to $\{m \mid m = \langle \Delta_{m'}, \mathbf{A}_i \rceil \mathcal{I}_{m'} \rangle \ \&\ m' \in M' \text{ or } m' \in M''\}$ and therefore to $\mathtt{red}_i(M') \cup \mathtt{red}_i(M'')$. [Monotonicity] It follows from additivity. □

The operation of focus allows for shifting from richer to simpler languages and it is, as we would intuitively expect: idempotent (the nesting of focus is reducible), normal (focusing the empty context yields the empty context), additive (the focus of a context obtained via joining of two contexts can be obtained also joining the focuses of the two contexts), monotonic (if a context is less general then another one, the focus of the first is also less general than the focus of the second one). Notice also that operation $\mathtt{red}_i$ yields the empty set of models when it is applied to a context M' the language of which is not an expansion of $\mathcal{L}_i$. This is indeed intuitive: the context obtained via focus of the context "dinosaurs" on the language of, say, "gourmet cuisine" should be empty.

2.2.5 Operations on contexts

We are now in the position to give a semantics to context constructs. In Definition 2.2 atomic contexts c are interpreted as sets of models on some language $\mathcal{L}_i$ for $0 \leq i \leq n$: $\mathbb{I}(c) = M \in \mathcal{P}(\mathbf{M}_0^\Delta) \cup \ldots \cup \mathcal{P}(\mathbf{M}_n^\Delta)$. The semantics of context constructs Ξ can be defined via inductive extension of $\mathbb{I}$.

Definition 2.4. *(Semantics of context constructs)*
Let ξ, ξ_1, ξ_2 be context constructs, then $\mathbb{I}^*$ is inductively defined as follows:

$$\begin{array}{rcl} \mathbb{I}^*(c) & = & \mathbb{I}(c) \ \textit{if} \ c \in \mathbf{c} \\ \mathbb{I}^*(\mathrm{fcs}_i\ \xi) & = & \mathtt{red}_i(\mathbb{I}^*(\xi)) \\ \mathbb{I}^*(\bot_i) & = & \emptyset \\ \mathbb{I}^*(\top_i) & = & \mathbf{M}_i^\Delta \\ \mathbb{I}^*(\xi_1 \curlyvee_i \xi_2) & = & \mathtt{red}_i(\mathbb{I}^*(\xi_1) \ \cup \ \mathbb{I}^*(\xi_2)). \end{array}$$

In the following, we will often use the lighter notation $\mathbb{I}(\xi)$ instead of $\mathbb{I}^*(\xi)$.

The focus operator fcs_i is interpreted on the context reduction operation introduced in Definition 2.3, i.e., as the restriction of the interpretation of its argument to language $\mathcal{L}_i$. The $\bot_i$ context is interpreted as the empty context (the same on each language); the $\top_i$ context is interpreted as the greatest, or most general, context on $\mathcal{L}_i$; the binary $\curlyvee_i$-composition of contexts is interpreted as the restriction on $\mathcal{L}_i$ of the lowest upper bound of the interpretations of the two contexts. By additivity of $\mathtt{red}_i$ (Proposition 2.1) this is equivalent to the lowest upper bound of the restriction of the interpretations of the two contexts on $\mathcal{L}_i$. The satisfaction relation for assertions in ct-models is defined in the next section.

2.2.6 Assertions

Let us first recall that a partial function $f : A \cdots \longrightarrow B$ is defined for argument a ($f \downarrow a$) iff $\{x \mid x = f(a)\} \neq \emptyset$. Now, the interpretation function $\mathcal{I}_m$ of a model m on a sublanguage $\mathcal{L}_i$ of $\mathcal{L}$ can always be thought of as a partial interpretation function $\mathcal{I}_m^*$ on $\mathcal{L}$ s.t. if $\mathcal{I}_m^*$ is defined for concept γ then $\mathcal{I}_m(\gamma) = \mathcal{I}_m^*(\gamma)$.

The semantics of assertions can be defined as follows.

Definition 2.5. *(Semantics of assertions: $\models$)*
Let ξ, ξ_1, ξ_2 be a context constructs, $\gamma, \gamma_1, \gamma_2$ concept description, then:

$$\mathbb{M} \models \mathtt{Dnt}(\xi : \gamma) \ \textit{iff} \ \forall m \in \mathbb{I}(\xi) : \mathcal{I}_m \downarrow \gamma \tag{2.7}$$

$$\mathbb{M} \models \mathtt{Dnt}(\xi : \rho) \ \textit{iff} \ \forall m \in \mathbb{I}(\xi) : \mathcal{I}_m \downarrow \rho \tag{2.8}$$

$$\begin{aligned} \mathbb{M} \models \xi : \gamma_1 \sqsubseteq \gamma_2 \ \textit{iff} \ & \mathbb{M} \models \mathtt{Dnt}(\xi : \gamma_1), \mathbb{M} \models \mathtt{Dnt}(\xi : \gamma_2) \\ & \textit{and} \ \forall m \in \mathbb{I}(\xi) : \ \mathcal{I}_m^*(\gamma_1) \subseteq \mathcal{I}_m^*(\gamma_2) \end{aligned} \tag{2.9}$$

$$\begin{aligned} \mathbb{M} \models \xi : \rho_1 \sqsubseteq \rho_2 \ \textit{iff} \ & \mathbb{M} \models \mathtt{Dnt}(\xi : \rho_1), \mathbb{M} \models \mathtt{Dnt}(\xi : \rho_2) \\ & \textit{and} \ \forall m \in \mathbb{I}(\xi) : \ \mathcal{I}_m^*(\rho_1) \subseteq \mathcal{I}_m^*(\rho_2) \end{aligned} \tag{2.10}$$

$$\mathbb{M} \models \xi_1 \preccurlyeq \xi_2 \ \textit{iff} \ \mathbb{I}(\xi_1) \subseteq \mathbb{I}(\xi_2). \tag{2.11}$$

Logical consequence and satisfiability are defined in the natural way. An assertion σ is satisfiable if there exists a ct-model $\mathbb{M}$ s.t. $\mathbb{M} \models \sigma$. An assertion σ is the logical consequence of a set of assertions Σ ($\Sigma \models \sigma$) if for all ct-models $\mathbb{M}$, it holds that if $\mathbb{M}$ models all assertions in Σ then $\mathbb{M}$ models σ.

Clauses 2.7 and 2.8 specify when a concept (respectively, a role) is meaningful with respect to a context. This is the case when the set of denotations attributed to that concept (role) by the models in the context is not empty. If concept γ is not expressible in the language of context ξ, then concept γ gets no denotation in context ξ. This happens because concept γ does not belong to the domain of the interpretation functions. The same holds for roles. Notice that, since the models in the denotation of a context are all on the same language, for any concept γ (or role ρ) either all the interpretations are defined for γ (or ρ), or none.

On contextual subsumption

Clauses 2.9 and 2.10 deal with satisfaction of contextual subsumptions. A contextual concept subsumption relation $\xi : \gamma_1 \sqsubseteq \gamma_2$ holds iff concepts γ_1 and γ_2 are defined in the models constituting context ξ, i.e., they receive a denotation in those models, and all the description logic models constituting that context interpret γ_1 as a subconcept of γ_2.

Clauses 2.9 and 2.10 build on the clause for the validity of a subsumption relation in DL. The semantics of contextual subsumption relaxes the standard semantics of subsumption in two directions:

1. It evaluates the validity of the subsumption w.r.t. models of a language which is not necessarily the language in which the concepts occurring in the subsumption are formulated.

2. It evaluates the validity of the subsumption w.r.t. a subset of all models on that language.

Point 1 calls for the use of the partial versions of the evaluation functions $\mathcal{I}_m$'s together with the convergence assertion (Dnt) of the relevant concept descriptions. Point 2 calls instead for a quantification on the elements of the set denoted by the context expression.

As to Point 1, clauses 2.9 and 2.10 could be reformulated, by means of a relational notation for functions $\mathcal{I}_m$'s, as follows:

$$
\begin{array}{lll}
\mathbb{M} \models \xi : \gamma_1 \sqsubseteq \gamma_2 & \text{iff} & \mathbb{M} \models \mathtt{Dnt}(\xi : \gamma_1), \mathbb{M} \models \mathtt{Dnt}(\xi : \gamma_2) \\
 & & \text{and} \\
 & & \forall m \in \mathbb{I}(\xi), \forall D, D' \in \mathcal{P}(\Delta) : ((\gamma_1, D), (\gamma_2, D') \in \mathcal{I}_m \Rightarrow D \subseteq D').
\end{array}
$$

that is, γ_1 and γ_2 are defined in the models constituting context ξ and for all models m of context ξ and subsets D and D' of domain Δ, if D is assigned to γ_1 and D' is assigned to γ_2 then D is a subset of D'. Clause 2.10 can be similarly reformulated.

It is also important to notice the effect of clauses 2.9 and 2.10 in Definition 2.5 on making a contextual subsumption unsatisfied by a model $\mathbb{M}$. In fact, $\mathbb{M} \not\models \xi$:

$\gamma_1 \sqsubseteq \gamma_2$ iff either $\mathbb{M} \not\models \texttt{Dnt}(\xi : \gamma_1)$ and $\mathbb{M} \not\models \texttt{Dnt}(\xi : \gamma_1)$, or it is not the case that $\forall m \in \mathbb{I}(\xi) : \mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$. The same holds for contextual role subsumption. Leaving technicalities aside, this means that for a contextual subsumption not to be satisfied, either the concepts (roles) involved do not get denotations in the context, or those concepts (roles) are not in a relation of subsumption with respect to that context. Notice that this distinction is what marks the difference between the slots labeled with "*not classifiable*" and respectively "not liable" in Table 2.1.

On context generality

Finally, clause 2.11 gives a semantics to the $\preccurlyeq$ relation between context descriptions. Formula $\xi_1 \preccurlyeq \xi_2$ means that the context denoted by ξ_1 contains at most all the models that ξ_2 contains, that is to say, ξ_1 is *at most as general as* ξ_2.

The $\preccurlyeq$ relation between context constructs is thus interpreted as a standard subset relation. Note that this relation, being interpreted on the $\subseteq$ relation, is reflexive, antisymmetric and transitive. In [Grossi and Dignum, 2004] a generality ordering with similar properties was imposed on the set of context identifiers. The interesting point is that, here, such an ordering emerges here from the semantics. Note also that this relation holds only between contexts specified on the same language and, as a consequence, it is not total since contexts can be incomparable when containing models for different languages.

By means of a context generality relation $\preccurlyeq$ and the context focus operations fcs_i, it is possible to represent a notion of *context concretization* providing an answer to part of our second research question. Intuitively, a context ξ_1 on language $\mathcal{L}_i$ is at least as concrete as context ξ_2 on language $\mathcal{L}_j$ if and only if $\mathcal{L}_j$ is a sublanguage of $\mathcal{L}_i$ and the focus of ξ_1 on $\mathcal{L}_j$ is at most as general as ξ_2:

$$\mathrm{fcs}_j(\xi_1) \preccurlyeq \xi_2. \tag{2.12}$$

Intuitively, Formula 2.12 expresses that context ξ_1 deals with more concepts than context ξ_2, and it is less general about the interpretation of the concepts of ξ_2. We might equivalently say that context ξ_2 is an abstraction of context ξ_1: it handles less concepts and it denotes more models than the ones in the reduction of ξ_1. It is striking how this formalization gets close to informal analysis of the notion of generality between contexts:

> "Whether one of the two contexts human beings in general and human beings in conditions where influenza viruses are present is more general than the other depends on what we mean by each one. If the first context includes some information about people and the second context includes that information plus further information about viruses, then the former is more general. If the first includes all information about people and the second some subset of it that has to do with viruses, then the latter is more general. Otherwise, the two are noncomparable" ([Shoham, 1991], p.398).

This relation will come back in the formalization of the scenarios in the next section and will be further discussed in Chapter 3.

2.2.7 Simplifying the syntax of context descriptions?

Before getting to the next section it is worth explaining why we did not introduce a more standard syntax for context descriptions, for instance by making use of a binary boolean operator $\vee$ interpreted as set-theoretic union, a nullary operator $\perp$ interpreted as $\emptyset$, the focus operators for each sublanguage $\mathcal{L}_i$ and possibly also a unary context negation operator $\neg$ interpreted as set-theoretic complement. Such a syntax would yield a context descriptions algebra:

$$\left\langle \Xi', \vee, \neg, \perp, \{fcs_i\}_{0 \leq i \leq n} \right\rangle .$$

In effect, this syntax is not supported by the actual semantics, and it can not thus be simplified, since the elements of context description algebras would not necessarily denote contexts (i.e., sets of models on a same language) once interpreted by the corresponding extension $\mathbb{I}'$ of $\mathbb{I}$:

$$\begin{array}{rcl} \mathbb{I}'(c) & = & \mathbb{I}(c) \text{ if } c \in \mathbf{c} \\ \mathbb{I}'(fcs_i\ \xi) & = & \mathtt{red}_i(\mathbb{I}'(\xi)) \\ \mathbb{I}'(\neg\xi) & = & \emptyset \\ \mathbb{I}'(\xi_1 \vee \xi_2) & = & \mathbb{I}'(\xi_1) \cup \mathbb{I}'(\xi_2). \end{array}$$

For instance, $\mathbb{I}'(\xi_1 \vee \xi_2)$ does not denote a context if $\mathbb{I}(\xi_1)$ and $\mathbb{I}(\xi_2)$ are sets of models on different languages. In fact, operation $\mathtt{red}_i$ is applied, in this case, also to sets of models on different languages.

In other words, context description algebras yield more context descriptions than the ones allowed by the BNF definition of Ξ in $\mathcal{L}^{CT}$ languages (see Table 2.2), that is, $\Xi \subseteq \Xi'$. It is easy to see that all context descriptions of $\mathcal{L}^{CT}$ languages can be translated via a function $\mathtt{tr}$ to elements of a context description algebra:

$$\begin{array}{rcl} \mathtt{tr}(c) & = & c \text{ if } c \in \mathbf{c} \\ \mathtt{tr}(fcs_i(\xi)) & = & fcs_i(\mathtt{tr}(\xi)) \\ \mathtt{tr}(\xi_1 \vee_i \xi_2) & = & fcs_i(\mathtt{tr}(\xi_1) \vee \mathtt{tr}(\xi_2)) \\ \mathtt{tr}(\bot_i) & = & \perp \\ \mathtt{tr}(\top_i) & = & fcs_i(\neg\perp) \end{array}$$

Notice that $\mathtt{tr}$ is such that all context descriptions in $\mathcal{L}^{CT}$ languages are translated to expressions where the first occurring symbol is a focus operator, except for $\perp$. For those context descriptions, i.e. for contexts in Ξ, it is easy to see that $\mathbb{I}^*$ and $\mathbb{I}'$ agree on their interpretation.

Proposition 2.2. *($\mathbb{I}^*$ and $\mathbb{I}'$ coincide on Ξ)*
For every ξ in Ξ: $\mathbb{I}^*(\xi) = \mathbb{I}'(\mathtt{tr}(\xi))$.

Proof. This can be proven by induction on the complexity of the context descriptions. The base hold by definition of $\mathtt{tr}$. For the step, the induction hypothesis is: $\mathbb{I}^*(\xi) = \mathbb{I}'(\mathtt{tr}(\xi))$. The interesting case is $\mathbb{I}^*(\xi_1 \curlyvee_i \xi_2)$. By Proposition 2.1 this corresponds to: $\mathtt{red}_i(\mathbb{I}^*(\xi_1) \cup \mathbb{I}^*(\xi_2))$. By induction hypothesis and by the definition of $\mathbb{I}'$ this is equal to: $\mathbb{I}'(\mathtt{tr}(\xi_1 \curlyvee_i \xi_2))$. □

This shows that, somehow, context descriptions in $\mathcal{L}^{CT}$ (recall Table 2.2) restrict context description algebra in order for context descriptions ξ to denote only sets of models on the same language.

2.3 Contextual Terminologies at Work

This section introduces and discusses the formalization of Examples 2.1 and 2.2 in the framework of contextual terminologies.

2.3.1 Formalizing the first scenario

We proceed now to the formalization of Example 2.1.

Example 2.3. *(Sufficient conditions for "liability") To formalize the first scenario within our setting a language $\mathcal{L}$ is needed, which contains the following atomic concepts:* $\mathtt{person}$, $\mathtt{liable}$, $\mathtt{vehicle}$, $\mathtt{car}$, $\mathtt{bike}$; *and the following atomic roles:* drive *and* wheel. *Four atomic contexts are at issue here: the context of the main regulation Reg, let us call it c_{Reg}; and the contexts of the municipal regulations M1, M2 and M3, let us call them c_{M1}, c_{M2} and c_{M3} respectively. Let us call $\mathcal{L}_0$ and $\mathcal{L}_1$ the two languages at issue.*

To model the desired situation, our ct-model should at least satisfy the following $\mathcal{L}^{CT}$ formulae:

$$c_{M1} \curlyvee_0 c_{M2} \curlyvee_0 c_{M3} \preccurlyeq c_{Reg} \tag{2.13}$$
$$c_{Reg} : \mathtt{person} \sqcap \exists \mathrm{drive}.\mathtt{vehicle} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \tag{2.14}$$
$$c_{M1} \curlyvee_1 c_{M2} \curlyvee_1 c_{M3} : \mathtt{car} \sqsubseteq \mathtt{vehicle} \tag{2.15}$$
$$c_{M1} \curlyvee_1 c_{M2} : \mathtt{bike} \sqsubseteq \mathtt{vehicle} \tag{2.16}$$
$$c_{M3} : \mathtt{bike} \sqsubseteq \neg\mathtt{vehicle} \tag{2.17}$$
$$c_{M1} \curlyvee_1 c_{M3} : \mathrm{wheel} \sqsubseteq \mathrm{drive} \tag{2.18}$$
$$c_{M2} : \mathrm{wheel} \sqsubseteq \overline{\mathrm{drive}}. \tag{2.19}$$

Formula 2.13 is an instance of Formula 2.12. It states that the three contexts c_{M1}, c_{M2}, c_{M3} are concrete variants of context c_{Reg} by saying that the context obtained by joining the three concrete contexts on language $\mathcal{L}_0$ is at most as general as context c_{Reg}, that is: $\rceil_0\mathbb{I}(c_{M1}) \cup \rceil_0\mathbb{I}(c_{M2}) \cup \rceil_0\mathbb{I}(c_{M3}) \subseteq \mathbb{I}(c_{Reg})$ (see Section 2.2.2). As we will see in the following, this makes c_{M1}, c_{M2} and c_{M3} inherit what holds in c_{Reg}. Formula 2.14 formalizes the abstract rule to the effect that persons driving vehicles (within public parks) are liable for a violation of the applicable regulation. Formulas 2.15-2.17 describe the different taxonomies holding in the three concrete contexts at issue,

while formulas 2.18 and 2.19 describe the different role hierarchies holding in those contexts. To discuss in some more depth the proposed formalization, let us first list some easily proven logical consequences of formulae 2.13-2.19. We will focus on subsumptions contextualized to monadic contexts, that is to say, we will show what the consequences of formulae 2.13-2.19 are at the level of the four contexts c_{Reg}, c_{M1}, c_{M2} and c_{M3} considered in isolation.
Assertions concerning c_{Reg}:

$$\begin{array}{rcl}
2.14 & \models & \mathtt{Dnt}(c_{Reg} : \mathtt{person}) \\
2.14 & \models & \mathtt{Dnt}(c_{Reg} : \mathtt{vehicle}) \\
2.14 & \models & \mathtt{Dnt}(c_{Reg} : \mathtt{liable}) \\
2.14 & \models & \mathtt{Dnt}(c_{Reg} : \text{drive})
\end{array}$$

Assertions concerning c_{M1}:

$$\begin{array}{rcl}
2.13, 2.14 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \\
2.15 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{car} \sqsubseteq \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \\
2.16 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \\
2.18 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{wheel}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \exists \text{drive}.\mathtt{bike} \\
2.16, 2.18 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{wheel}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \\
2.14, 2.15 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{car} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \\
2.13, 2.14, 2.16 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \\
2.13, 2.14, 2.16, 2.18 & \models & c_{M1} : \mathtt{person} \sqcap \exists \text{wheel}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \\
2.15, 2.16, 2.18, 2.19 & \models & \mathtt{Dnt}(c_{M1} : \mathtt{car}) \\
2.15, 2.16, 2.18, 2.19 & \models & \mathtt{Dnt}(c_{M1} : \mathtt{bike}) \\
2.15, 2.16, 2.18, 2.19 & \models & \mathtt{Dnt}(c_{M1} : \text{wheel}) \\
2.13, 2.14 & \models & \mathtt{Dnt}(c_{M1} : \mathtt{person}) \\
2.13, 2.14 & \models & \mathtt{Dnt}(c_{M1} : \mathtt{vehicle}) \\
2.13, 2.14 & \models & \mathtt{Dnt}(c_{M1} : \mathtt{liable}) \\
2.13, 2.14 & \models & \mathtt{Dnt}(c_{M1} : \text{drive})
\end{array}$$

Assertions concerning c_{M2}:

$$\begin{array}{rcl}
2.13, 2.14 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \\
2.15 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{car} \sqsubseteq \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \\
2.16 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \exists \text{drive}.\mathtt{vehicle} \\
2.19 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{wheel}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \exists \overline{\text{drive}}.\mathtt{bike} \\
2.16, 2.19 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{wheel}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \exists \overline{\text{drive}}.\mathtt{vehicle} \\
2.14, 2.15 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{car} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable} \\
2.13, 2.14, 2.16 & \models & c_{M2} : \mathtt{person} \sqcap \exists \text{drive}.\mathtt{bike} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable}
\end{array}$$

$$\begin{aligned}
2.15, 2.16, 2.18, 2.19 &\models \mathtt{Dnt}(c_{M2} : \mathtt{car})\\
2.15, 2.16, 2.18, 2.19 &\models \mathtt{Dnt}(c_{M2} : \mathtt{bike})\\
2.15, 2.16, 2.18, 2.19 &\models \mathtt{Dnt}(c_{M2} : \mathtt{wheel})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M2} : \mathtt{person})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M2} : \mathtt{vehicle})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M2} : \mathtt{liable})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M2} : \mathtt{drive})
\end{aligned}$$

Assertions concerning c_{M3}:

$$\begin{aligned}
2.13, 2.14 &\models c_{M3} : \mathtt{person} \sqcap \exists\mathtt{drive.vehicle} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable}\\
2.15 &\models c_{M3} : \mathtt{person} \sqcap \exists\mathtt{drive.car} \sqsubseteq \mathtt{person} \sqcap \exists\mathtt{drive.vehicle}\\
2.17 &\models c_{M3} : \mathtt{person} \sqcap \exists\mathtt{drive.bike} \sqsubseteq \mathtt{person} \sqcap \exists\mathtt{drive.}\neg\mathtt{vehicle}\\
2.18 &\models c_{M3} : \mathtt{person} \sqcap \exists\mathtt{wheel.bike} \sqsubseteq \mathtt{person} \sqcap \exists\mathtt{drive.bike}\\
2.17, 2.18 &\models c_{M3} : \mathtt{person} \sqcap \exists\mathtt{wheel.bike} \sqsubseteq \mathtt{person} \sqcap \exists\mathtt{drive.}\neg\mathtt{vehicle}\\
2.14, 2.15 &\models c_{M3} : \mathtt{person} \sqcap \exists\mathtt{drive.car} \sqsubseteq \mathtt{person} \sqcap \mathtt{liable}
\end{aligned}$$

$$\begin{aligned}
2.15, 2.16, 2.18, 2.19 &\models \mathtt{Dnt}(c_{M3} : \mathtt{car})\\
2.15, 2.16, 2.18, 2.19 &\models \mathtt{Dnt}(c_{M3} : \mathtt{bike})\\
2.15, 2.16, 2.18, 2.19 &\models \mathtt{Dnt}(c_{M3} : \mathtt{wheel})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M3} : \mathtt{person})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M3} : \mathtt{vehicle})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M3} : \mathtt{liable})\\
2.13, 2.14 &\models \mathtt{Dnt}(c_{M3} : \mathtt{drive})
\end{aligned}$$

Proofs are omitted. The list displays four sets of formulae grouped on the basis of the context to which they pertain. Let us have a closer look to them.

The first group of formulae pertains the abstract context c_{Reg}. Only the abstract notions of person, liability, driving and vehicle are handled in that context, while the complex concepts which are central in the scenario (person driving a car, person driving a bicycle, person wheeling a bicycle) are not necessarily well-defined with respect to c_{Reg}. For instance, $\mathtt{Dnt}(c_{Reg} : \mathtt{car})$ is not a logical consequence of Formulae 2.13-2.19.

As to the consequences pertaining the three concrete contexts c_{M1}, c_{M2} and c_{M3}, notice first of all that all concepts get a denotation. Notice then that the first consequence of each group results from the generality relation expressed in 2.13, by means of which the content of 2.14 is shown to hold also in the three concrete contexts: in simple words, contexts c_{M1}, c_{M2} and c_{M3} inherit the general rule stating the liability of persons driving vehicles. Via this inherited rule, and via 2.15, it is shown that, in all contexts, who drives a car is also held liable.

As to cars and driving cars then, all contexts agree. Where differences arise is in relation with how the concept of bicycle and the role of wheeling are handled. In context c_{M1}, we have that it does not matter if somebody wheels or actually drives a bicycle, because in both cases this would count as driving a vehicle, and therefore as violating the regulation. In fact, in this context, a bicycle is a vehicle (2.16) and to wheel is a way of driving (2.17). Context c_{M2}, instead, expresses a different view. Since bicycles count as vehicles (2.16), to drive a bicycle is still a ground for liability. On the other hand, to wheel is actually classified as a way of refraining from driving (2.19), and therefore, there is no ground for considering persons wheeling bicycles to count as persons driving vehicles, and therefore to commit a violation. In fact, Formula $c_{M2} : \texttt{person} \sqcap \exists \text{wheel}.\texttt{bike} \sqsubseteq \texttt{person} \sqcap \texttt{liable}$ is not a logical consequence of Formulae 2.13-2.19. Context c_{M3} yields yet another terminology. Here bicycles are classified as objects which are not vehicles (2.17). Therefore, although to wheel is conceived as a way of driving (2.18), both to drive and to wheel a bicycle does not determine liability. None of $c_{M3} : \texttt{person} \sqcap \exists \text{drive}.\texttt{bike} \sqsubseteq \texttt{person} \sqcap \texttt{liable}$ and $c_{M3} : \texttt{person} \sqcap \exists \text{wheel}.\texttt{bike} \sqsubseteq \texttt{person} \sqcap \texttt{liable}$ is a logical consequence.

With respect to this, it is instructive to notice that even though both in c_{M2} and c_{M3} to wheel a bicycle is not a sufficient reason for being held liable, this holds for two different reasons: in c_{M2} because of 2.19, and in c_{M3} because of 2.17. Finally, in each of the concrete contexts, the concepts car and bike, and the role wheel all get a meaning (last consequence of each group). This illustrates how our framework is able to cope with some quite subtle nuances that characterize contextual terminologies.

2.3.2 A model of the scenario

In this section we expose a simple ct-model satisfying 2.13-2.19. Let us stipulate that the models m that constitute our interpretation of contexts identifiers consist of a domain $\Delta = \{a, b, c, d, e, f, g\}$. The contexts can be interpreted on two relevant languages (let us call them $\mathcal{L}_0$ and $\mathcal{L}_1$) s.t. $\mathbf{A}_0 = \{\texttt{person}, \texttt{liable}, \texttt{vehicle}\}$, $\mathbf{R}_0 = \{\text{drive}\}$ and $\mathbf{A}_1 = \{\texttt{person}, \texttt{liable}, \texttt{vehicle}, \texttt{car}, \texttt{bike}\}$, $\mathbf{R}_1 = \{\text{drive}, \text{wheel}\}$. That is to say, an abstract language concerning only persons, liability, vehicles and the action of driving, and a more detailed language concerning, besides liable persons, vehicles and driving, also cars, bicycles and the action of wheeling. Being $\mathcal{L}_0$ and $\mathcal{L}_1$ the two languages at issue and Δ the domain of the contexts, the domain of the ct-models is $\mathbf{M}_0^{\Delta} \cup \mathbf{M}_1^{\Delta}$. A ct-model would then be, for instance, a structure $\left\langle \mathbf{M}_0^{\Delta} \cup \mathbf{M}_1^{\Delta}, \mathbb{I} \right\rangle$ where $\mathbb{I}$ is such that:

- $\mathbb{I}(c_{M1}) = \{m_1, m_2\} \subseteq \mathbf{M}_1$ s.t. $\mathcal{I}_{m_1}(\texttt{person}) = \{e, f, g\}$, $\mathcal{I}_{m_1}(\texttt{vehicle}) = \{a, b, c, d\}$, $\mathcal{I}_{m_1}(\texttt{bike}) = \{a, b\}$, $\mathcal{I}_{m_1}(\texttt{car}) = \{c, d\}$, $\mathcal{I}_{m_1}(\text{drive}) = \{(e, a), (f, c)\}$, $\mathcal{I}_{m_1}(\text{wheel}) = \{(e, a)\}$, $\mathcal{I}_{m_1}(\texttt{liable}) = \{e, f\}$ and $\mathcal{I}_{m_2}$ agrees with $\mathcal{I}_{m_1}$ on the interpretation of person, bike, car, vehicle and $\mathcal{I}_{m_2}(\text{drive}) = \{(f, c), (g, d)\}$, $\mathcal{I}_{m_2}(\text{wheel}) = \{(g, d)\}$, $\mathcal{I}_{m_2}(\texttt{liable}) = \{f, g\}$.
- $\mathbb{I}(c_{M2}) = \{m_3, m_4\} \subseteq \mathbf{M}_1$ s.t. $\mathcal{I}_{m_3}$ and $\mathcal{I}_{m_4}$ agree with $\mathcal{I}_{m_1}$ on the interpretation of person, bike, car, vehicle and $\mathcal{I}_{m_3}(\text{drive}) = \{(f, d), (g, a)\}$, $\mathcal{I}_{m_3}(\text{wheel}) = \{(e, a)\}$, $\mathcal{I}_{m_3}(\texttt{liable}) = \{f, g\}$ and $\mathcal{I}_{m_4}(\text{drive}) = \{(e, c)\}$, $\mathcal{I}_{m_4}(\text{wheel}) = \{(f, a)\}$, $\mathcal{I}_{m_4}(\texttt{liable}) = \{e\}$.

- $\mathbb{I}(c_{M3}) = \{m_5\} \subseteq \mathbf{M}_1$ s.t. $\mathcal{I}_{m_5}$ agrees with $\mathcal{I}_{m_1}$ on the interpretation of `person`, `bike`, `car`. In addition, $\mathcal{I}_{m_5}(\texttt{vehicle}) = \{c, d\}$, $\mathcal{I}_{m_5}(\text{drive}) = \{(e, a), (f, c), (g, d)\}$, $\mathcal{I}_{m_1}(\text{wheel}) = \{(e, a)\}$, $\mathcal{I}_{m_1}(\texttt{liable}) = \{f, g\}$.
- $\mathbb{I}(c_{Reg}) = \{m \mid m = \langle \Delta_m, \mathbf{A}_0 \cup \mathbf{R}_0 \rceil \mathcal{I}_i \rangle$ and $1 \leq i \leq 5\}$, that is, c_{Reg} is interpreted by the model as the union of all models constituting c_{M1}, c_{M2} and c_{M3} restricted to the language $\mathcal{L}_0$.

The model makes an interesting feature of the semantics explicit. In contexts c_{M1} and c_{M2} the set of liable persons do not coincide in the two models composing the context; nevertheless only persons driving vehicles are indeed liable. This clearly shows that contexts can be viewed as clusters of possible situations all instantiating the same terminology. This intuition will be studied in modal logic in Chapter 4.

2.3.3 Formalizing the second scenario

The formalization of the scenario introduced in Example 2.2 follows.

Example 2.4. *(Categorizing "teenagers on skates") The global language $\mathcal{L}$ contains the following atomic concepts:* `conv`*,* `person`*,* `obj`*,* `vehicle`*,* `teenager`*,* `skate`*,* `self_prop`*; and the following atomic role:* transp. *Three are the atomic contexts at issue here: the context of the main regulation Reg, let us call it c_{Reg}; the contexts of the municipal regulations M1 and M2, let us call them c_{M1} and c_{M2} respectively. Let us call $\mathcal{L}_0$ and $\mathcal{L}_1$ the two languages at issue. To model the desired situation, a ct-model should then at least satisfy the following $\mathcal{L}^{CT}$ formulae:*

$$c_{M1} \curlyvee_0 c_{M2} \preccurlyeq c_{Reg} \tag{2.20}$$
$$c_{Reg} : \texttt{vehicle} \sqsubseteq \texttt{conv} \sqcap \forall \text{transp}.(\texttt{person} \sqcup \texttt{obj}) \tag{2.21}$$
$$c_{M1} : \texttt{vehicle} \sqsubseteq \texttt{self\_prop} \tag{2.22}$$
$$c_{M1} : \texttt{conv} \sqcap \exists \text{transp}.(\texttt{person} \sqcup \texttt{obj}) \sqcap \texttt{self\_prop} \sqsubseteq \texttt{vehicle} \tag{2.23}$$
$$c_{M2} : \texttt{conv} \sqcap \exists \text{transp}.(\texttt{person} \sqcup \texttt{obj}) \sqsubseteq \texttt{vehicle} \tag{2.24}$$
$$c_{M1} \curlyvee_1 c_{M2} : \texttt{teenager} \sqsubseteq \texttt{person} \tag{2.25}$$
$$c_{M1} \curlyvee_1 c_{M2} : \texttt{skate} \sqsubseteq \texttt{conv} \tag{2.26}$$
$$c_{M1} \curlyvee_1 c_{M2} : \texttt{skate} \sqsubseteq \neg\texttt{self\_prop} \tag{2.27}$$

We discuss the formalization of this scenario in fewer details than the previous one, stressing only the most important aspects. Formulae 2.20 is analogous to formula 2.13. Formula 2.21 represents the abstract constraints that context c_{Reg} imposes on the concept `vehicle`.

Formulae 2.22, 2.23 and 2.24 express the additional constraints on the concept `vehicle` holding in context c_{M1} and c_{M2} respectively: both contexts specify sufficient conditions and context c_{M1} adds also new necessary ones (2.22). Formulae 2.25 and 2.26 state the intuitive background knowledge common to the two concrete contexts. The point of the scenario consists in showing how teenagers on skateboards are conceptualized in the three contexts, that is to say: how are concept

$\mathtt{skate} \sqcap \exists\text{transp}.\mathtt{teenager}$ and concept $\mathtt{vehicle}$ related in each context? This can be easily shown via some relevant logical consequences of (2.20)-(2.27):

$$
\begin{array}{rcl}
2.20, 2.21, 2.22, 2.23 & \models & c_{M1} : \mathtt{conv} \sqcap \mathtt{self\_prop} \sqcap \\
 & & \exists\text{transp}.(\mathtt{person} \sqcup \mathtt{obj}) \equiv \mathtt{vehicle} \\
2.20, 2.21, 2.22, 2.23, 2.25, 2.26, 2.27 & \models & c_{M1} : \mathtt{skate} \sqcap \\
 & & \exists\text{transp}.\mathtt{teenager} \sqsubseteq \neg\mathtt{vehicle} \\
2.25, 2.26 & \models & \mathtt{Dnt}(c_{M1} : \mathtt{skate} \sqcap \exists\text{transp}.\mathtt{teenager}) \\
 & & \\
2.20, 2.21, 2.24 & \models & c_{M2} : \mathtt{conv} \sqcap \\
 & & \exists\text{transp}.(\mathtt{person} \sqcup \mathtt{obj}) \equiv \mathtt{vehicle} \\
2.20, 2.21, 2.24, 2.25, 2.26 & \models & c_{M2} : \mathtt{skate} \sqcap \exists\text{transp}.\mathtt{teenager} \sqsubseteq \mathtt{vehicle} \\
2.25, 2.26 & \models & \mathtt{Dnt}(c_{M2} : \mathtt{skate} \sqcap \exists\text{transp}.\mathtt{teenager}).
\end{array}
$$

Like in the previous example, the abstract context c_{Reg} does not necessarily categorize the concept at issue. In the two concrete contexts c_{M1} and c_{M2} two different definitions of $\mathtt{vehicle}$ hold, and therefore two different conceptualizations of the concept $\mathtt{skate} \sqcap \exists\text{transp}.\mathtt{teenager}$: since skateboards are not, in c_{M1}, self-propelled, they are not only non classifiable as vehicles, but, more strongly, they are actually classifiable as objects which are not vehicles. The contrary holds in c_{M2} since vehicle is there defined as conveyance transporting people or objects.

2.4 Intermezzo: Core and Penumbra

Let us go back to the quote from [Hart, 1958] by which we introduced the notion of open-texture. There, we read that for any legal (or, more broadly, normative) concept: "There must be a *core* of settled meaning, but there will be, as well, a *penumbra* of debatable cases in which words are neither obviously applicable nor obviously ruled out". What is the part of a denotation of a concept which remains context independent? What is the part which varies instead? The framework presented supports the formalization of the notions of core and penumbra —as described in [Hart, 1958]— in a natural way.

2.4.1 Core and penumbra formally defined

Before defining the notions of core and penumbra it is useful to introduce two more notions: the denotation of a concept in a context, and the range of a concept in a context.

Definition 2.6. *($\mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi)$)*

$$\mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi) =_{def} \{D \mid \exists m \in \mathbb{I}(\xi) : (\gamma, D) \in \mathcal{I}_m\}$$

where $\gamma \in \Gamma$, $\xi \in \Xi$ *and* $\mathbb{I}$ *is the interpretation function of* $\mathbb{M}$. *Notice that the definition is based on viewing interpretation functions* $\mathcal{I}_m$*'s as relations on* $\Gamma \times \mathcal{P}(\Delta)$.

Definition 2.6 makes explicit that, in the formal setting presented, concepts get multiple denotations within contexts. This feature corresponds to a very precise answer to the issue of the representation of vagueness. We will come back to this in more detail in Section 2.6.2.

Notice also that if concept γ is not expressible in the language of context ξ, then $\mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi) = \emptyset$, that is, concept γ gets no denotation at all in context ξ. This happens because concept γ does not belong to the domain of the interpretation functions[5]. Therefore, there exists no interpretation for that concept in the models of ξ. Formula 2.6 allows to capture the distinction between concepts which lack denotation ($\mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi) = \emptyset$), and concepts which have a denotation which is empty ($\mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi) = \{\emptyset\}$).

On the ground of the notion of denotation of a concept in a context, the notion of range of a concept in a context can be defined by means of the operation of union on a family of sets:

Definition 2.7. *(*$\mathfrak{Range}_{\mathbb{M}}(\gamma, \xi)$*)*

$$\mathfrak{Range}_{\mathbb{M}}(\gamma, \xi) =_{def} \bigcup \mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi)$$

where $\gamma \in \Gamma$, $\xi \in \Xi$ *and* $\mathbb{I}$ *is the interpretation function of* $\mathbb{M}$.

The notion of range defines the whole span of variation that the interpretation of a concept γ gets within a context ξ. An individual belongs to the range of a concept in a context iff it belongs to at least one denotation of that concept yielded by some interpretation in the context. Notice that notion of range flattens the distinction between concepts that do not get any denotation and concepts that get empty denotations in all models of the context. In both cases their range is just the empty set.

The notion of core of a concept in a context is then defined via the operation of intersection on a family of sets.

Definition 2.8. *(*$\mathfrak{Core}(\gamma, \xi_1, \xi_2)$*)*

$$\mathfrak{Core}_{\mathbb{M}}(\gamma, \xi) =_{def} \bigcap \mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi)$$

where $\gamma \in \Gamma$, $\xi \in \Xi$ *and* $\mathbb{I}$ *is the interpretation function of* $\mathbb{M}$.

Intuitively, the definition takes just the conjunction of the union of all the interpretations of γ in the context at issue. Let us get back to Example 2.1. Consider the ct-model $\mathbb{M} = \langle \mathbf{M}_0 \cup \mathbf{M}_1, \mathbb{I} \rangle$ exposed in Section 2.3.2. We have that $\mathfrak{Core}(\texttt{vehicle}, c_{M1} \curlyvee_1 c_{M2} \curlyvee_1 c_{M3}) = \{c, d\}$, that is, the core of the concept `vehicle` coincides, in that context, with the range of the concept `car` in that same context: $\mathfrak{Range}(\texttt{car}, c_{M1} \curlyvee_1 c_{M2} \curlyvee_1 c_{M3}) =$

[5]It might be instructive to note that if the context denoted by ξ is the empty context then $\mathfrak{Denotation}_{\mathbb{M}}(\gamma, \xi) = \emptyset$ for any concept γ.

$\{c, d\}$. Concept `car` in the three contexts c_{M1}, c_{M2} and c_{M3} lies in the core of concept `vehicle`.

The notion of penumbra is now easily definable.

Definition 2.9. *($\mathfrak{Penumbra}(\gamma, \xi_1, \xi_2)$)*
The "penumbra" of concept γ w.r.t. contexts ξ_1, ξ_2 on language $\mathcal{L}_i$ is defined as:

$$\mathfrak{Penumbra}_{\mathbb{M}}(\gamma, \xi) =_{def} \mathfrak{Range}_{\mathbb{M}}(\gamma, \xi) \setminus \mathfrak{Core}(\gamma, \xi).$$

where $\gamma \in \Gamma$, $\xi \in \Xi$ and $\mathbb{I}$ is the interpretation function of $\mathbb{M}$.

Thus, a "penumbral meaning" is nothing else but the set of individuals on which the contextual interpretation of the concept varies. Referring back again to Example 2.1: $\mathfrak{Penumbra}(\texttt{vehicle}, c_{M1} \curlyvee_1 c_{M2} \curlyvee_1 c_{M3}) = \{a, b\}$, that is to say, the penumbra of the concept `vehicle` ranges over those individuals that belong to at least one denotation of the concept `vehicle` yielded by some $\mathcal{I}_m \in \mathbb{I}(c_{M1} \curlyvee_1 c_{M2} \curlyvee_1 c_{M3})$ but which are not instances of the core of `vehicle`. Notice that the penumbra coincides in this case with the range of the concept bicycle itself: $\mathfrak{Range}(\texttt{bike}, c_{M1} \curlyvee_1 c_{M2} \curlyvee_1 c_{M3}) = \{a, b\}$. In the ct-model $\mathbb{M}$ of Section 2.3.2 we have thus that the range of car lies in the core of the concept vehicle, while the range of bicycle lies in the penumbra of that concept.

2.4.2 The modal flavor of core and penumbra

The notions of range, core and penumbra can be restated as follows:

$$x \in \mathfrak{Range}_{\mathbb{M}}(\gamma, \xi) \quad \text{iff} \quad \exists m \in \mathbb{I}(\xi) : x \in \mathcal{I}_m(\gamma) \tag{2.28}$$

$$x \in \mathfrak{Core}_{\mathbb{M}}(\gamma, \xi) \quad \text{iff} \quad \forall m \in \mathbb{I}(\xi) : x \in \mathcal{I}_m(\gamma) \tag{2.29}$$

$$x \in \mathfrak{Penumbra}_{\mathbb{M}}(\gamma, \xi) \quad \text{iff} \quad \exists m \in \mathbb{I}(\xi) : x \in \mathcal{I}_m(\gamma) \;\&\; \exists m \in \mathbb{I}(\xi) : x \notin \mathcal{I}_m(\gamma). \tag{2.30}$$

In other words: something lies in the range of γ iff it is a possible instance of γ according to the interpretations in $\mathbb{I}(\xi)$; something lies in the core of γ iff it is necessarily an instance of γ according to the interpretations in $\mathbb{I}(\xi)$; something lies in the penumbra of γ iff it is a contingent instance of γ according to the interpretations in $\mathbb{I}(\xi)$.

It is suggestive to have a look at the properties of these readings of range, core and penumbra. The following property is easily proven.

$$x \in \mathfrak{Core}_{\mathbb{M}}(\gamma_1 \sqcap \gamma_2, \xi) \quad \Leftrightarrow \quad x \in \mathfrak{Core}_{\mathbb{M}}(\gamma_1, \xi) \;\&\; x \in \mathfrak{Core}_{\mathbb{M}}(\gamma_2, \xi) \tag{2.31}$$

Formula 2.31 just states that x is always an instance of $\gamma_1 \sqcap \gamma_2$ according to ξ iff it is always an instance of γ_1 and always an instance of γ_2.

Notice also that it is not the case that if $x \in \mathfrak{Core}_{\mathbb{M}}(\gamma, \xi)$ then $x \in \mathcal{I}_m(\gamma)$ for any m. In other words, it is not the case that if x is necessarily an instance of γ according to ξ then x is an instance of γ for any interpretation $\mathcal{I}_m$ of γ. In fact, it might be the case that x is an instance of $\neg\gamma$ according to a different context ξ_1. It is also not the case that if $x \in \mathfrak{Core}_{\mathbb{M}}(\gamma, \xi)$ then $x \in \mathfrak{Range}_{\mathbb{M}}(\gamma, \xi)$, since ξ can be interpreted by $\mathbb{I}$ as the empty context.

Finally, suppose expressions $\mathfrak{Core}_{\mathbb{M}}(\gamma, \xi)$ and $\mathfrak{Range}_{\mathbb{M}}(\gamma, \xi)$ to be well-formed concept descriptions as well. This would not be so weird, since the interpretation of range, core, and penumbra is, like for any concept γ, a subset of the domain Δ on which the DL models in $\mathbb{M}$ are built. It would then be possible to nest those expressions and talk about, for instance, the core of the core of a concept γ in context ξ. It is so, however, that such nestings would be reducible:

$$x \in \mathfrak{Core}_{\mathbb{M}}(\gamma, \xi_1) \quad \Rightarrow \quad x \in \mathfrak{Core}_{\mathbb{M}}(\mathfrak{Core}_{\mathbb{M}}(\gamma, \xi_1), \xi_2) \tag{2.32}$$

$$x \in \mathfrak{Range}_{\mathbb{M}}(\gamma, \xi_1) \quad \Rightarrow \quad x \in \mathfrak{Core}_{\mathbb{M}}(\mathfrak{Range}_{\mathbb{M}}(\gamma, \xi_1), \xi_2) \tag{2.33}$$

Formulae 2.32 and 2.33 state that if x is always (or possibly) an instance of γ in ξ_1 then x lies also in the core of $\mathfrak{Core}_{\mathbb{M}}(\gamma, \xi_1)$ (or of $\mathfrak{Range}_{\mathbb{M}}(\gamma, \xi_1)$) in context ξ_2. To put it otherwise, if something lies in the range or the core of a concept in context ξ_1 then it does so independently of any other context, or more precisely, of any model m in $\mathbb{M}$. This is not surprising, since range core and penumbra denote, in effect, properties of elements of domain Δ which depend only on the labeling imposed by $\mathbb{I}$ on sets of models in $\mathbb{M}$. So, even though the interpretations of each element in Δ varies from context to context, range, core and penumbra are instead global notions in ct-models.

At this point it is suggestive to note: first, that Formula 2.31 clearly reminds of system **K**; second, that Formulae 2.31-2.33 are reminiscent of the modal axioms of multi-modal logic $\mathbf{K45}_n^{ij}$, which will be chosen as logic of context in the formal analysis of counts-as developed in Chapter 4[6]. We will see that, although the use of $\mathbf{K45}_n^{ij}$ will be independently motivated by isolating a suitable class of Kripke frames, the idea underlying the logic will be the same: contexts as sets of models. In this chapter the models at issue have been DL ones; in Chapter 4 contexts will be studied as sets of possible worlds, i.e., as sets of propositional models.

2.4.3 Arbitrary concepts vs. open-textured ones

The notion of core presented provides a formal handle on Hart's notion of open-texture as presented for instance in the quote reported in Section 2.1.3. It is important to notice that the existence of the core of a concept γ with respect to a non-empty context is what marks the difference between γ being arbitrary and γ being open-textured (with respect to that context). This draws a line between the often confused notions of arbitrariness and open-texture. Arbitrary concepts have empty core, while the interpretation of open-textured one is not allowed to vary indefinitely[7].

[6]See the axiomatics presented in Section 4.2.5. The axioms we are referring to are axioms 4^{ij} and 5^{ij}.

[7]It is worth noticing that open-texture has been studied by making use of algebraic methods also in [Lindahl and Odelstad, 2006, 2007], where open-textured concepts are called open intermediaries, that is, concepts that mediate the application of legal consequences to facts but whose application is to some extent open to courts' decisions. Although the formal background of this theory is radically different from ours, we do think there are several points of contact between the two theories which deserve future investigations.

This provides an answer to the second part of our second research question.

Let us also have a look at how a core can be constrained in contextual terminologies. This usually happens by subsumptions holding in an abstract context which are inherited by concretizations of the abstract context. We can single out the following pattern:

$$\begin{array}{rl} \mathrm{fcs}_i(\xi_1) \preccurlyeq \xi_2 & \text{“}\xi_1 \text{ is a concrete version of } \xi_2 \text{ on } \mathcal{L}_i\text{”} \\ \xi_2 : \gamma_1 \sqsubseteq \gamma_2 & \text{“In } \xi_2, \gamma_2 \text{ is a necessary condition of } \gamma_1\text{”} \end{array}$$

The first formula is precisely the sort of generality relation discussed in Section 2.2.6. The second one is just a contextual subsumption relation. In our semantics it follows that ξ_1 inherits that subsumption and, consequently, any interpretation of γ_1 in ξ_1 should be such that it is subsumed by the interpretation of γ_2. This means nothing but that the range of γ_1 in ξ_1 is contained in the range of γ_2. If γ_2 describes a concept whose interpretation is relatively stable (e.g., “self-propelled conveyance” in Example 2.2), then concept γ_1 results also to be clearly constrained in its possible interpretations (e.g., “vehicle” in Example 2.2).

This pattern is ubiquitous in institutional regulations where abstract pronouncements establish the boundaries for more concrete ones. We will come back to arbitrariness and open-texture in Chapter 5 while discussing abstract and concrete norms in institutions.

2.5 Terminologies and Deontics

Up to now, the focus has been on dealing exclusively with how concepts occurring in institutional regulations are structured to form the terminologies which institutions use in order to conceptualize the domains they are supposed to regulate. In this section we show how deontic notions such as obligation and permission can be expressed in the framework of contextual terminologies.

Deontic logic formalizes reasoning about “the distinction between what *ideally* is the case on the one hand, and what *actually* is the case on the other”([Jones and Sergot, 1992]). The standard deontic logic is the modal system **KD** in which $\Box\phi$ and $\Diamond\phi$ formulae are interpreted as, respectively, ϕ is obligatory (i.e. it holds in all “ideal” worlds) and ϕ is permitted (i.e. it holds in some “ideal” worlds). A plethora of alternative approaches to deontic notions is available in the literature[8].

Aim of this section is to discuss possible representations of basic deontic notions in DL and thus in $\mathcal{L}^{CT}$ languages. Two main options are introduced and discussed.

2.5.1 First option: a reduction strategy

The most straightforward way to deal with deontics in a $\mathcal{L}^{CT}$ language consists of developing a version of the reduction approach to deontic logic which was first proposed by Anderson in [Anderson, 1957, 1958] within a modal logic setting. The

[8]See [Lomuscio and Nute, 2004] and [Goble and Meyer, 2006] for recent contributions in this field.

reduction strategy is based on the intuition according to which the fact that something is obligatory means that its negation "necessarily" implies a violation (of the relevant set of norms or deontic constraints). The nature of the reduction lies in how this reference to a "necessity" is formally modeled. Various alternative reductions are studied in [d'Altan et al., 1993; Krabbendam and Meyer, 1999; Lomuscio and Sergot, 2003; Grossi et al., 2005d]. In Example 2.3 we have in fact already represented a norm with Formula 2.14:

$$c_{Reg} : \texttt{person} \sqcap \exists \texttt{drive}.\texttt{vehicle} \sqsubseteq \texttt{person} \sqcap \texttt{liable}$$

Intuitively, in context c_{Reg} it is necessarily the case that persons driving a vehicle are liable. The concept `liable` plays a deontic role exactly as the atom V (violation) in Anderson's reduction. In general, the application of a reduction strategy to our framework for representing norms would lead to a class of contextual subsumption statements defining, within each context, the notion of violation, of liability, or of related concepts. From this standpoint the regulative component of normative systems can be viewed as a specification of how each institution categorizes non-compliant objects. This option will be studied in detail also from a modal logic perspective in Section 4.7.3.

Along these lines, a yet easier way to represent deontics in a $\mathcal{L}^{CT}$ language from a reduction perspective would be to consider sorts of "ideal" counterparts of contexts. Given a context ξ providing classification of the concept of violation or liability, its ideal counterpart ξ_{ideal} would consist in the set of models $m \in \mathbb{I}(\xi)$ and such that $\mathcal{I}_m(\texttt{liable}) = \emptyset$. That is to say:

$$\xi_{ideal} : \texttt{liable} \equiv \bot.$$

This condition states that in such contexts the notion of violation corresponds to an inconsistency and therefore, whatever happens to be classified as yielding a violation is, in such contexts, a non existent entity. In other words, whatever holds in an ideal counterpart ξ_{ideal} can be consequently viewed as what ideally is the case. This idea will be studied from a modal logic perspective in relation with the notion of counts-as in Section 4.7.4.

2.5.2 Second option: "deontic roles"

Another option would consist in exploiting the correspondence between description logic and modal logic to which we already pointed in Section 2.1.1. Doing this, we would be able to represent standard deontic logic [von Wright, 1951], i.e., logic **KD** in a description logic fashion.

This can be achieved introducing a special role R_{ideal} denoting the relation between individuals and their "ideal" counterparts, in the exact same fashion in which Kripke semantics relates possible worlds in the standard interpretation of deontic logic. This special role needs then to be axiomatized as follows. For all i s.t. $1 \leq i \leq n$:

$$\top_i : \exists R_{ideal}.\top \equiv \top.$$

That is to say, for all top contexts on each language i, role R_{ideal} defines a serial relation between individuals.

Within such a setting, the obligation in context c_{Reg} for people driving vehicles not to enter public parks can be expressed as follows:

$$c_{Reg} : \texttt{person} \sqsubseteq \forall R_{ideal}.\neg(\exists \text{drive}.\texttt{vehicle}).$$

Literally, in to c_{Reg} if something is a person, then it is classified as something the ideal counterpart of which is always something which does not drive a vehicle.

2.5.3 Reduction or explicit deontics?

We have now two ways for expressing norms in a $\mathcal{L}^{CT}$ language. The main difference resides in the expressivity of the two solutions. Namely, the reduction approach sketched in Section 2.5.1 does not allow for expressing nested obligations: "it ought to be the case that it ought to be the case that persons do not drive vehicles in public parks". Obligations being classifications of a plain atom representing violation, cannot appear as a term in classifications themselves. In other words, we cannot express the fact that an obligation is obligatory for the simple reason that obligations are themselves classifications.

This can be done, instead, within the second approach making use of special serial roles. Special roles allow to encode a notion of obligation out of the classification itself at the syntactic level of concepts: the concept $\forall R_{ideal}.\neg\exists \text{drive}.\texttt{vehicle}$ denotes, literally, all individuals which should not be driving vehicles. Given this, the fact that such an obligation is obligatory can be easily represented nesting the application of the special role: $\forall R_{ideal}.(\forall R_{ideal}.\neg\exists \text{drive}.\texttt{vehicle})$. Intuitively, this latter concept denotes all individuals all the ideal counterparts of which are such that they ought not to drive vehicles. What is worth stressing, which neatly displays the core difference between the two variants, is their intuitive reading. Representing norms via reduction amounts to classify forbidden properties under a violation atom: what ought not to be the case is what happens to be subsumed under such a concept. Representing norms via special roles, instead, amounts to represent obligatory properties via complex concepts built up from those special roles: what ought to be the case is what holds for all "ideal" individuals.

The common feature of these two options should also be stressed. Whether represented via a reduction approach or via special roles, deontic notions are represented as subsumption statements, i.e. universal properties of DL models. This is aligned with the perspective already introduced by the quote opening the chapter: deontic notions as universal criteria. More on this issue will be discussed in Section 5.7.

2.6 Related Work

This section puts the notion of contextual terminology in perspective with work on vagueness and defeasibility. The section is closed with a remark about contextual terminologies and the notion of counts-as.

2.6.1 Contexts, fuzzy and rough sets

In this section we relate the present proposal with some more standard approaches to vagueness, namely approaches making use of fuzzy sets ([Wygralak, 1996]) or rough sets ([Pawlak, 1991]).

The most characteristic feature of our approach, with respect to fuzzy or rough set theories, consists of considering vagueness as an inherently semantic phenomenon. Vagueness arises from the referring of a language to structures modeling reality, and not from those structures themselves. That is to say, the truth denotation of a predicate is, in our approach, always definite and crisp, even if multiple. Consequently, no degree of membership is considered, as in fuzzy logic, and no representation of sets in terms of approximations is used, as in rough set theory.

Let us use a simple example in order to make this distinction evident. Consider the vague monadic predicate or, to use a description logic terminology, the concept `tall_person`. Fuzzy approaches would determine the denotation of this predicate as a fuzzy set, i.e., as the set of elements with membership degree contained in the interval $]0, 1]$. Standard rough set theory approaches would characterize this denotation not directly, but on the basis of a given partition of the universe (the set of all individuals) and a lower and upper approximation provided in terms of that partition. For instance, a trivial partition might be the one consisting of the following three concepts: `tall>2m`, `1.60m≤tall≤2m`, `tall<1.60m`. Concept `tall_person` would then be approximated by means of the lower approximation `tall>2m` (the elements of a set that are definitely also members of the to be approximated set), and the upper approximation `1.60m≤tall≤2m` $\sqcup$ `tall>2m` (the elements of a set that may be also members of the to be approximated set). In this rough set representation, set `1.60m≤tall≤2m` constitutes the so called *boundary* of `tall_person`. Within our approach instead, the set `tall_person` can be represented crisply and without approximations. The key feature is that `tall_person` obtains multiple crisp interpretations, at least one for each context: in the context of Dutch standards, concept `tall_person` does not subsume concept `1.60m≤tall≤2m`, whereas it does in the context of pygmy standards. According to our approach, vagueness resides then in the contextual nature of interpretation rather than in the concepts themselves.

It is nevertheless easy to spot some similarities, in particular with respect to rough set theory. The notions of "core" and "penumbra" have much in common with the notions of, respectively, *lower approximation* and *boundary* developed in rough set theory: each of these pairs of notions denotes what is always, and respectively, in some cases, an instance of a given concept. But the characterization of the last pair is based on a partition of the universe denoting the equivalence classes imposed by a set of given known properties. The notions of "core" and "penumbra", instead, are yielded by the consideration of many contextual interpretations of the same concept.

With respect to fuzzy approaches, notice that sets $\mathfrak{Core}$ can be viewed exactly as the sets of instances having a membership degree equal to one, while sets $\mathfrak{Penumbra}$

can be viewed as the sets of instances with degree of membership between zero and one. Furthermore, sets $\mathfrak{Penumbra}$ could be partitioned in sets X_n each containing instances that occur in a fixed number n of models constituting the "penumbra", thus determining a total and, notice, discrete ordering on membership: instances occurring in only one model in the "penumbra" will belong to the denotation of the concept at the minimum degree of membership, while instances occurring in the "core" at the maximum one.

2.6.2 Contexts and superevaluationism

The proposed analysis of contextual terminologies can be placed within those analyses of vagueness, developed in the area of philosophical logic, which distinguish between *de re* and *de dicto* views of vagueness ([Varzi, 2000]), the first holding that referents themselves are vague and therefore that vagueness constitutes something objective, whereas the second holding that it is the way referents are established that determines vagueness. Fuzzy set approaches lie within a *de re* conception of vagueness, while our approach is grounded on the alternative *de dicto* view. In philosophical logic, a formal theory has been developed which formalizes this *de dicto* approach to vagueness, the so called *superevaluationism* ([van Fraassen, 1966]). In this view, when interpreting vague terms, we consider the many possible ways in which those terms can be interpreted:

> "Whatever it is that we do to determine the 'intended' interpretation of our language determines not one interpretation but a range of interpretations. The range depends on context [...]" ([Lewis, 1999]).

As it is evident from Section 2.2.2, this intuition also backs our semantics. What our approach adds to formal accounts of *superevaluationism* such as [van Fraassen, 1966; Fine, 1975] consists in the explicit use of contexts as specific formal objects clustering the possible ways terms can be interpreted: contexts are exactly the range of interpretations which are admitted for the concepts at issue.

2.6.3 Contexts and defeasibility

The quote reproduced in Section 2.1.3, by means of which we introduced the issue of open-texture, goes on relating the "problems of the penumbra" with non-deductive forms of reasoning:

> "If a penumbra of uncertainty must surround all legal rules, then their application to specific cases in the penumbral area cannot be a matter of logical deduction, and so deductive reasoning, which for generations has been cherished as the very perfection of human reasoning, cannot serve as a model for what judges, or indeed anyone, should do in bringing particular cases under general rules. In this area men cannot live by deduction alone. And it follows that if legal arguments and legal decisions of penumbral questions are to be rational, their rationality

> must lie in something other than a logical relation to premises. So if it is rational or "sound" to argue and to decide that for the purposes of this rule an airplane is not a vehicle, this argument must be sound or rational without being logically conclusive" ([Hart, 1958], pp.607-608).

In fact, the "problems of the penumbra" have been extensively approached in logic especially from the perspective of the formalization of defeasible reasoning: the regional rule "all vehicles are banned from public parks" is defeated by the regulation of the first municipality stating that "all vehicles that are bicycles are allowed in the park" thus establishing an *exception* to the general directive. The formalization of norms via non-monotonic techniques (see [Prakken, 1997] for an overview) emphasizes the existence of exceptions to norms while understanding abstract terms in the standard way (all instances of bicycles are always vehicles). It has also been proposed to view subsumption rules as defaults: "normally, if something is a bicycle, then it is a vehicle" (for example [Royakkers and Dignum, 1997; Grossi and Dignum, 2004]).

These approaches, despite being effective in capturing reasoning patterns occurring for instance in Examples 2.1 and 2.2, are not adequate for analyzing problems related with the *meaning* of the terms that trigger those reasoning patterns. Those reasoning patterns can be viewed as defeasible because the meaning of the terms involved is not definite, it is open-textured and context dependent[9]. Statements such as "according to (in the context of) the public parks regulation of the first municipality bicycles are not vehicles, according to (in the context of) the public parks regulation of the second one bicycles are vehicles" have been interpreted as follows: "the subsumption of the concept `bicycle` under the concept `vehicle` holds in the context of the first municipality, but not in the context of the second one". A defeasible reasoning analysis leads to a quite different reading, which flattens the meaning of concepts and handles its variations by means of the notion of exception: "every non-exceptional instance of `bicycle` is an instance of `vehicle`". Bringing contexts into play allows instead for a neat characterization of the notions of "core" and "penumbra" of the meaning of a concept, a characterization which is not obtainable via the use of a notion of exception.

Contexts are in effect a viable tool for analyzing defeasibility. A context represents what is taken to be true —in our case subsumption statements— by drawing some conclusions, and defeasibility arises by shifting the context:

> "[...] what is commonly called nonmonotonic reasoning consists of two separate components. The first one is that on some grounds we may prefer to believe some assertions and reason with them as if they were true. The second component consists of changes in our preferences to believe these assertions due to newly discovered facts" ([Meyer and van der Hoek, 1991], p.399).

In [Meyer and van der Hoek, 1991], this intuition leads the authors to study defeasibility as the phenomenon arising by concluding something on the grounds of a set

[9] The issue of the relationship between contextuality and defeasibility has been raised also in [Akman and Surav., 1996].

of 'working beliefs', i.e., a context, which can then possibly be changed in order to draw different conclusions. They represent this insight in a multi-modal doxastic logic allowing for inconsistent belief sets, i.e., logic $\mathbf{K45_n}$. We will see that such logic bears essential similarities with the logic of context we will present in Chapter 4 and which has been anticipated in Section 2.4.2. To rephrase the title of [Meyer and van der Hoek, 1991], contextuality is thus a "monotonic means" for analyzing nonmonotonic reasoning. We will get back to [Meyer and van der Hoek, 1991] and logic $\mathbf{K45_n}$ in Section 4.7.2.

Before closing the section it is worth noting that also circumscription ([McCarthy, 1980]), probably the oldest model-theoretic approach to defeasibility, makes use of a notion of context in exactly the same sense introduced by LMS and used here: contexts as sets of models. In circumscription, context is what is obtained by an operation of minimization of the extension of given predicates applied to the set of models of a given theory.

2.6.4 Contextual terminologies and counts-as

One last remark is in order before concluding the chapter. It has not been by accident that the locution "counts as" has here and there appeared in the chapter. Contextual terminologies are strictly related to Searlean counts-as.

In fact, this chapter can be thought of as an attempt to provide a precise formal semantics to statements of the type "X counts as Y in context C". Counts-as statements can be viewed as what makes the terminology of an institution explicit. In this perspective, the semantics of counts-as statements is the one of contextual subsumption relations: $\xi : \gamma_1 \sqsubseteq \gamma_2$.

Chapter 4 pushes this view of counts-as further, showing that understanding counts-as statements as contextual subsumptions corresponds to only one view on what counts-as statements mean. In particular, counts-as statements will be studied also as context-defining subsumptions, that is, as those sets of subsumptions $\gamma_1 \sqsubseteq \gamma_2$ which define the context ξ to which they pertain as contextual subsumptions $\xi : \gamma_1 \sqsubseteq \gamma_2$. In fact, contexts represent the set of interpretations which a normative system assumes as possible for the concepts it deals with. In turn, what is assumed as a possible interpretation depends on the rules of the normative system, that is, on the set of $\gamma_1 \sqsubseteq \gamma_2$ by means of which it is specified. To put it in yet another way, normative systems consist of a set of rules —their constitutive rules— specifying their terminology. As such, they determine a set of interpretations, i.e., the set of models which satisfy its terminology. This set of models is what we call the context of the counts-as statements concerning that normative system. This thesis, which has just been sketched, will be systematically unraveled in Chapter 4.

2.7 Conclusions

By starting with a quote from [Husserl, 1988] we motivated the use of terminological logics for the analysis of normative systems. Then, by referring to the problem of open-texture, we pointed at the issue of contextuality in terminological analysis. This led us to devise a formal framework for representing contextual terminologies via a contextualized version of description logic semantics.

The framework has been used for formalizing two examples of "perspective" reasoning (Examples 2.3 and 2.4) in which a same set of entities is categorized in different ways. It has then been shown how to provide an articulate account of context interplay, which will be further investigated in the next chapter. Furthermore, the notions of "core" and "penumbra" of the meaning of a concept, which have been introduced at the beginning in relation with open-texture, have obtained a formal characterization (Section 2.4) in the framework. This has provided a rigorous way for distinguishing arbitrary concepts from open-textured ones: arbitrary concepts are vague concepts without a core, while open-textured ones are vague concepts endowed with a core (Section 2.4.3).

The upshot of these formal investigations for the theory of institutions have then been made explicit. In particular, it has been shown what the logical pattern is which institutions use to settle the boundaries of abstract notions in order to constrain their concrete interpretations (Section 2.4.3). Finally, deontic notions have also been addressed from this terminological perspective (Section 2.5).

Some words have been spent to put the exposed results in perspective with formal analysis of vagueness, such as rough sets, fuzzy sets and superevaluationism (Sections 2.6.1 and 2.6.2), and with formal analysis of defeasible reasoning (Section 2.6.3). Finally, it has been stressed that the results presented can be read as an attempt to give a first formal analysis of the semantics of counts-as statements, which will be further developed in modal logic in Chapter 4.

Chapter 3

Contexts as Algebraic Entities

"Tacere è la nostra virtù."

"We have a talent for being silent."

C. Pavese "I Mari del Sud", l.5

We investigate the implications that the framework exposed in the previous chapter bears for context theory. In particular, we want to make explicit what notion of context emerges from our analysis of contextual terminologies.

The chapter extends what was presented in [Grossi et al., 2005b]. The exposition is structured as follows. Section 3.1 provides some preliminary considerations and an overview of the content of the chapter. Section 3.2 shows how ct-models (Definition 2.2) can be fruitfully reduced to more familiar structures such as Boolean Algebras with Operators. This view of contexts is original and, to our knowledge, has never been advanced in the literature on context theory. This will also allow us to relate the view of contexts in contextual terminologies to the one maintained in the local model semantics (see Section 2.1.6). In Section 3.3 a family of modal logics is introduced, called release logics, and the contextual reasoning part of ct-models is shown to be an instance of reasoning in such logics. This provides the inter-contextual reasoning part of contextual terminologies with a modal proof theory. Some concluding remarks follow in Section 3.4.

3.1 Preliminaries

This section exposes the basic intuition of the chapter and sketches an overview of the results to be presented.

3.1.1 "The talent for being silent"

In [McCarthy, 1986] the statement about the need for addressing "contexts as abstract mathematical entities" ([McCarthy, 1986], p.1) was first put forth. The chapter will show that the analysis of context in contextual terminologies presupposes a view of context as *algebraic entities*, and more precisely, contexts as elements of Boolean Algebras with Operators ([Jónsson and Tarski, 1951, 1952]). This result will allow us to compare in some more detail our work with the local model semantics for contextual reasoning which was briefly exposed in Section 2.1.6.

Key intuition of the chapter is to consider the language of each context as what makes the context be indifferent about something, rather than what makes the context be able to express something:

> "The expressive power [...] determines not so much what can be said, but what can be left unsaid" ([Levesque and Brachman, 1987], p.82).

Here, with expressive power we do not mean the logical expressive power, i.e., what operators are available in the language, but the properly descriptive one, i.e., the non-logical alphabet of the language (the set of atomic concepts and roles). The intuition backing the chapter is that, by choosing the non-logical alphabet of a language, we give to the contexts in that language —to use the poetic expression from [Pavese, 1936]— the "talent for being silent" about some issues.

For the design of institutions, this 'talent' is essential since the norms of institutions are expressed in languages which are often selected with the precise aim of disregarding irrelevant aspects. The chapter provides insights on the logical properties of this linguistic selection lying at the ground of norm formulation.

3.1.2 Overview of the content

The chapter shows how to transform ct-models into special kinds of algebras of sets of DL models, which we call *context algebras* (Definition 3.4), preserving the validity of concept subsumptions (Proposition 3.8). It is then shown that these algebras can be thought of as models of a special kind of modal logic (Corollary 3.1), thus grounding the conception of contexts (in contextual terminologies) as modal notions.

We will first introduce a notion of equivalence between models with respect to a sublanguage (*sublanguage equivalence*). Intuitively, two DL models are said to be equivalent up to what can be expressed in a language $\mathcal{L}_i$ if they agree on all subsumptions between concepts expressible in that language. The operation $\mathtt{red}_i$ of context reduction (Definition 2.3) can be 'emulated' by means of this notion of sublanguage equivalence (Proposition 3.5). Roughly speaking, the intuition consists in the following: given a set of models X in language $\mathcal{L}$, the set of models $\mathtt{red}_i(X)$ corresponds to the set of models on $\mathcal{L}$ which are $\mathcal{L}_i$-equivalent with the models in X. By exploiting this result, ct-models can be given up in favor of more standard structures, which we call context algebras.

3.2 On Contexts in Contextual Terminologies

This section shows that ct-models, as they have been introduced in Definition 2.2, can be represented as structures of a more familiar kind, by making use of some simple operations on sets of models. The results presented in the chapter are stated for ct-models for terminologies expressed on $\mathcal{ALC}$[1]. Different DL languages could have been used as well.

3.2.1 Equivalent $\mathcal{ALC}$ models

Two models m and m' for an $\mathcal{ALC}$ language $\mathcal{L}$ are equivalent if they satisfy the same set of inclusion axioms between any two $\mathcal{ALC}$ concepts γ_1 and γ_2 expressible in $\mathcal{L}$: $\mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$ iff $\mathcal{I}_{m'}(\gamma_1) \subseteq \mathcal{I}_{m'}(\gamma_2)$. If m and m' are equivalent ($m \sim m'$) then they cannot be distinguished by an $\mathcal{ALC}$ TBox since they satisfy the very same terminology.

This notion of equivalence has to do with the indistinguishability of two $\mathcal{ALC}$ models as far as the logical constructs, or grammar, of a certain language are concerned. What if we take into consideration also the strictly descriptive alphabet of the language, i.e., the non-logical part of it? Two models, which are not equivalent with respect to a given descriptive alphabet (a given set of atomic concepts and roles), may become equivalent if only a sub-alphabet (a subset of atomic concepts and roles) is considered.

To get this idea formal let us first have a brief look at the way a set of sublanguages of a given $\mathcal{ALC}$ language is structured. Let $\mathcal{L}$ be an $\mathcal{ALC}$ at most countable language, and let us denote with $\mathcal{L}_i$ any of its sublanguages, i.e., languages defined on a set of atomic concepts $\mathbf{A_i} \subseteq \mathbf{A}$ and on a set of atomic roles $\mathbf{R_i} \subseteq \mathbf{R}$ where $\mathbf{A}$ and $\mathbf{R}$ are the set of atomic concepts and, respectively, roles of $\mathcal{L}$. Now let $\mathfrak{Sub}(\mathcal{L})$ be the set of all the $\mathcal{ALC}$ sublanguages $\mathcal{L}_i$ of $\mathcal{L}$[2]. If we allow both $\mathbf{A_i}$ and $\mathbf{R_i}$ for any i to be possibly empty, it is immediately clear that the structure $\langle \mathfrak{Sub}(\mathcal{L}), \cup, -, \mathcal{L}, \emptyset \rangle$ is a set algebra and therefore a Boolean Algebra. Leaving technicalities aside, this just means that by choosing a descriptive alphabet, a set of sublanguages is consequently chosen which is structured according to a Boolean Algebra. This observation takes us a step further toward the notion of sublanguage equivalence.

Definition 3.1. *(Sublanguage equivalence)*
Two models m and m' for an $\mathcal{ALC}$ language $\mathcal{L}$ are equivalent w.r.t. sublanguage $\mathcal{L}_i$ if they satisfy the same set of inclusion statements between $\mathcal{ALC}$ concepts expressible using the alphabet of $\mathcal{L}_i$. For any $\gamma_1, \gamma_2 \in \mathcal{L}_i$: $\mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$ iff $\mathcal{I}_{m'}(\gamma_1) \subseteq \mathcal{I}_{m'}(\gamma_2)$. If m and m' are equivalent w.r.t. $\mathcal{L}_i$ ($m \sim_i m'$) then they cannot be distinguished by an $\mathcal{ALC}$ TBox expressed on $\mathcal{L}_i$.

[1]Language $\mathcal{ALC}$ is the 'standard' DL language and it has been presented in Section 2.1.2.

[2]Obviously, if $\mathcal{L}$ is infinite then $|\mathfrak{Sub}(\mathcal{L})| > \aleph_0$. For our purposes, as the example discussed in Chapter 2 show, we are typically interested in finite languages.

The definition makes precise the idea of two $\mathcal{ALC}$ models agreeing up to what is expressible on a given alphabet. To put it another way, it formalizes the idea that two $\mathcal{ALC}$ models m and m' are equivalent modulo the alphabet in the complement $-\mathcal{L}_i$ (i.e., $\mathcal{L}\backslash\mathcal{L}_i$) of the sublanguage considered: m is indistinguishable from m' if we disregard the descriptive alphabet in $-\mathcal{L}_i$. Notice that if $m \sim_i m'$ and $\mathcal{L}_i = \mathcal{L}$ (i.e., the maximal element in $\mathfrak{Sub}(\mathcal{L})$) then $\sim_i = \sim$, that is, $\sim_i$ is the standard equivalence between $\mathcal{ALC}$ models.

Proposition 3.1. *(Properties of $\sim_i$)*
Let m and m' be two models for the $\mathcal{ALC}$ language $\mathcal{L}$ on domain Δ. The following holds:

1. *For every sublanguage $\mathcal{L}_i$ of $\mathcal{L}$, relation $\sim_i$ is an equivalence relation on the set of all models $\mathbf{M}^\Delta$ of language $\mathcal{L}$ on domain Δ.*
2. *For all sublanguages $\mathcal{L}_i$ and $\mathcal{L}_j$ of $\mathcal{L}$: if $\mathcal{L}_i \subseteq \mathcal{L}_j$ then $\sim_j \subseteq \sim_i$. It follows that for every sublanguage $\mathcal{L}_i$ of $\mathcal{L}$: $\sim \subseteq \sim_i$, that is, standard equivalence implies sublanguage equivalence.*

Proof. Claim (1) is straightforwardly proven. It is easy to see that: identity is a subrelation of $\sim_i$ for any sublanguage $\mathcal{L}_i$; and that $\sim_i \circ \sim_i$ and $\sim_i^{-1}$ are subrelations of $\sim_i$ for any sublanguage $\mathcal{L}_i$. Claim (2) is proven by considering that, if $\mathcal{L}_i$ is a sublanguage of $\mathcal{L}_j$ and $m \sim_j m'$, then for all concepts $\gamma_1, \gamma_2 \in \mathcal{L}_i$: $\mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$ iff $\mathcal{I}_{m'}(\gamma_1) \subseteq \mathcal{I}_{m'}(\gamma_2)$. Hence, $m \sim_i m'$. □

Sublanguage equivalence can therefore be used to cluster the set of models $\mathbf{M}^\Delta$ of an $\mathcal{ALC}$ language $\mathcal{L}$.

Definition 3.2. *(Clustering sublanguage-equivalent models)*
Given a set of models $X \subseteq \mathbf{M}^\Delta$ in a language $\mathcal{L}$ and a language $\mathcal{L}_i \subseteq \mathcal{L}$, the cluster of models in $\mathcal{L}$ which are $\mathcal{L}_i$-equivalent to the models in X with respect to sublanguage $\mathcal{L}_i$ is defined as follows:

$$\mathsf{eq}_i(X) = \{m \in \mathbf{M}^\Delta \mid \exists m' \in X : m \sim_i m'\}.$$

Notice that if $\mathcal{L}_i = \mathcal{L}$ then $\mathsf{eq}_i(X)$ yields a closure of X under $\mathcal{ALC}$ standard equivalence, i.e., the set of m's which are equivalent with at least an element of X. By omitting the index in $\mathsf{eq}_i(X)$ we indicate that the language on which the operation eq_i is applied is the same language of the models in X, i.e., $\mathcal{L}_i$.

In practice, function eq_i treats all the elements of its argument as representatives of corresponding clusters of $\mathcal{L}_i$-equivalent models and joins these clusters together.

Proposition 3.2. *(Properties of eq_i)*
Operation eq_i enjoys the following properties for any $X \subseteq \mathbf{M}^\Delta$ in a language $\mathcal{L}$ and sublanguages $\mathcal{L}_i$ and $\mathcal{L}_j$ of $\mathcal{L}$:

- *Inclusion (reflexivity): $X \subseteq \mathsf{eq}_i(X)$.*
- *Additivity: $\mathsf{eq}_i(X \cup Y) = \mathsf{eq}_i(X) \cup \mathsf{eq}_i(Y)$.*
- *Normality: $\mathsf{eq}_i(\emptyset) = \emptyset$.*

- *Idempotency:* $\mathtt{eq}_i(\mathtt{eq}_i(X)) = \mathtt{eq}_i(X)$.
- *If* $\mathcal{L}_i \subseteq \mathcal{L}_j$ *then* $\mathtt{eq}_j(X) \subseteq \mathtt{eq}_i(X)$*. Hence:* $\mathtt{eq}_i(\mathtt{eq}(X)) = \mathtt{eq}_i(X)$.

Proof. Follows from Proposition 3.1 and Definition 3.2. □

Besides, it follows from additivity that $\mathtt{eq}_i$ is monotonic ($X \subseteq Y \Rightarrow \mathtt{eq}_i(X) \subseteq \mathtt{eq}_i(Y)$); from inclusion and idempotency that $\mathtt{eq}_i$ enjoys cumulative transitivity ($X \subseteq Y \subseteq \mathtt{eq}_i(X) \Rightarrow \mathtt{eq}_i(Y) \subseteq \mathtt{eq}_i(X)$). Since $\mathtt{eq}_i$ is reflexive, idempotent and monotonic, it is thus a closure operation. Finally, a consequence of reflexivity is that, if $\mathcal{L}$ is the maximal language considered and $\mathbf{M}^{\Delta}$ the set of its models, then $\mathtt{eq}_i(\mathbf{M}^{\Delta}) = \mathbf{M}^{\Delta}$.

It is important to notice that these properties are exactly the properties proven to hold for the operation $\mathtt{red}_i$ in Proposition 2.1 plus inclusion, which does not hold for $\mathtt{red}_i$. Inclusion points precisely to the essential difference between $\mathtt{eq}_i$ and $\mathtt{red}_i$ applied to a set of models X: while the first one yields a set of models in the same language, the second one yields a set of models in a sublanguage of the language of the models in X whose intersection with X is therefore trivially empty. Central aim of the chapter is precisely to emulate $\mathtt{red}_i$ with an operation which satisfies inclusion, allowing us to get rid of the complex domain of ct-models, i.e., the sets of all models in languages in $\mathfrak{Sub}(\mathcal{L})$, and representing it in terms of the set of all models of the maximal language in $\mathfrak{Sub}(\mathcal{L})$.

3.2.2 Reduction and sublanguage equivalence

Among all the context operators we have considered in Chapter 2, the only genuinely non-boolean one was the focus operator fcs_i where i refers to the sublanguage $\mathcal{L}_i$ on which the context in the scope of the operator is 'focused' (see Remark 2.2.7). Semantics to that operator has been given in term of a model-theoretic operation of reduction, which we now briefly recall. The function $\mathtt{red}_i : \bigcup_{i \in \mathfrak{Sub}(\mathcal{L})} \mathcal{P}(\mathbf{M}_i^{\Delta}) \longrightarrow \mathcal{P}(\mathbf{M}_i^{\Delta})$ defined as follows:

$$\mathtt{red}_i(X) = \{m \in \mathbf{M}_i^{\Delta} \mid \exists m' \in X : \rceil_i m' = m\}$$

where X is a set of models of sub-language $\mathcal{L}_i$ of $\mathcal{L}$, i.e., $X \subseteq \mathbf{M}_i^{\Delta}$ for some $i \in \mathfrak{Sub}(\mathcal{L})$. Functions $\{\mathtt{red}_i\}_{0 \leq i \leq n}$ yield for a given X the set all of models of $\mathcal{L}_i$ which are restrictions of at least one model in X.

Aim of this section is to show that the reduction function $\mathtt{red}_i$ and the sublanguage equivalence function $\mathtt{eq}_i$ are strict relatives.

Definition 3.3. *(Extending models)*
Consider two $\mathcal{ALC}$ languages $\mathcal{L}_i \subseteq \mathcal{L}$ and denote as usual the set of models of $\mathcal{L}_i$ as $\mathbf{M}_i^{\Delta}$ and of $\mathcal{L}$ as $\mathbf{M}^{\Delta}$. The extension function $\mathtt{ext} : \bigcup_{i \in \mathfrak{Sub}(\mathcal{L})} \mathcal{P}(\mathbf{M}_i^{\Delta}) \longrightarrow \mathcal{P}(\mathbf{M}^{\Delta})$ is defined as follows:

$$\mathtt{ext}(X) = \{m \in \mathbf{M}^{\Delta} \mid \exists m' \in X : \rceil_i m = m'\}$$

where X is a set of models of sub-language $\mathcal{L}_i$ of $\mathcal{L}$, i.e., $X \subseteq \mathbf{M}_i^{\Delta}$ for some i s.t. $0 \leq i \leq n$.

Function $\mathtt{ext}$ yields for a given X the set of all models of $\mathcal{L}$ whose restrictions to $\mathcal{L}_i$ are in X, or, in other words, all the extensions in $\mathbf{M}^\Delta$ of the models in X. It is worth noticing that, given a set $X \subseteq \mathbf{M}_i^\Delta$, any concept γ which is not expressible in $\mathcal{L}_i$ gets in $\mathtt{ext}(X)$ all possible denotations on Δ, from $\emptyset$ to Δ itself.

Function $\mathtt{ext}$ enjoys the following properties.

Proposition 3.3. *(Properties of* $\mathtt{ext}$*)*
For any $X \subseteq \mathbf{M}^\Delta$ *in* $\mathcal{L}$*:*

- $\mathtt{ext}(X) = X$, *i.e., the extension of a set of models* $X \subseteq \mathbf{M}^\Delta$ *is* X *itself;*
- $\mathtt{ext}(\emptyset) = \emptyset$ *and* $\mathtt{ext}(\mathbf{M}^\Delta) = \mathbf{M}^\Delta$*;*
- $\mathtt{ext}(\mathbf{M}_i^\Delta) = \mathbf{M}^\Delta$ *, i.e., by extending the set of all models of a sublanguage* $\mathcal{L}_i$ *we obtain the set of all models of* $\mathcal{L}$*.*

Proof. The first property is directly proven from Definition 3.3 by considering that the reduction of the domain D of a function f to the set D itself yields f: if D is the domain of f then $D\rceil f = f$. The second and third properties follow directly from the definition of $\mathtt{ext}$. □

In other words, the proposition shows that the elements of $\mathcal{P}(\mathbf{M}^\Delta)$ are fixpoints for $\mathtt{ext}$. A second relevant property is that the application of functions $\mathtt{ext}$ and $\mathtt{eq}_i$ can be swapped.

Proposition 3.4. *(*$\mathtt{eq}_i$ *and* $\mathtt{ext}$*)*
It holds that, for any $\mathcal{L}_i, \mathcal{L}_j$ *s.t.* $\mathcal{L}_i \subseteq \mathcal{L}_j \subseteq \mathcal{L}$*, and any* $X \subseteq \mathbf{M}_j^\Delta$*:*

$$\mathtt{ext}(\mathtt{eq}_i(X)) = \mathtt{eq}_i(\mathtt{ext}(X)).$$

Proof. We prove the following fact:

$$\forall m \in \mathbf{M}^\Delta, m \in \mathtt{ext}(\mathtt{eq}_i(X)) \Leftrightarrow m \in \mathtt{eq}_i(\mathtt{ext}(X)).$$

($\Rightarrow$) Assume $m \in \mathtt{ext}(\mathtt{eq}_i(X))$. It follows: $\exists m' \in \mathbf{M}_j^\Delta, m'' \in X$ s.t. $m' = \rceil_j m$ and $m' \sim_i m''$. It is easily seen that it is always possible to build a model $m''' \in \mathbf{M}^\Delta$ s.t. $\rceil_i m''' = m''$ and thus s.t. $m \sim_i m'''$. Hence, $m \in \mathtt{eq}_i(\mathtt{ext}(X))$.
($\Leftarrow$) Assume $m \in \mathtt{eq}(\mathtt{ext}(X))$. It follows: $\exists m''' \in \mathbf{M}^\Delta, m'' \in X$ s.t. $m'' = \rceil_j m'''$ and $m \sim_i m'''$. It is then easy to build a model $m' \in \mathbf{M}_j^\Delta$ s.t. $m' = \rceil_j m$ and thus s.t. $m'' \sim m'$. Hence, $m \in \mathtt{ext}(\mathtt{eq}_i(X))$. □

The proposition just states that, given a set of models $X \subseteq \mathbf{M}_j^\Delta$, if I first consider all the models that agree with at least one model in X up to $\mathcal{L}_i$, and then all their extensions in $\mathcal{L}$ I obtain the same set which I would obtain by first taking the extension of X in $\mathcal{L}$, and then taking all the models in $\mathcal{L}$ that agree, up to $\mathcal{L}_i$, with at least one model in the extension.

At this point everything is in place for proving a key result of this section.

Proposition 3.5. *($\mathtt{red}_i$ and $\mathtt{eq}_i$)*
It holds that, for any $X \subseteq \mathbf{M}_j^\Delta$ with $\mathcal{L}_j \subseteq \mathcal{L}$:

$$\mathtt{ext}(\mathtt{eq}_i(X)) = \mathtt{ext}(\mathtt{eq}(\mathtt{red}_i(X)))$$

Proof. We prove the following fact:

$$\forall m \in \mathbf{M}^\Delta, m \in \mathtt{ext}(\mathtt{eq}_i(X)) \Leftrightarrow m \in \mathtt{ext}(\mathtt{eq}(\mathtt{red}_i(X))).$$

($\Rightarrow$) Assume $m \in \mathtt{ext}(\mathtt{eq}_i(X))$. It follows: $\exists m' \in X, m'' \in \mathbf{M}_j^\Delta$ s.t. $m'' =\rceil_j m$ and $m' \sim_i m''$. We have also that: $\exists m''', m'''' \in \mathbf{M}_i^\Delta$ s.t. $m''' =\rceil_i m'$, $m'''' =\rceil_i m''$. Because of this, and since $m' \sim_i m''$, we have that $m''' \sim m''''$. But $m'''' =\rceil_i m$ and hence $m \in \mathtt{ext}(\mathtt{eq}(\mathtt{red}_i(X)))$.
($\Leftarrow$) Assume $m \in \mathtt{ext}(\mathtt{eq}(\mathtt{red}_i(X))))$. It follows: $\exists m''', m'''' \in \mathbf{M}_i^\Delta, m' \in X$ s.t. $m''' \sim m''''$, $m''' =\rceil_i m'$ and $m'''' =\rceil_i m$. Now it should be the case that $\exists m'' \in \mathbf{M}_j^\Delta$ s.t. $m'''' =\rceil_i m''$ and $m'' =\rceil_j m$, because such a model m'' can always be built given that $m'''' =\rceil_i m$. Since $m''' \sim m''''$, it holds that $m' \sim_i m''$. Hence, $m \in \mathtt{ext}(\mathtt{eq}_i(X))$. □

Roughly speaking, the proposition states that the set of all $\mathcal{L}_i$-equivalent models of X is equal —after normalization via $\mathtt{ext}$— to the set of all models which are equivalent to the reduction of the models in X. To put it otherwise, function $\mathtt{eq}_i \circ \mathtt{ext}$ is equal to function $\mathtt{red}_i \circ \mathtt{eq} \circ \mathtt{ext}$.

It follows that sublanguage equivalence on models of the maximal language $\mathcal{L}$ corresponds to the composition of the extension operation with the clustering under equivalence and the restriction function: for any $X \subseteq \mathbf{M}^\Delta$, it is the case that $\mathtt{eq}_i(X) = \mathtt{ext}(\mathtt{eq}(\mathtt{red}_i(X)))$.

3.2.3 Context algebras

A context algebra is the algebra on the contexts defined by the models of an $\mathcal{ALC}$ language $\mathcal{L}$ on a domain Δ, plus the family of the sublanguage equivalence operations $\mathtt{eq}_i$ of the sublanguages $\mathcal{L}_i$ of $\mathcal{L}$.

Definition 3.4. *(Context algebras)*
A context algebra of an $\mathcal{ALC}$ language $\mathcal{L}$ is a structure

$$\mathfrak{A} = \langle Cxt, \cup, -, \emptyset, \{\mathtt{eq}_i\}_{0 \leq i \leq n} \rangle$$

where the set Cxt:

- *is a subset of $\mathcal{P}(\mathbf{M}^\Delta)$, containing $\emptyset$ and $\mathbf{M}^\Delta$, i.e., the set of models of $\mathcal{L}$ on domain Δ;*
- *is closed under the usual boolean operations $\cup$ (union) and $-$ (complement);*
- *is closed under the $\mathtt{eq}_i$ operations in $\{\mathtt{eq}_i\}_{0 \leq i \leq n}$, where indexes i denote sublanguages $\mathcal{L}_i$ of $\mathcal{L}$.*

Context algebras are algebras of an interesting kind, namely, Boolean Algebras with Operators.

Proposition 3.6. *(Properties of context algebras)*
Context algebras are Boolean Algebras with operators.

Proof. That structure $\langle Cxt, \cup, -, \emptyset \rangle$ is a Boolean Algebra is obvious since it is a set algebra. That the the $\mathtt{eq}_i$ operations in $\{\mathtt{eq}_i\}_{0 \leq i \leq n}$ are normal and additive follows from proposition. 3.2. $\square$

Context algebras are, in effect, instances of complex algebras ([Blackburn et al., 2001]) obtained by extending the set algebras on $\mathbf{M}^{\Delta}$ with a family of $\mathtt{eq}_i$ operators, which have been defined on the basis of the relation $\sim_i$ of sublanguage equivalence (Definition 3.2). This clearly points to a strict relationship between context algebras and modal logics, which will be investigated in Section 3.3 of this chapter.

Before doing this we will show, in the next section, that every interpretation $\mathbb{I}$ of context descriptions ξ in ct-models yields an interpretation of the same context descriptions on a context algebra (Proposition 3.7) and that this interpretation preserves the validity of contextual subsumptions (Proposition 3.8).

3.2.4 Ct-models as context algebras

We now show that every ct-model $\mathbb{M}$ can be thought of as a context algebra $\mathfrak{A}^{\mathbb{M}}$. First recollect that the interpretation function in ct-models is a function $\mathbb{I} : \mathbf{c} \longrightarrow \mathcal{P}(\mathbf{M}_0^{\Delta}) \cup \ldots \cup \mathcal{P}(\mathbf{M}_n^{\Delta})$ where $\mathbf{c}$ is the set of context identifiers and n the number of sublanguages $\mathcal{L}_i$ considered. The context algebra $\mathfrak{A}^{\mathbb{M}} = \left\langle Cxt^{\mathbb{M}}, \cup, -, \emptyset, \{\mathtt{eq}_i\}_{0 \leq i \leq n} \right\rangle$ of a ct-model $\mathbb{M}$ will be obtained by performing some specific operations on the context interpretation function $\mathbb{I}$ of $\mathbb{M}$.

Let us define a function $\mathtt{alg}^{\mathbb{I}} : \mathbf{c} \longrightarrow Cxt^{\mathbb{M}}$ as follows:

$$\mathtt{alg}^{\mathbb{I}} = \mathbb{I} \circ \mathtt{eq} \circ \mathtt{ext}.$$

Intuitively, function $\mathtt{alg}^{\mathbb{I}}$ 'builds' elements of $Cxt^{\mathbb{M}}$ from the context interpretation function $\mathbb{I}$ of $\mathbb{M}$ by applying the following procedure: first, the set of all the models of $\mathcal{L}_i$ —assuming $\mathcal{L}_i$ to be the of the context assigned by $\mathbb{I}$— which are $\mathcal{L}_i$-equivalent to at least one model in $\mathbb{I}(c)$ is built; then, this set is extended to a set of models of $\mathcal{L}$.

Now, to prove that $\mathtt{alg}^{\mathbb{I}}$ yields indeed the the algebra we want it needs to be proven that for all context descriptions ξ, $\mathtt{alg}^{\mathbb{I}^*}(\xi) \in Cxt^{\mathbb{M}}$, where $\mathtt{alg}^{\mathbb{I}^*}$ is defined in the same fashion of $\mathtt{alg}^{\mathbb{I}}$ by taking the inductive extension $\mathbb{I}^*$ of $\mathbb{I}$ (see Definition 2.4), that is:

$$\mathtt{alg}^{\mathbb{I}^*} = \mathbb{I}^* \circ \mathtt{eq} \circ \mathtt{ext}$$

To put it otherwise, function $\mathtt{alg}^{\mathbb{I}^*} : \Xi \longrightarrow Cxt^{\mathbb{M}}$ interprets all context descriptions ξ on the support $Cxt^{\mathbb{M}}$ of $\mathfrak{A}^{\mathbb{M}}$.

Proposition 3.7. *($\mathfrak{A}^{\mathbb{M}}$ is the context algebra of $\mathbb{M}$)*
For all context descriptions $\xi \in \Xi$, it holds that $\mathtt{alg}^{\mathbb{I}^*}(\xi) \in Cxt^{\mathbb{M}}$

Proof. By induction on the complexity of context expressions ξ.

[B] The atomic context case holds by definition of $\mathtt{alg}^{\mathbb{I}^*}$. As to the nullary operators, it is easy to see that:

$$\begin{aligned}\mathtt{ext}(eq(\mathbb{I}^*(\bot_i))) &= \mathtt{ext}(eq(\emptyset))\ \text{[Semantics of } \bot_i]\\ &= \mathtt{ext}(\emptyset)\ \text{[Proposition 3.2]}\\ &= \emptyset\ \text{[Definition 3.3]}\end{aligned}$$

which is in $Cxt^{\mathbb{M}}$ by definition.

$$\begin{aligned}\mathtt{ext}(eq(\mathbb{I}^*(\top_i))) &= \mathtt{ext}(eq\mathbf{M}_i^{\Delta}))\ \text{[Semantics of } \top_i]\\ &= \mathtt{ext}(\mathbf{M}_i^{\Delta})\ \text{[Proposition 3.2]}\\ &= \mathbf{M}^{\Delta}\ \text{[Definition 3.3]}\end{aligned}$$

which is also in $Cxt^{\mathbb{M}}$ by definition.

[S] As to the focus operator:
if $\mathtt{alg}^{\mathbb{I}^*}(\xi) \in Cxt^{\mathbb{M}}$ then $\mathtt{alg}^{\mathbb{I}^*}(fcs_i(\xi)) \in Cxt^{\mathbb{M}}$.

$$\begin{aligned}\mathtt{ext}(eq(\mathbb{I}^*(fcs_i(\xi)))) &= \mathtt{ext}(eq(\mathtt{red}_i\mathbb{I}^*(\xi))\ \text{[Semantics of } fcs_i(\xi)]\\ &= \mathtt{ext}(eq_i(\mathbb{I}^*(\xi)))\ \text{[Proposition 3.5]}\\ &= eq_i(\mathtt{ext}(\mathbb{I}^*(\xi)))\ \text{[Proposition 3.4]}\\ &= eq_i(eq(\mathtt{ext}(\mathbb{I}^*(\xi))))\ \text{[Proposition 3.2]}\\ &= eq_i(\mathtt{ext}(eq(\mathbb{I}^*(\xi))))\ \text{[Proposition 3.4]}\\ &= eq_i(\mathtt{alg}^{\mathbb{I}^*}(\xi))\ \text{[Definition of } \mathtt{alg}^{\mathbb{I}^*}]\end{aligned}$$

Therefore, by the induction hypothesis and Definition 3.4, $\mathtt{alg}^{\mathbb{I}^*}(fcs_i\xi) \in Cxt$. The$\curlyvee_i$-case follows from this last result by application of Proposition 3.2 and of the properties of $\mathtt{ext}$. □

Context descriptions $\xi \in \Xi$ can therefore be interpreted on context algebras via function $\mathtt{alg}^{\mathbb{I}^*}$. The following section shows that this interpretation preserves the validity of contextual subsumptions.

3.2.5 Invariance of contextual subsumptions

We have now to prove that contexts in a ct-model $\mathbb{M}$ and the corresponding contexts in a context algebra $\mathfrak{A}^{\mathbb{M}}$ built via $\mathtt{alg}^{\mathbb{I}^*}$ satisfy exactly the same subsumptions between $\mathcal{ALC}$ concepts.

Proposition 3.8. *($\mathcal{ALC}$ contextual validity is preserved in $\mathfrak{A}^{\mathbb{M}}$)*
Let $\mathbb{M}$ be a ct-model in an $\mathcal{ALC}$ language $\mathcal{L}$ and let $\{\mathcal{L}_i\}_{0\leq i\leq n}$ be the set of sublanguages considered in $\mathbb{M}$. For all context expressions ξ, and all $\mathcal{ALC}$ concepts γ_1, γ_2 in $\mathcal{L}$: $\mathbb{M} \models \xi : \gamma_1 \sqsubseteq \gamma_2$ iff $\forall m \in \mathtt{alg}^{\mathbb{I}^*}(\xi) : \mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$.

Proof. We prove the following claim:

$$\mathbb{M} \models \xi : \gamma_1 \sqsubseteq \gamma_2 \quad \text{iff} \quad \forall m \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi))) : \mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$$

First of all recall the semantics of contextual concept subsumption in ct-models (Definition 2.2.6).

[Right to left] We assume per absurdum: $\forall m \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi))) : \mathcal{I}_m(\gamma_1) \subseteq \mathcal{I}_m(\gamma_2)$ and $\mathbb{M} \not\models \xi : \gamma_1 \sqsubseteq \gamma_2$. This is the case: either a) because the interpretation functions $\mathcal{I}_m$'s of models $m \in \mathbb{I}^*(\xi)$ are not defined on γ_1 or on γ_2; or b) because they are defined on those concepts but $\exists m \in \mathbb{I}^*(\xi)$ s.t. $\mathcal{I}_m(\gamma_1) \not\subseteq \mathcal{I}_m(\gamma_2)$. Notice first that the interpretation functions $\mathcal{I}_m$'s of models $m \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi)))$ are always defined on γ_1 and γ_2 since they are interpretations of the global language $\mathcal{L}$. Now, suppose a) is the case and in particular that γ_1 is the concept which does not get denotation in the models in $\mathbb{I}^*(\xi)$. The models in $\mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi)))$ assign to γ_1 all possible denotations on domain Δ. In particular, $\forall m \in \mathbb{I}^*(\xi), \exists m', m'' \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi)))$ s.t. $\mathcal{I}_{m'}(\gamma_1) = \emptyset$ and $\mathcal{I}_{m'}(\gamma_1) = \Delta$. The same considerations apply if it is γ_2 that does not get a denotation in $\mathbb{I}^*(\xi)$. Hence, if a) is the case then $\exists m \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi)))$ s.t. $\mathcal{I}_m(\gamma_1) \not\subseteq \mathcal{I}_m(\gamma_2)$, which implies a contradiction. If b) is the case, suppose then that $\mathbb{I}^*(\xi) \subseteq \mathbf{M}_i^{\Delta}$ and thus that $\gamma_1, \gamma_2 \in \mathcal{L}_i$. Since $\exists m \in \mathbb{I}^*(\xi)$ s.t. $\mathcal{I}_m(\gamma_1) \not\subseteq \mathcal{I}_m(\gamma_2)$ it follows that $\exists m' \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi)))$ and $\exists m'' \in \mathtt{eq}(\mathbb{I}^*(\xi))$ s.t. $m'' =\rceil_i m'$ and $m'' \sim m$. Hence, $\exists m' \in \mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi))) : \mathcal{I}_{m'}(\gamma_1) \not\subseteq \mathcal{I}_{m'}(\gamma_2)$, which is impossible.

[Left to right] The direction from left to right is easily proven considering that function $\mathtt{eq} \circ \mathtt{ext}$ leaves the subsumptions holding between the concepts interpreted in $\mathbb{I}^*(\xi)$ untouched, since the models in $\mathtt{ext}(\mathtt{eq}(\mathbb{I}^*(\xi)))$ are built by first taking equivalent models and then expanding them to the global language. Hence, the $\mathcal{L}_i$ subsumptions holding in $\mathbb{I}^*(\xi)$ hold by construction also in $\mathtt{alg}^{\mathbb{I}^*}(\xi)$. □

As far as contextual validity of subsumption statements is concerned, ct-models and their context algebras are therefore equivalent. The interesting aspect of this proposition can be noticed in the proof. The fact that a contextual subsumption $\xi : \gamma_1 \sqsubseteq \gamma_2$ does not hold in a ct-model $\mathbb{M}$ because either γ_1 or γ_2 does not belong to the language of $\mathbb{I}^*(\xi)$ is equivalent to the fact that in $\mathtt{alg}^{\mathbb{I}^*}(\xi)$ it is neither the case that $\gamma_1 \sqsubseteq \gamma_2$ nor that $\gamma_1 \not\sqsubseteq \gamma_2$. That is exactly how the interpretation of context descriptions on context algebras can model the linguistic dependence of contexts without needing to consider models on different languages.

This consideration also points at the reason why function $\mathtt{alg}^{\mathbb{I}^*}$ has been so designed. Roughly speaking, function $\mathtt{alg}^{\mathbb{I}^*}$ makes it so that each model m in a sublanguage $\mathcal{L}_i$ of the global language $\mathcal{L}$ can be thought of as the set of models of $\mathcal{L}$ which are equivalent to m w.r.t. what can be expressed in $\mathcal{L}_i$. In a way, sublanguage equivalence, which is nothing but a family of equivalence relations (see Proposition 3.2), provides the tool for expressing language dependence without actually considering different languages. This might be regarded as an interesting result in itself. It shows that if, on the one hand, "a general treatment of contexts may indeed wish to exempt contexts from the obligation to interpret every assertion" ([Shoham, 1991], p.400), on the other hand, in order to represent that contexts do not interpret every assertion, we are exempted from using partial functions and similar machinery and we can simply resort to equivalence relations.

3.2.6 Context algebra and context theory

This section shows what are the consequences, at a theoretical level, of the perspective on context emerging from the previous section with respect to the local model semantics (LMS) of [Ghidini and Giunchiglia, 2001].

In Section 2.1.5 it has already been observed that our formal approach to context borrows the key idea of viewing contexts as sets of models from LMS. It has also been anticipated in Section 2.1.7 that the main difference between our view on contexts and the one yielded by LMS resides in the formal analysis of context interplay.

In our view, context interplay can be algebraically described. In LMS, instead, context interplay is not analyzed in terms of operations on contexts. In fact, no syntax is introduced in order to express operators on context expressions, and only atomic context descriptions are used.

As it was already exposed in Section 2.1.6 of the previous chapter, in LMS the central notion for the analysis of context interplay is the notion of bridge rule. Bridge rules are intended as inter-contextual inference rules: "if ϕ_1 holds in c_1, then infer that ϕ_2 holds in c_2". They are made sound by appropriately defined compatibility relations holding between sets of models. However, depending on the compatibility relation chosen to link the contexts, virtually any bridge rule can be made sound. Bridge rules are, in this sense, not necessarily logical but they could be just domain-specific rules.

Instead, by explicitly introducing context operations we can aim at the specification of logical inter-contextual inference rules based on those operations. This suggests the possibility to import to contextual reasoning the standard distinction between inference rules which are domain-specific, and inference rules which are instead logical. This clearly happens by interpreting the focus operator fcs_i as the operation eq_i. Corresponding laws of context interaction become thus readily available from the properties of eq_i. For example, $\xi \preccurlyeq fcs_i(\xi)$ which follows from the inclusion property of eq_i and which just states that every context is a subcontext of the context obtained by abstracting from what is not expressible in $\mathcal{L}_i$. This suggests the following bridge rule for contextual terminologies:

$$\frac{\vdash fcs_i(\xi) : \gamma_1 \sqsubseteq \gamma_2}{\vdash \xi : \gamma_1 \sqsubseteq \gamma_2} \tag{3.1}$$

that is, if a subsumption $\gamma_1 \sqsubseteq \gamma_2$ holds in the focus of ξ then it can be inferred that $\gamma_1 \sqsubseteq \gamma_2$ holds also in ξ, since it is always a subcontext of its focus. Like any bridge rule Formula 3.1 expresses that a context and its focus "agree up to a certain extent", but the reason of that agreement lies on the structural properties of contexts and not on an extrinsically specified relation on the set of contexts.

Bridge rules in the fashion of Formula 3.1, just like the conception of contexts as elements of a Boolean Algebra with Operators (see Proposition 3.6), seem thus to suggest the existence of a precise logic of inter-contextual reasoning for contextual terminologies. The next section is devoted to this issue.

3.3 Release

This section studies, from a logical point of view, the "talent for being silent" about some issues. Being silent about some issues is interpreted as releasing part of the language in which we still want to say something.

3.3.1 Propositional release logic

Propositional release logics (PRL) have been first introduced and studied in [Krabbendam and Meyer, 1999, 2000] in order to provide a modal logic characterization of the notion of *irrelevancy*. Irrelevancies are, in short, those aspects which we can choose to ignore.

Irrelevancy is represented via modal *release operators*, specifying what is relevant to the current situation and what can instead be ignored as noise. Release operators are indexed by an abstract 'issue' denoting what is considered to be irrelevant for evaluating the formula in the scope of the operator:

- $\triangle_I\, \phi$: 'formula ϕ holds in all states where issue I is irrelevant' or 'ϕ holds in all states modulo the set of issues I' or 'ϕ necessarily holds while *releasing* issue I'.

- $\triangledown_I \phi$: 'formula ϕ holds in at least one of the states where issue I is irrelevant' or 'ϕ necessarily holds while *releasing* issue I'.

Issues can be in principle anything, but their essential feature is that they yield equivalence relations which cluster the states in the model. An issue I is conceived as something that determines a partition of the domain in clusters of states which agree on everything but I, or which are equivalent modulo I. Release operators are interpreted on these equivalence relations. As such, propositional release logic can be thought of as a "logic of controlled ignorance" ([Krabbendam and Meyer, 1999]). They are logic of *ignorance*, because release operators are in fact formally analogous to epistemic operators clustering the domain in epistemically indistinguishable states. They are a logic of *controlled* ignorance because these epistemic operators are indexed by issues representing what we choose to ignore. Release operators represent what we would know, and thus what we would ignore, by choosing to disregard some issues.

3.3.2 PRL: syntax, semantics, axiomatics

The syntax of PRL is the syntax of a standard multi-modal language $\mathcal{L}_n$ ([Blackburn et al., 2001]) where n is the cardinality of the set `Iss` of releasable issues. The alphabet of $\mathcal{L}_n$ contains: an at most countable set $\mathbb{P}$ of propositional atoms p; the set of boolean connectives $\{\neg, \wedge, \vee, \rightarrow\}$; a finite non-empty set `Iss` of issues. Metavariables $I, J, \ldots$ are used for denoting elements of `Iss`. The set of well formed formulae ϕ of $\mathcal{L}_n$ is defined by the usual BNF:

$$\phi ::= \top \mid p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid \phi_1 \rightarrow \phi_2 \mid \ \triangle_I\ \phi \mid \triangledown_I \phi.$$

where I denotes elements in `Iss`.

One last important feature of PRL should be addressed before getting to the semantics. We have seen that modal operators are indexed by an issue denoting what is disregarded when evaluating the formula in the scope of the operator. The finite set `Iss` of these issues is structured as a partial order, that is to say, $\langle \texttt{Iss}, \leqq\rangle$ is a structure on the non-empty set `Iss`, where $\leqq$ ("being a sub-issue of") is a binary relation on `Iss` which is reflexive, transitive and antisymmetric. The aim of the partial order is to induce a structure on the equivalence relations denoting the release of each issue in `Iss`: if $I \leq J$ then the clusters of states obtained by releasing J contain the clusters of states obtained by releasing I. Intuitively, if I is a sub-issue of J then by disregarding J, I is also disregarded. This aspect is made explicit in the models which, for the rest, are just Kripke models.

Definition 3.5. *(PRL models)*
Let $\langle \texttt{Iss}, \leqq\rangle$ be a partial order. A PRL model is a structure $\mathcal{M} = \langle \mathcal{F}, \mathcal{I}\rangle$ where:

- *$\mathcal{F}$ is a frame, i.e., a structure $\langle W, \{R_I\}_{\texttt{Iss}}\rangle$, where W is a non-empty set of states (or possible worlds) and $\{R_I\}_{\texttt{Iss}}$ is a family of equivalence relations such that: if $I \leqq J$ then $R_J \subseteq R_I$;*
- *$\mathcal{I}$ is an evaluation function $\mathcal{I} : \mathbb{P} \longrightarrow \mathcal{P}(W)$ associating to each atom the set of states which make it true.*

PRL models are therefore just $\mathbf{S5}_n$ models with the further constraint that the granularity of the equivalence relations follows the partial order defined on the set of issues: the $\leqq$-smaller is the issue released, the more granular is the partition obtained via the associated equivalence relation.

The satisfaction relation is standard.

Definition 3.6. *(Satisfaction for PRL models)*
Let $\mathcal{M}$ be a PRL model.

$$\begin{array}{rll}
\mathcal{M}, w \models \top & & \\
\mathcal{M}, w \models p & \textit{iff} & w \in \mathcal{I}(p) \\
\mathcal{M}, w \models \neg\phi_1 & \textit{iff} & \text{NOT } \mathcal{M}, w \models \phi_1 \\
\mathcal{M}, w \models \phi_1 \wedge \phi_2 & \textit{iff} & \mathcal{M}, w \models \phi_1 \text{ AND } \mathcal{M}, w \models \phi_2 \\
\mathcal{M}, w \models \phi_1 \vee \phi_2 & \textit{iff} & \mathcal{M}, w \models \phi_1 \text{ OR } \mathcal{M}, w \models \phi_2 \\
\mathcal{M}, w \models \phi_1 \rightarrow \phi_2 & \textit{iff} & \mathcal{M}, w \models \phi_1 \text{ IMPLIES } \mathcal{M}, w \models \phi_2 \\
\mathcal{M}, w \models \triangle_I\ \phi & \textit{iff} & \forall\, w', wR_Iw' : \mathcal{M}, w' \models \phi \\
\mathcal{M}, w \models \triangledown_I \phi & \textit{iff} & \exists w', wR_Iw' : \mathcal{M}, w' \models \phi.
\end{array}$$

where $I \in$ `Iss`. As usual, a formula ϕ is said to be valid in a model $\mathcal{M}$, in symbols $\mathcal{M} \models \phi$, iff for all w in W, $\mathcal{M}, w \models \phi$. It is said to be valid in a frame $\mathcal{F}$ ($\mathcal{F} \models \phi$) if it is valid in all models based on that frame. Finally, it is said to be valid on a class of frames F *(*$\mathrm{F} \models \phi$*) if it is valid in every frame $\mathcal{F}$ in* F.

Finally, the axiomatics runs as follows and it amounts to a multi modal **S5** plus the `PO` (partial order) axiom:

(`P`)	all tautologies of propositional calculus
(`K`)	$\triangle_I (\phi_1 \to \phi_2) \to (\triangle_I \phi_1 \to \triangle_I \phi_2)$
(`T`)	$\triangle_I \phi \to \phi$
(`4`)	$\triangle_I \phi \to \triangle_I \triangle_I \phi$
(`5`)	$\nabla_I \phi \to \triangle_I \nabla_I \phi$
(`PO`)	$\triangle_I \phi \to \triangle_J \phi \quad \text{IF } J \leqq I$
(`Dual`)	$\nabla_I \phi \leftrightarrow \neg \triangle_I \neg \phi$
(`MP`)	IF $\vdash \phi_1$ AND $\vdash \phi_1 \to \phi_2$ THEN $\vdash \phi_2$
($\mathtt{N}^I$)	IF $\vdash \phi$ THEN $\vdash \triangle_I \phi$

where $I, J \in$ `Iss`.

A proof of the soundness and completeness of this axiomatics w.r.t. to the semantics presented in Definition 3.6 is exposed in [Krabbendam and Meyer, 2000].

3.3.3 PRL with Boolean Algebras

The partial order structure of a PRL logic is reflected in the axiomatics by axiom `PO`, and in the semantics by a partial order on the accessibility relations. By adding structure to the partial order on the set of issues more validities can be derived which mirror that structure. Interesting for our purposes is the case when `Iss` is structured according to a Boolean Algebra. The following propositions lists some of the PRL validities holding in that case.

Proposition 3.9. *(Validities of PRL with BA)*
Let `Iss` *be ordered as* $\langle \mathtt{Iss}, \sqcup, \sqcap, -, 1, 0, \leqq \rangle$*, where* $\langle \mathtt{Iss}, \sqcup, \sqcap, -, 1, 0 \rangle$ *is a Boolean Algebra. The following formulae can be derived in PRL:*

$$\triangle_1 \phi \to \triangle_I \phi \tag{3.2}$$

$$\triangle_I \phi \to \triangle_0 \phi \tag{3.3}$$

$$\triangle_{I \sqcup J} \phi \to (\triangle_I \phi \wedge \triangle_J \phi) \tag{3.4}$$

$$\triangle_{I \sqcup J} \phi \to (\triangle_I \triangle_J \phi \wedge \triangle_J \triangle_I \phi) \tag{3.5}$$

$$\triangle_I \phi \vee \triangle_J \phi \to \triangle_{I \sqcap J} \phi \tag{3.6}$$

$$\triangle_I \phi \leftrightarrow \triangle_{--I} \phi \tag{3.7}$$

$$(\triangle_I \phi \to \triangle_J \phi) \leftrightarrow (\triangle_{-J} \phi \to \triangle_{-J} \phi) \tag{3.8}$$

Proof. The desired derivations are easily obtainable: some (Formulae 3.2, 3.3, 3.4) are just instances of P0, some (Formulae 3.6, 3.7, 3.8) can be proven by application of P0 and propositional logic. Formula 3.5 is derived by applying P0, 4 and propositional logic. □

3.3.4 Contextual reasoning as reasoning in PRL

This section shows that reasoning with context descriptions in $\mathcal{L}^{CT}$ languages is an instance of release reasoning. This can be done by just showing that the context algebra $\mathfrak{A}^{\mathbb{M}}$ of a ct-model $\mathbb{M}$ can be viewed as a PRL model.

The modal language of such model would consists of the context descriptions Ξ', i.e., context descriptions built with boolean connectives (denoted by the standard $\vee, \neg$) and focus operators (denoted now by release operators $\vartriangle_{-i}$) from the set $\mathbf{c}$ of atomic indexes. Notice that is precisely the syntax of the context descriptions algebra introduced by Section 2.2.7 in Chapter 2.

Definition 3.7. *(Context release models)*
A context release model of a ct-model $\mathbb{M}$ is a structure $\mathcal{M}^{\mathbb{M}} = \langle W, \{R_i\}_{i \in \mathtt{Iss}}, \mathcal{I} \rangle$ where:

- $W = \mathbf{M}^{\Delta}$*, i.e., the set of states is the set of models for language $\mathcal{L}$ on domain Δ;*
- $\{R_i\}_{i \in \mathtt{Iss}} = \{\sim_{-i}\}_{i \in \mathfrak{Sub}(\mathcal{L})}$*, that is, as accessibility relations we take the set of $\mathcal{L}_i$-equivalence relations for every sublanguage in $\mathfrak{Sub}(\mathcal{L})$;*
- $\mathcal{I} : \Xi' \longrightarrow \mathcal{P}(\mathbf{M}^{\Delta})$ *such that* $\Xi \rceil \mathcal{I} = \mathtt{alg}^{\mathbb{I}^*}$*, that is, the interpretation function $\mathcal{I}$ is an extension of function $\mathtt{alg}^{\mathbb{I}^*}$ to Ξ'.*

Notice that the release issues Iss are the complements $-\mathcal{L}_i$ of the sublanguages in $\mathfrak{Sub}(\mathcal{L})$. In fact, what is released is just what cannot be expressed. The accessibility relations should therefore be taken to be the sublanguage-equivalence relations $\sim_{-i}$.

Notice also that the evaluation function is such that context descriptions in Ξ are interpreted according to $\mathtt{alg}^{\mathbb{I}^*}$. In other words, $\mathcal{I}$ is a homomorphism from the context description algebra introduced in Section 2.2.7 to a context algebra which interprets the relevant context descriptions Ξ via $\mathtt{alg}^{\mathbb{I}^*}$. The intuition of Definition 3.7 is just to view context descriptions as modal formulae, where focus operators fcs_i are denoted by release operators $\vartriangle_{-i}$, and the meaning of those formulae as homomorphism. To recap the whole argument in three equations, for all $\xi \in \Xi$:

$$\mathcal{I}(\vartriangle_{-i} \xi) = \mathtt{alg}^{\mathbb{I}^*}(fcs_i(\xi)) = eq_i(\mathtt{alg}^{\mathbb{I}^*}(\xi)) = \{w \in W \mid \exists w' \in \mathcal{I}(\xi) : w \sim_i w'\}.$$

At this point, by inspecting the frames on which context release models are built, it is easy to see that context release models are models for PRL.

Corollary 3.1. *(Contexts and PRL)*
Context release models are PRL models.

Proof. Follows directly from Proposition 3.2 and Definition 3.7. □

Context algebras built from ct-models can be viewed as PRL models. This also means that context algebras belong to the variety of Boolean Algebras with Operators where the axioms of PRL logic are valid. Leaving technicalities aside, context descriptions for contextual terminologies behave as formulae of a modal release logic and the logic of context interplay, i.e., the logic of inter-contextual reasoning in contextual terminologies, is modal release logic.

3.3.5 Towards a proof theory for contextual terminologies

The results presented in this chapter provide contextual terminologies with a proof-theory for inter-contextual reasoning. This section illustrates this aspect further. First of all recall (see Section 2.2.7) that context descriptions in $\mathcal{L}^{CT}$ languages can be reformulated in a context description algebra, and therefore in the syntax of modal logic and in particular: $\xi_1 \curlyvee_i \xi_2$ corresponds to $\vartriangle_{-i} (\xi_1 \vee \xi_2)$, and $\xi_1 \preccurlyeq \xi_2$ to $\neg\xi_1 \vee \xi_2$, i.e., $\xi_1 \rightarrow \xi_2$.

Let us now get back to Example 2.1. In its formalization in Example 2.3 we have shown that contexts c_{M1}, c_{M2} and c_{M3} inherited the classifications holding in the abstract context c_{Reg}. Somehow, ct-models supported a form of inheritance from more abstract contexts to more concrete one. We are now in the position to formally represent relations between concrete and abstract contexts simply as the model validity of implications such as:

$$c_{M1} \rightarrow c_{Reg}.$$

In addition, we can show that such relations can follow from the transitivity of the implication between context descriptions (i.e., context inclusion) and the fact that release operators (i.e., the focus operator) enjoy axiom D:

Assumption	$\vartriangle_{-0} (c_{M1} \vee c_{M2} \vee c_{M3}) \rightarrow c_{Reg}$	(3.9)
Instance of D	$c_{M1} \vee c_{M2} \vee c_{M3} \rightarrow \vartriangle_{-0} (c_{M1} \vee c_{M2} \vee c_{M3})$	(3.10)
Conclusion	$c_{M1} \vee c_{M2} \vee c_{M3} \rightarrow c_{Reg}$	(3.11)

Formula 3.9 restates in the syntax of modal logic Formula 2.13:

$$c_{M1} \curlyvee_0 c_{M2} \curlyvee_0 c_{M3} \preccurlyeq c_{Reg}$$

that is, the context inclusion statement saying that the focus on language $\mathcal{L}_0$ of the union of the three concrete contexts c_{M1}, c_{M2} and c_{M3} is a subcontext of the abstract context c_{Reg}. From this assumption and axiom D (Formula 3.10) it follows by propositional logic Formula 3.11, which states that c_{M1}, c_{M2} and c_{M3} are subcontexts of c_{Reg}. It is worth noticing that Formula 3.11 had no counterpart in the representation of the scenario in $\mathcal{L}^{CT}$ languages due to the fact that ct-models interpret context descriptions on sets of models of different languages. Contexts c_{M1}, c_{M2} and c_{M3} were interpreted in a different language than the one of c_{Reg}. As a consequence, it was simply not the case that the union of c_{M1}, c_{M2} and c_{M3} yielded a subset of the models of c_{Reg}. It was just a set of models fot another language. Now, what modal syntax interpreted on context algebras makes explicit is that, in fact, the union of c_{M1}, c_{M2} and c_{M3} defines a set of models (situations, worlds) which is smaller than the set of models (situations, worlds) which is defined by c_{Reg} and this precisely because the language of c_{Reg} releases many terms, or, to get back to our opening quote, "stays silent" about many issues. To put it in a nutshell, it is shown that being abstract (i.e. using less rich languages) means being more general (i.e., having more models). This observation completes our answer to the second research question.

3.3.6 Future work

The validity of Formula 3.11 under the assumption of Formula 3.9 justifies the inheritance of what holds in c_{Reg}, i.e., the bigger context, to each of c_{M1}, c_{M2} and c_{M3}, i.e., the smaller contexts. However, 'what holds in a context' cannot be expressed in the release logic presented, since it deals only with context descriptions and not with concept descriptions and subsumption relations in $\mathcal{ALC}$. In other words, we cannot yet give a modal counterpart of the whole reasoning happening in Example 2.1 and 2.2 but just of its inter-contextual component. Would it be possible to import also the intra-contextual reasoning dimension into a release model in order to obtain Kripke structures that would at the same time model the context logic as well as the various $\mathcal{ALC}$ terminologies holding in those contexts? This is topic of future research.

3.4 Conclusions

The chapter has shown that contexts in contextual terminologies can be conceived of as algebraic entities (Proposition 3.7 and 3.8), and that the logic of the descriptions of these entities can be viewed as modal release logic (Corollary 3.1). This provides a specific thesis about the study of contexts as "abstract mathematical entities" as first advocated in [McCarthy, 1986]. It also shows that if on the one hand "we may indeed wish to exempt contexts from the obligation to interpret every assertion" ([Shoham, 1991], p.400), on the other hand, in order to represent that contexts do not interpret every assertion, we can simply recur to the use of equivalence relations to model a form of 'controlled ignorance' about what is disregarded.

Finally, these results have also provided us with a formal handle for comparing of our approach to contextuality with the one advanced by local model semantics (Section 3.2.6).

Chapter 4

The Form of Social Reality

"Was jedes Bild, welcher Form immer, mit der Wirklichkeit gemein haben muss, um sie überhaupt – richtig oder falsch – abbilden zu können, ist die logische Form, das ist, die Form der Wirklichkeit."

"What every picture, of whatever form, must have in common with reality, in order to be able to depict it —correctly or incorrectly— in any way at all, is logical form, i.e., the form of reality."

L. Wittgenstein, "Logisch-Philosophische Abhandlung", 2.18

Even though "nobody has ever seen a state" not even "in a picture[1]", the present chapter does try to take one, though only in the formal sense suggested by the above quote from [Wittgenstein, 1921].

The analysis of social reality which is put forth in [Searle, 1969] and [Searle, 1995] aims at taking a hold on the complexity of social phenomena in terms of one essential basic brick: the notion of constitutive rule. According to Searle, the "*construction of social reality*" takes place by means of systems of constitutive rules. The paradigmatic syntax of these rules has the form of "counts-as" statements:

> [...] "institutions" are systems of constitutive rules. Every institutional fact is underlain by a (system of) rule(s) of the form "X counts as Y in context C" ([Searle, 1969], pp.51-52).

Despite the richness of Searle's analysis —and probably because of it— his work does not result in a rigorous theory of constitutive rules, nor does it go further in systematically analyzing the logical behaviour of counts-as statements. The present

[1] Recollect the quote from [Debray, 1997] opening Chapter 1.

chapter pursues exactly this line by proposing an analysis of counts-as statements in formal logic. As such, this analysis can be thought of as displaying what "*the (logical) form of social reality*" is. In doing this, Searle's work is constantly referred to and used as starting point of our analysis.

The chapter articulates an answer to the third research question. It analyzes three different senses in which it can be said that "X counts as Y in context C". For each of these different senses of counts-as a formal semantics is developed by making use of standard modal logic techniques. These investigations presented stem from the acknowledgment that 'counts-as can be said in many ways', and shows that these 'many ways' all have a precise formal semantics. They have to be thought of, essentially, as investigations in concept analysis by means of formal logic. From a methodological point of view, we will proceed as recommended in this quote.

> "[. . .] it seems to me obvious that the only rational approach to such problems [of concept analysis] would be the following: [1] We should reconcile ourselves with the fact that we are confronted, not with one concept, but with several different concepts which are denoted by one word; [2] we should try to make these concepts as clear as possible (by means of definition, or of an axiomatic procedure, or in some other way); [3] to avoid further confusions, we should agree to use different terms for different concepts; and then we may proceed to a quiet and systematic study of all concepts involved, which will exhibit their main properties and mutual relations" ([Tarski, 1944], p. 355).

All points from 1 to 3 emphasized in the quote will be addressed. We will disentangle some of the meanings of counts-as and study them formally, by means of modal logic, showing what their properties are and how they relate to one another.

The chapter builds on the series of papers [Grossi et al., 2005d, 2006f,g; Grossi. et al., 2007]. The exposition is organized as follows. In Section 4.1 the thesis underpinning the whole analysis is presented and motivated on the ground of what already discussed in Chapter 2: counts-as statements will be studied as the basic bricks of contextual terminologies. In Section 4.2 a modal logic characterization of counts-as statements as contextual classificatory statements is defined, and the syntactic properties of the resulting counts-as operator are investigated. In Section 4.3 some aspects of counts-as are touched upon which are not captured in the contextual classificatory reading. Then, following hints from legal and social theory, two stronger readings of the statements are set forth and the grounds for their modal logic characterization are laid. These two readings of counts-as, which we call *proper classificatory counts-as* and, respectively, *constitutive counts-as* are formally studied in Sections 4.4 and 4.5. The relationships between the three readings are studied in Section 4.6. In Section 4.7 the formalisms developed are related to other areas of philosophical and applied logic showing that the notion of counts-as bear, from a formal point of view, relevant similarities with a number of different notions in non-monotonic, epistemic and deontic logic. Finally, in Section 4.8 the results of our formal analysis are throughly compared with work on the formal analysis of counts-as available in the literature. The comparison will be led both from a model-theoretic point of view, and from the point of view of the structural properties enjoyed by the various syntaxes of counts-as operators to be found in the literature. Section 4.9 recapitulates some of our central results. Soundness and completeness of the logics introduced are proven in Appendix A.

4.1 Analytical Background

In this section we expose the analytical background of the chapter by recapitulating some of the relevant findings of Chapter 2 along the lines already made explicit in Section 2.6.4.

4.1.1 Subsumption as the basic ingredient of counts-as

In legal theory the non-regulative component of normative systems has been labeled in ways that emphasize a classificatory, as opposed to a normative or regulative, character: *determinative rules* ([Von Wright, 1963]), *conceptual rules* ([Bulygin, 1992]), *qualification norms* ([Peczenik, 1989]), *definitional norms* ([Jones and Sergot, 1992]). Constitutive rules are definitional in character:

> "The rules for checkmate or touchdown must 'define' *checkmate in chess* or *touchdown in American Football* [. . .]"([Searle, 1969], p.34).

Focusing on this feature, a first reading of counts-as is readily available: counts-as statements express classifications. For example, they express what is classified to be a checkmate in chess, or a touchdown in American Football.

In the light of this classificatory view of counts-as, the way to establish a formal characterization of counts-as statements is extremely plain: if counts-as statements yield classifications, this means that they function as conceptual subsumption relations, that is, counts-as statements assert just that a concept X is a subconcept of a concept Y. Via such classificatory statements, normative systems can establish the ontology they use in order to distribute obligations, rights, prohibitions, permissions. Vehicles are not admitted in public parks (general norm), but then, if bicycles are vehicles (classification), bicycles are not admitted in public parks (specific norm). The term `vehicle` works in this case as a sort of "middle term" ([Lindahl, 2004; Atkinson and Bench-Capon, 2005]), mediating between the general and the specific norm, and the counts-as statement provides the classification necessary for the derivation to soundly take place. This interplay phenomenon between regulative and non-regulative components, between norms and classificatory statements, is a crucial feature of normative systems ([Alchourrón and Bulygin, 1971]). As we already suggested in Section 2.6.4, it is our claim that this interplay works on a classificatory, i.e., terminological basis and that counts-as is a kind of basic brick of these terminologies. Each normative system, via its constitutive rules, states a set of classifications, i.e., a terminology which provides a conceptualization of the domain of entities it is supposed to regulate.

4.1.2 Contextual subsumption and modal logic

The intuition of viewing counts-as statements as subsumptions was, in fact, already stated in [Jones and Sergot, 1996]:

> "There are usually constraints within any institution according to which certain states of affairs of a given type count as, or *are to be classified as*, states of affairs of another given type" ([Jones and Sergot, 1996], p.431).

Notice that the quote puts forth a neat modal intuition. The constraint laid by a counts-as statement "X counts as Y" expresses that the states of a affairs of a given type X are included in the states of a given type Y. This amounts to a conceptual subsumption $X \sqsubseteq Y$ between state types X and Y. By recollecting that a subsumption $X \sqsubseteq Y$ is satisfied by an interpretation $\mathcal{I}$ if and only if $\mathcal{I}(X) \subseteq \mathcal{I}(Y)$, it becomes easy to see that $X \sqsubseteq Y$ corresponds, in modal propositional logic, to a material implication held to be valid in a model $\mathcal{M}$ containing interpretation $\mathcal{I}$: $\mathcal{M} \models X \rightarrow Y$.

A subsumption considered to hold unconditionally would then be formalized on a modal language containing an operator $[u]$ interpreted on universal frames, i.e., frames s.t. $\forall w, w' \in W : wR_1w'$. Now, the class Univ of universal frames is characterized by **S5** logic ([Blackburn et al., 2001]), and subsumption statements can therefore be represented as strict implications in logic **S5**: $[u]X \rightarrow Y$.

The terminology of each normative system can thus be represented via models built on frames from the class Univ of universal frames. The point is what happens if we want to consider, under the same modal formalism, a variety of such structures, that is to say, if we want to represent many terminologies belonging to different normative system specifications. The classifications expressed by counts-as statements are not absolute. They only hold in relation with a context, that is, in relation with the normative system to which they pertain: "X counts as Y *in context* C". How to fit contexts in this modal picture?

Technically we are interested in a logic that can "locally" behave like an **S5** logic but that can "globally" behave in a weaker way allowing for different and possibly inconsistent classificatory representations at the same time. In other words, we should find a multi-modal logic enabling as many modalities as the to be represented contexts, and retaining for these modalities as many characteristics of **S5** as possible, but at the same time allowing for the satisfiability of expressions such as: $[i](X \rightarrow Y) \wedge \neg[j](X \rightarrow Y)$. The following section addresses this issue.

4.2 Modal logic of Classificatory Counts-as

On the grounds of the intuitions exposed above this section introduces a modal logic of contextual subsumption.

4.2.1 Preliminaries

We first introduce the languages we are going to work with: propositional n-modal languages $\mathcal{L}_n$ ([Blackburn et al., 2001]). The alphabet of $\mathcal{L}_n$ contains: an at most countable set $\mathbb{P}$ of propositional atoms p; the set of boolean connectives $\{\neg, \wedge, \vee, \rightarrow\}$; a finite non-empty set of n context indexes C, and the operators $[\ ]$ and $\langle\ \rangle$. Metavariables $i, j, ...$ are used for denoting elements of C. The set of well formed formulae ϕ of $\mathcal{L}_n$ is then defined by the following BNF:

$$\phi ::= \top \mid p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid \phi_1 \rightarrow \phi_2 \mid [i]\phi \mid \langle i \rangle \phi.$$

We will refer to formulae in which at least one modal operator occurs as *modalized* formulae. Modalized formulae in which all non-logical symbols occur in the scope of a modal operator are called *contextual* formulae. Formulae in which no modal operator occurs are called instead *objective*, and we denote them using the metavariables $\gamma_1, \gamma_2, \ldots$.

Semantics for these languages is given via structures $\mathcal{M} = \langle \mathcal{F}, \mathcal{I} \rangle$, where:

- $\mathcal{F}$ is a frame, i.e., a structure $\langle W, \{R_i\}_{i \in C} \rangle$, where W is a non-empty set of states (or possible worlds) and $\{R_i\}_{i \in C}$ is a family of n accessibility relations ($|C| = n$). We will refer to the set of accessible worlds from a world w via a relation R_i as $r_i(w)$.
- $\mathcal{I}$ is an evaluation function $\mathcal{I} : \mathbb{P} \longrightarrow \mathcal{P}(W)$ associating to each atom the set of states which make it true.

Satisfaction for these languages is then defined as follows:

$$\begin{array}{rll}
\mathcal{M}, w \models \top & & \\
\mathcal{M}, w \models p & \text{iff} & w \in \mathcal{I}(p) \\
\mathcal{M}, w \models \neg\phi_1 & \text{iff} & \text{NOT } \mathcal{M}, w \models \phi_1 \\
\mathcal{M}, w \models \phi_1 \wedge \phi_2 & \text{iff} & \mathcal{M}, w \models \phi_1 \text{ AND } \mathcal{M}, w \models \phi_2 \\
\mathcal{M}, w \models \phi_1 \vee \phi_2 & \text{iff} & \mathcal{M}, w \models \phi_1 \text{ OR } \mathcal{M}, w \models \phi_2 \\
\mathcal{M}, w \models \phi_1 \rightarrow \phi_2 & \text{iff} & \mathcal{M}, w \models \phi_1 \text{ IMPLIES } \mathcal{M}, w \models \phi_2 \\
\mathcal{M}, w \models [i]\phi & \text{iff} & \forall\, w' \in r_i(w) : \mathcal{M}, w' \models \phi \\
\mathcal{M}, w \models \langle i \rangle \phi & \text{iff} & \exists\, w' \in r_i(w) : \mathcal{M}, w' \models \phi.
\end{array}$$

A formula ϕ is said to be valid in a model $\mathcal{M}$, in symbols $\mathcal{M} \models \phi$, iff for all w in W, $\mathcal{M}, w \models \phi$. It is said to be valid in a frame $\mathcal{F}$ ($\mathcal{F} \models \phi$) if it is valid in all models based on that frame. Finally, it is said to be valid on a class of frames F (F $\models \phi$) if it is valid in every frame $\mathcal{F}$ in F.

4.2.2 Frames with contexts

The frames we are looking for are in fact a well investigated class of frames, namely, the class of *secondarily universal* frames.

Definition 4.1. *(Secondarily universal frames)*
A frame $\mathcal{F} = \langle W, \{R_i\}_{i \in C} \rangle$ is secondarily universal if:

- *For all $i \in C$ and $w \in W$, R_i is universal on $r_i(w)$;*
- *For all $w, w' \in W$, $r_i(w) = r_i(w')$.*

Intuitively, in secondarily universal frames every world has access via R_i to the same set of worlds, which we denote as W_i, and R_i is also universal on that set. The following representation result holds.

Proposition 4.1. *(Representation of secondarily universal frames)*
A relation R_i on W is secondarily universal iff there exists a set $W_i \subseteq W$ such that for all w, w', wR_iw' iff $w' \in W_i$.

Proof. The right to left direction is straightforward. From left to right: for every $w, w' \in W$ it holds, by the definition of function r (see Section 4.2.1), that wR_iw' iff $w' \in r_i(w)$. Since R_i is secondarily universal, it holds that for every $w, w'' \in W$, $r_i(w) = r_i(w'')$. It is now enough to stipulate $W_i = r_i(w'')$ for any w'' to obtain the desired result: there exists a set $W_i \subseteq W$ such that for all w, w', wR_iw' iff $w' \in W_i$. □

Leaving technicalities aside, Definition 4.1 forces relations in $\{R_i\}_{i\in C}$ to cluster the domain of the frame in (possibly empty) sets of worlds, one for each accessibility relation, and then defines these accessibility relations in such a way that the sets of accessible worlds correspond, for each world in W, to the clusters. Yet an easier way to express this is to say that these frames define, via $\{R_i\}_{i\in C}$, n contexts, i.e., n sets of worlds. Noticeably, this class of frames implements in a straightforward way the thesis developed in context theory according to which contexts can be soundly represented as sets of possible worlds (see [Stalnaker, 1998]), which is, in turn, just a variation on the conception of contexts as sets of models. In Chapter 2 contexts were sets of DL models, here they have become sets of propositional models.

Proposition 4.1 guarantees that secondarily universal frames can be conveniently represented replacing the family $\{R_i\}_{i\in C}$ of accessibility relations with the family of sets $\{W_i\}_{i\in C}$. This brings us to the class of Cxt frames.

Definition 4.2. (Cxt *frames)*
A Cxt *frame $\mathcal{F}$ is a structure $\mathcal{F} = \langle W, \{W_i\}_{i\in C}\rangle$, where W is a set non-empty of states and $\{W_i\}_{i\in C}$ is a family of subsets of W*[2]*.*

Models for a multi-modal language $\mathcal{L}_n$ can be built on Cxt frames in the obvious way. These are called Cxt models. The satisfaction relation results in the following.

Definition 4.3. (Satisfaction based on Cxt frames)
Let $\mathcal{M}$ be a model built on a Cxt *frame.*

$$\begin{array}{lll} \mathcal{M}, w \models [i]\phi & \textit{iff} & \forall\, w' \in W_i : \mathcal{M}, w' \models \phi \\ \mathcal{M}, w \models \langle i\rangle\, \phi & \textit{iff} & \exists w' \in W_i : \mathcal{M}, w' \models \phi. \end{array}$$

The obvious boolean clauses are omitted.

[2]We call these structures frames even though, technically speaking, they are not since they only implicitly contain accessibility relations. In effect, they are just multi-sets, or bags, of subsets of the domain W.

It is instructive to stress the shift in the semantics of the modal operators induced by Cxt frames. While in the general semantics for $\mathcal{L}_n$ languages, the truth of $[i]$ and $\langle i \rangle$ formulae depends on the state from which the formula is evaluated (see Section 4.2.1), the truth of such formulae with respect to Cxt frames abstracts from the point of evaluation: in other words truth implies validity.

Proposition 4.2. *(Truth of contextual formulae implies validity)*
Let $\mathcal{M}$ be a model built on a Cxt *frame. It holds for any $w \in W$ that:*

$$\mathcal{M}, w \models [i]\phi \quad \text{IMPLIES} \quad \mathcal{M} \models [i]\phi$$
$$\mathcal{M}, w \models \langle i \rangle \phi \quad \text{IMPLIES} \quad \mathcal{M} \models \langle i \rangle \phi$$

that is, truth of contextual formulae implies their validity.

Proof. Let us prove the first claim. Suppose $\mathcal{M}, w \models [i]\phi$ but $\mathcal{M} \not\models [i]\phi$. It follows that $\exists w' \in W_i$ such that $\mathcal{M}, w' \not\models \phi$, and therefore, by Definition 4.3, $\mathcal{M}, w \not\models [i]\phi$ which is impossible. The proof of the second claim is analogous. □

Proposition 4.2 reflects the idea that what is true or false in a context does not depend on the state of evaluation, and this is what we would intuitively expect especially for contexts interpreted as normative systems: what holds in the context of a given normative system is not determined by the point of evaluation but just by the system in itself, i.e., by its (constitutive) rules[3]. The following interesting property follows.

Corollary 4.1. *(Triviality or absurdity of contextual formulae)*
At a model level, any contextual formula $[i]\phi$ (respectively, $\langle i \rangle \phi$) is either equivalent to $\top$ or to $\bot$. In other words, contextual formulae express only global truths or falsities of a model.

Proof. Follows directly from Proposition 4.2: $\mathcal{M}, w \models [i]\phi$ iff $\forall w \in W, \mathcal{M}, w \models [i]\phi$. The same holds for $\langle i \rangle \phi$. □

Contextual formulae hold either in all worlds or in none. To put it another way, contextual formulae cannot express local properties of a model. This result will turn out useful when discussing the connection between contexts and sets of fourmlae in Section 4.3.

One last essential characteristic of Cxt frames, which marks the difference with Univ frames, consists in the fact that they do not make the T scheme ($[i]\phi \rightarrow \phi$) valid. This is indeed what one would first of all expect from a formalization of a notion of contextuality or locality via modal operators: if something holds in a context, it

[3]There nevertheless exist propositional modal-like logics modeling different conceptions of context and which do not enjoy this property, i.e., logics according to which truth and falsehood in a context do depend on the point of evaluation. Contexts of this type are, for instance, the context of the beliefs of agents: in the context of the beliefs of agent i it can be the case that the context of the beliefs of agent j does not contain a certain belief while in the context of the beliefs of agent k it does. In these cases, a feature of contextuality is to express a point of view on other contexts. Contexts of this type are investigated in [Buvač and Mason, 1993; Buvač et al., 1994].

does not necessarily hold in general. On the other hand, T remains valid in a sort of contextualized formulation. It is easy to prove that formulae $[i]([i]\phi \rightarrow \phi)$ are valid in Cxt frames.

4.2.3 Cxt frames and PRL models

It is natural at this point to spend a few words about the relation between Cxt frames and the models of propositional release logics which have been used in Section 3.3 of the previous chapter to interpret context description algebras.

That the two approaches are related is quite obvious. In both cases contexts are modeled just as sets of states from a domain W. In fact, a Cxt frame can be thought of as the incorporation at a frame level of the interpretation, by a PRL model, of a finite set C of atomic context identifiers. Let $\mathcal{I}$ be the interpretation function of a PRL model, then a Cxt frame $\mathcal{F} = \langle W, \{W_i\}_{i \in C} \rangle$ can always be built such that for every $i \in C$: $W_i = \mathcal{I}(i)$. What Cxt frames do is to fix as constant the interpretation of a finite set of 'special propositions', that is, the interpretation of the atomic context labels c.

Notice also that, as a consequence, Cxt frames could be in principle enriched in order to incorporate the algebraic structure displayed by context descriptions interpreted on PRL models. This extension is, however, of secondary importance for the analysis of counts-as.

4.2.4 Cxt models and contextual terminologies

We have thus found a class of frames which models the desired conception of context. It is worth showing now how given a finite number n of models built on universal frames, each of which satisfies possibly different sets of contextual formulae, a Cxt model can be built preserving the truth of those formulae. Intuitively, this will show that Cxt model can represent the different terminologies stated by different normative systems. The proof is simple and deals with two issues.

1. Consider a model $\mathcal{M}_1$ based on a Univ frame for the language $\mathcal{L}_1$ with the only modal operators being $[1]$ and $\langle 1 \rangle$. This is a model for **S5**, making exactly the set of contextual formulae Φ on $\mathcal{L}_1$ true, the set Φ representing a set of classifications. There always exists a super-model $\mathcal{M}_0$ of $\mathcal{M}_1$ based on a Cxt frame for $\mathcal{L}_1$ and such that it makes exactly Φ true. Apart from technicalities, this means that the terminology (in this case represented by Φ) of a normative system considered in isolation, can be represented as the set of $[1]$-formulae which are true in a $\mathcal{M}_0$ model (Proposition 4.3).

2. Consider a family $\{\mathcal{M}_i\}_{1 \leq i \leq n}$ of **S5** models for $\mathcal{L}_1$ built on Univ frames. There always exists a model $\mathcal{M}_0$ based on Cxt for $\mathcal{L}_n$ such that $\mathcal{M}_i \models \phi$ iff $\mathcal{M}_0 \models \phi'$, where ϕ is contextual and ϕ' is obtained from ϕ uniformly substituting occurrences of the $[1]$ and $\langle 1 \rangle$-operators with occurrences of $[i]$ and $\langle i \rangle$, with $i \in C$. Obviously, an appropriate set of indexes c should be chosen so that $|C| = n$. This is just the multi-modal extension of the previous result: the

terminology of each of n different normative systems can be represented as the set of $[i]$-formulae (with $i \in C$) which are true in a $\mathcal{M}_0$ model (Proposition 4.4).

Proposition 4.3. (From Univ to Cxt models)
Consider the model $\mathcal{M}_1$ for $\mathcal{L}_1$ s.t. $\mathcal{M}_1 = \langle W_1, R_1, \mathcal{I}_1 \rangle$ with R_1 universal on W_1. There always exists a super-model $\mathcal{M}_0 = \langle W_0, R_0, \mathcal{I}_0 \rangle$ for $\mathcal{L}_1$, with R_0 secondarily universal on W_0, and such that: $\mathcal{M}_1 \models \phi$ iff $\mathcal{M}_0 \models \phi$, where ϕ is a contextual *formula.*

Proof. The proposition is proven showing that it is always possible to obtain the desired structure $\mathcal{M}_0 = \langle W_0, R_0, \mathcal{I}_0 \rangle$. Model $\mathcal{M}_0$ is such that: $W_0 \supset W_1$; $R_1 = R_0 \cap W_1 \times W_1$ and $\forall w \in W_0, r_0(w) = W_1$; $\mathcal{I}_1 = \mathcal{I}_0 \lceil W_1$. The obtained structure $\mathcal{M}_0$ is, by construction, secondarily universal. It follows therefore that $\mathcal{M}_1 \models \phi$ iff $\mathcal{M}_0 \models \phi$, with ϕ contextual. □

Proposition 4.4. (From many Univ models to one Cxt model)
Consider now a set of models $\mathcal{M}_1, \ldots, \mathcal{M}_n$ for $\mathcal{L}_1$ on Univ *frames and a language $\mathcal{L}_n$ such that $|C| = n$. Then there always exists a model $\mathcal{M}_0$ for $\mathcal{L}_n$ on a* Cxt *frame s.t. $\mathcal{M}_0 = \langle W_0, \{R_{0i}\}_{i \in C}, \mathcal{I}_0 \rangle$ on $\mathcal{L}_n$ and $\mathcal{M}_i \models \phi$ iff $\mathcal{M}_0 \models \phi'$, where ϕ is contextual and ϕ' is obtained from ϕ uniformally substituting occurrences of the* $[1]$ *and* $\langle 1 \rangle$*-operators with occurrences of* $[i]$ *and* $\langle i \rangle$*, for every $i \in C$.*

Proof. Model $\mathcal{M}_0$ can be obtained applying, for model $\mathcal{M}_i = \langle W_i, R_i, \mathcal{I}_i \rangle$ the construction used in proving Proposition 1: $W_0 \supseteq \bigcup_{i \in C} W_i$; $R_i = R_{0i} \cap W_i \times W_i$; $\mathcal{I}_i = \mathcal{I}_0 \lceil W_i$. □

Proposition 4.3 and 4.4 prove that secondarily universal frames are in fact the structures we are looking for since they can represent a number of universal relations at the same time. By Proposition 4.1 this result directly applies also to Cxt models.

It is worth noticing that Cxt frames cannot just represent models for sets of **S5** contextual formulae, but they can also represent that such models do not exist. They can represent the absence of models for a set of subsumptions, i.e., the empty context. This is the case for frames where for a given $i \in C$, $W_i = \emptyset$. This is an interesting expressive feature allowing for capturing, within the framework, also the notion of an inconsistent classification. In our setting this simply amounts to accept the possibility of normative systems issuing inconsistent terminologies[4].

4.2.5 Axiomatics

The natural question is now: is there a logic characterizing the class of Cxt frames? The answer is positive and the system at issue corresponds to logic $\mathbf{K45^{ij}_n}$. It is

[4]In [Grossi et al., 2005d] we ruled this possibility out considering only Cxt frames with non-empty contexts. Those frames yield logic $\mathbf{KD45^{ij}_n}$, which is also a well investigated system (see [Nayak, 1994; Lomuscio and Sergot, 2003]).

axiomatized via the following axioms and rules schemata:

(P)	all tautologies of propositional calculus
(K)	$[i](\phi_1 \rightarrow \phi_2) \rightarrow ([i]\phi_1 \rightarrow [i]\phi_2)$
(4^{ij})	$[i]\phi \rightarrow [j][i]\phi$
(5^{ij})	$\neg[i]\phi \rightarrow [j]\neg[i]\phi$
(Dual)	$\langle i\rangle \phi \leftrightarrow \neg[i]\neg\phi$
(MP)	IF $\vdash \phi_1$ AND $\vdash \phi_1 \rightarrow \phi_2$ THEN $\vdash \phi_2$
(N^i)	IF $\vdash \phi$ THEN $\vdash [i]\phi$

where i, j denote elements of the set of indexes C. The system is a multi-modal homogeneous **K45** with the two interaction axioms 4^{ij} and 5^{ij}.

A remark is in order here especially with respect to axiomata 4^{ij} and 5^{ij}. In fact, what the two schemata do, consists in making the nesting of the operators reducible which, leaving technicalities aside, means that truth and falsehood in contexts ($[i]\phi$ and $\neg[i]\phi$) are somehow absolute because they remain invariant even if evaluated from another context ($[j][i]\phi$ and $[j]\neg[i]\phi$). In other words, they express the fact that whether something holds in a context i is not something that a context j can influence. This is nothing but the syntactic counterpart of the property we discussed in relation with Definition 4.3 in Section 4.2.2.

Logic $\mathbf{K45_n^{ij}}$ is sound and complete with respect the class of CXT frames. This result is proven in Appendix A. This logic is thus the candidate to provide a logical characterization of the notion of counts-as as it was exposed in Section 4.1.

At this stage the logical machinery is semantically and syntactically worked out and it can be put to work.

4.2.6 Classificatory counts-as formalized

Using a multi-modal logic $\mathbf{K45_n^{ij}}$ on a language $\mathcal{L}_n$, the classificatory view on counts-as statements can be given the following formal characterization.

Definition 4.4. (Classificatory counts-as: $\Rightarrow_c^{cl}$)
"γ_1 counts as γ_2 in context c", with γ_1 and γ_2 objective formulae, is formalized in a multi-modal language $\mathcal{L}_n$ as the strict implication in logic $\mathbf{K45_n^{ij}}$:

$$\gamma_1 \Rightarrow_c^{cl} \gamma_2 \ := \ [c](\gamma_1 \rightarrow \gamma_2)$$

Notice that the definition constrains the counts-as conditionals $\Rightarrow_c^{cl}$ to hold only between objective formulae. This limitation is motivated by viewing counts-as statements as concept subsumption statements (i.e., $X \rightarrow Y$) which are contextualized as a whole (i.e., $[c](X \rightarrow Y)$). Hence, the formulae occurring in counts-as statements are just boolean compounds and do not contain any modal flavor since what is contextualized is the implication as a whole. These considerations have also a more technical side. Classificatory counts-as statements concern the subsumption of local properties, i.e., those typically expressed by objective formulae, while contextualized formulae denote global ones (Proposition 4.2).

These are some of the resulting properties of $\Rightarrow_c^{cl}$.

Proposition 4.5. (Properties of $\Rightarrow_c^{cl}$)
The following formulae and rules are valid in Cxt *frames:*

$$\gamma_2 \leftrightarrow \gamma_3 \ / \ (\gamma_1 \Rightarrow_c^{cl} \gamma_2) \leftrightarrow (\gamma_1 \Rightarrow_c^{cl} \gamma_3) \tag{4.1}$$

$$\gamma_1 \leftrightarrow \gamma_3 \ / \ (\gamma_1 \Rightarrow_c^{cl} \gamma_2) \leftrightarrow (\gamma_3 \Rightarrow_c^{cl} \gamma_2) \tag{4.2}$$

$$((\gamma_1 \Rightarrow_c^{cl} \gamma_2) \wedge (\gamma_1 \Rightarrow_c^{cl} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow_c^{cl} (\gamma_2 \wedge \gamma_3)) \tag{4.3}$$

$$((\gamma_1 \Rightarrow_c^{cl} \gamma_2) \wedge (\gamma_3 \Rightarrow_c^{cl} \gamma_2)) \rightarrow ((\gamma_1 \vee \gamma_3) \Rightarrow_c^{cl} \gamma_2) \tag{4.4}$$

$$\gamma \Rightarrow_c^{cl} \gamma \tag{4.5}$$

$$(\gamma_1 \Rightarrow_c^{cl} \gamma_2) \wedge (\gamma_2 \Rightarrow_c^{cl} \gamma_3) \rightarrow (\gamma_1 \Rightarrow_c^{cl} \gamma_3) \tag{4.6}$$

$$(\gamma_1 \Rightarrow_c^{cl} \gamma_2) \wedge (\gamma_2 \Rightarrow_c^{cl} \gamma_1) \rightarrow [c](\gamma_1 \leftrightarrow \gamma_2) \tag{4.7}$$

$$(\gamma_1 \Rightarrow_c^{cl} \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \Rightarrow_c^{cl} \gamma_2) \tag{4.8}$$

$$(\gamma_1 \Rightarrow_c^{cl} \gamma_2) \rightarrow (\gamma_1 \Rightarrow_c^{cl} \gamma_2 \vee \gamma_3) \tag{4.9}$$

Proof. We provide only the deduction of Formula 4.8 as an example:

$$\begin{array}{lcl}
1. & (\mathtt{P}) & (\gamma_1 \rightarrow \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2) \\
2. & (\mathtt{N}), 1 & [c]((\gamma_1 \rightarrow \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)) \\
3. & (\mathtt{K}) & [c]((\gamma_1 \rightarrow \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)) \\
 & & \rightarrow [c](\gamma_1 \rightarrow \gamma_2) \rightarrow [c](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2) \\
4. & (\mathtt{MP}), 3, 1 & [c](\gamma_1 \rightarrow \gamma_2) \rightarrow [c](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2) \\
5. & (\text{Def. } 4.4), 4 & (\gamma_1 \Rightarrow_c^{cl} \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \Rightarrow_c^{cl} \gamma_2)
\end{array}$$

All other proofs are just as straightforward via application of Definition 4.4 and using the given axiomatization of $\mathbf{K45_n^{ij}}$. □

This system validates all the intuitive syntactic constraints isolated in [Jones and Sergot, 1996] (Formulae 4.1-4.4). Besides, the analysis shows that counts-as conditionals, once they are viewed as conditionals of a classificatory nature, naturally satisfy reflexivity (Formula 4.5), transitivity (Formula 4.6), a form of "contextualized" antisymmetry (Formula 4.7), strengthening of the antecedent (Formula 4.8) and weakening of the consequent (Formula 4.9). We will systematically come back to the discussion of these structural properties in Section 4.8.2.

4.3 Counts-as Beyond Contextual Classification

The previous Section has provided a formal analysis of the classificatory view of counts-as (Definition 4.4), i.e., counts-as statements as implications holding in a set of states, explicating what logical properties are to be accepted once such an analytical option on the semantics of counts-as is assumed (Proposition 4.5). The classificatory view does not exhaust, however, the whole spectrum of readings of what counts-as statements mean.

In particular, the interpretation of counts-as in merely classificatory terms does not do justice to the notion which is stressed in the label "constitutive rule", that is, the notion of *constitution*. Rules are said to be constitutive in at least two distinct senses:

1. In the sense that they bring about something new, they introduce classifications which are not already considered to be hold. In this view counts-as statements express sorts of strong, or proper, classifications. This particular meaning of counts-as statements is further discussed in Section 4.3.1 under the name of *proper classificatory counts-as* and formally analyzed in Sections 4.3.3 and 4.4.

2. In the sense that they define the normative system, or institution, to which they pertain. This sense of constitution corresponds to the idea that a constitutive rule is a necessary condition for the existence of normative system it contributes to define. For instance, the game chess is constituted by its rules:

   > "Chess is the game it is in virtue of all its rules" ([Wittgenstein, 1958], §197).

   Without the rules of chess, the game chess would not be conceivable. This acception of the term constitution is further expanded upon in Section 4.3.2 under the name of *constitutive counts-as* and formally analyzed in Sections 4.3.4 and 4.5.

To put it in a few words, the distinction amounts to the difference between something being constituted, that is, something as being the *result of constitution* (1), and something *constituting* a normative system or a context (2).

4.3.1 Counts-as statements as proper classifications

The analytic literature on constitutive norms often comes to emphasize the following feature: counts-as statements are not just classifications but "new" classifications, that is, classifications which would not hold without the normative system stating them:

> "Where the rule is purely regulative, behaviour which is in accordance with the rule could be given the same description or specification (the same answer to the question "What did he do?") whether or not the rule existed, provided the description or specification makes no explicit reference to the rule. But where the rule (or system of rules) is constitutive, behaviour which is in accordance with the rule can receive specifications or descriptions which it could not receive if the rule did not exist" ([Searle, 1969], p.35).

The novelty of the classifications stated via counts-as is stressed by various different names that the statements based on constitutive rules have obtained in the literature. For instance, it is said that the occurrence of X "conventionally generates" the occurrence of Y ([Goldman, 1976]) given the existence of a corresponding rule, or that the occurrence of Y "supervenes" on the occurrence of X ([Hage and Verheij, 1999]).

In this view, counts-as statements do not only state contextual classifications, but they state new classifications which would not otherwise hold.

Observation 4.1. *Counts-as statements are classifications which hold with respect to a context (set of situations) but which do not hold in general (i.e., with respect to all situations).*

We call counts-as statements intended in the sense of Observation 4.1 *proper contextual classifications*. In other words, X counts as Y in context C because X is classified as Y in C but also because this does not hold in general, i.e., in the global context. In this sense the notion of proper contextual classification captures the idea, of rules denoting the bringing about of something new. The classification is brought about, or constituted, by the normative system itself which specifies the context of the classification, and it would not hold without it.

In this view, it is obvious to expect that from "X counts as Y" it does not follow, for instance, that $X \wedge Z$ counts as Y, because it might well be that to obtain Y from $X \wedge Z$ no constitutive rule is needed at all. To provide a simple example: given that in the context of the animal ethics movement all animals count as (in a proper classificatory sense) individuals bearing rights, it does not follow that all animals which are humans count as (in a proper classificatory sense) individuals bearing rights, because the fact that humans are individuals bearing rights is a global ethical truth common to all ethical systems, which would therefore also hold without the animal ethics movement.

For analogous reasons, it does not seem intuitive to say, in the view of proper contextual classification, that X counts as X, since "X is X" is just a tautology, or that if X counts as Y, then X counts as Y or Z. In the light of these considerations, Formulae 4.5, 4.8 and 4.9 in Proposition 4.5 should turn out to be invalid in a formal characterization of proper contextual classification.

4.3.2 Counts-as statements as constitutive rules

Consider the following inference: it is a rule of normative system Γ that conveyances transporting people or goods count as vehicles; it is always the case that bikes count as conveyances transporting people or goods but not that bikes count as vehicles; therefore, according to normative system Γ, bikes count as vehicles. This is an instance of a typical normative reasoning pattern: from the rule of a given normative system and a common-sense fact, another fact is inferred which holds with respect to that normative system. The count-as locution occurs three times. However, the second premise states a generally acknowledged classification, and the conclusion states a "new" classification which is considered to hold with respect to the given normative system. The second premise is a contextual classification concerning a global or universal context, the conclusion is a proper contextual classification in the sense clarified in the previous section. What about the first premise? Obviously, it also expresses a classification which is brought about by the normative system, a proper contextual classification thus. There is however something more. It explicitly states a constitutive rule of a normative system: "conveyances transporting people or goods are classified as vehicles" is one of the rules of Γ. This semantic ingredient is not captured by the notions of contextual and proper contextual classification. It involves two essential aspects.

The first one, as already noticed, is that counts-as statements of this type are always part of a *set* of similar statements, a set of rules.

> "Rules are constitutive if and only if they are part of a set of rules. Strictly speaking, there is no such thing as *a* rule that is constitutive in isolation" ([Ricciardi, 1997], p.5).

In other words, constitutive are only sets of rules. Hence, a constitutive rule cannot be isolated from the set to which it belongs, since it is constitutive only in as much it is part of that set. It is worth stressing how close this consideration lies to the warning raised in [Makinson, 1999]: "no logic of norms without attention to a system of which they form part". Constitutivity is in essence a matter of sets of rules.

The second aspect concerns the relation between, on the one hand, the notion of a set of rules Γ, i.e., normative system or institution, and on the other hand the notion of set of situations c, or context c. The set of classifications stated as constitutive rules by a normative system (for instance, "conveyances transporting people or goods count as vehicles") can be seen as exactly the set of situations which make that set of classifications true. Hence, the set of constitutive rules of any normative system can be seen as a set of situations. And a set of situations (or states, or models) —we have seen in the previous chapters— is what is called a context in much literature on context theory. To put it in a nutshell, a context is a set of situations, and if the constitutive rules of a given normative system Γ are satisfied by all and only the situations in a given set, then that set of situations is *the context defined by* Γ. This simple observation allows us to think of contexts as "systems of constitutive rules" ([Searle, 1969], p.51). Notice that this is no exotic thought. In fact, this idea has been neatly advanced —informally— in some literature on the theory of institutions:

> "A set of constitutive rules defines a logical space" ([Ricciardi, 1997], p.6).

A logical space is, needless to say, nothing but a set of states, i.e., a context. Getting back to the above example: the statement "according to Γ, bikes count as vehicles" is read as "in the set of situations defined by the rules of system Γ, `bike` is a subconcept of `vehicle`".

Counts-as statements used to express constitutive rules have a different meaning from the counts-as statements which are instead used for expressing what follows from the existence of a constitutive rule. We call the first ones *constitutive counts-as* *statements*, while the second ones are *proper classificatory counts-as statements* if they stress the classification to be something brought about by the context, or simply *classificatory counts-as statements* if they just denote the fact that the classification holds in a given context. In other words, when statements "X counts as Y in the context c of normative system Γ" are read as constitutive rules, what is meant is that the classification of X under Y is considered to be an explicit promulgation of the normative system Γ defining context c. Instead, when they are read as proper classificatory statements they are meant to denote classifications that are constituted, or brought about, by the context at issue. Finally, when they are read as mere contextual classification, they are meant to denote classificatory statements that are just the case in the given context.

"In normative system Γ, conveyances transporting people or goods count as vehicles"	**Constitutive**
"It is always the case that bikes count as conveyances transporting people or goods"	**Classificatory**
"In normative system Γ bikes count as vehicles"	**Proper Classificatory**

The discussion above is recapitulated in the following observation which will be used as guideline for the formal analysis to follow.

Observation 4.2. *A constitutive counts-as statement is a proper contextual classification such that: (a) it is an element of the set of rules specifying a given normative system* Γ*; (b) the set of rules of* Γ *defines the context (set of situations) to which the counts-as statement pertains.*

Before proceeding with the formal analysis, it is worth noting that some literature on legal theory considers counts-as statements to be special kinds of constitutive rules and rejects a full identification between constitutive rules and counts-as statements. For example, [Sartor, 2006] considers counts-as statements to typically concern the constitution of state-of-affairs which have no duration (e.g., committing a crime) while constitutive rules concern the constitution of state-of-affairs with duration, i.e., which can start and cease to hold (e.g., being a citizen). This is of course a terminological matter, and we chose to solve it by sticking to the Searlean view, where the identification "constitutive rule = counts-as" is quite clearly stated. Besides, it should also be said that, in order to introduce such a distinction between counts-as and constitutive rules, distinctions should also be introduced which allow to distinguish the specific nature of the X and Y terms occurring in the rules. The propositional logic setting assumed here abstracts from such distinctions by viewing X and Y simply as propositions whose further logical structure is left unspecified.

4.3.3 Getting formal: from classification to proper classification

As usual, model-theoretic considerations can give us crucial hints to formalize our intuitions. Let us define the set $\mathbb{T}(X)$ of all formulae which, given a model, are satisfied by all worlds in a set of worlds X:

$$\mathbb{T}(X) = \{\phi \mid \forall w \in X : \mathcal{M}, w \models \phi\}.$$

and let $\mathbb{T}^{\rightarrow}(X)$ be the set of all implications between objective formulae γ_1 and γ_2 which are satisfied by all worlds in a set of worlds X:

$$\mathbb{T}^{\rightarrow}(X) = \{\gamma_1 \rightarrow \gamma_2 \mid \forall w \in X : \mathcal{M}, w \models \gamma_1 \rightarrow \gamma_2\}.$$

Obviously, for every X: $\mathbb{T}^{\rightarrow}(X) \subseteq \mathbb{T}(X)$. In the classificatory reading, given a model $\mathcal{M}$ where the set of worlds $W_c \subseteq W$ models context c, the set of all classificatory counts-as statements holding in c, which we denote as $\mathbb{CL}(W_c)$, can be defined as the set $\mathbb{T}^{\rightarrow}(W_c)$:

$$\mathbb{CL}(W_c) := \mathbb{T}^{\rightarrow}(W_c).$$

Hence, it is easy to see that: $\mathbb{T}^{\rightarrow}(W) \subseteq \mathbb{CL}(W_c) \subseteq \mathbb{T}(W_c)$. In other words, the set of classificatory counts-as statements is:

- A subset of all the truths of W_c;
- A superset of all conditional truths of W, that is, of the "global" or "universal" context of model $\mathcal{M}$.

While the first point represents a quite banal semantic constraint to which any formal characterization of counts-as should adhere, the second one is much more questionable. Indeed, what is true anyway is not characteristic of any context (except of the global one), and it cannot be properly said to represent any new truth. In other words, interpreting counts-as as statements as mere classifications, as it has been done in Section 4.2, make them inherit all trivial classifications which hold globally in the model.

These considerations suggest a readily available strategy to specify the set of proper classificatory counts-as holding in a context c on the basis of $\mathbb{T}^{\rightarrow}(W_c)$. The problem boils down to eliminate from the set of classificatory counts-as $\mathbb{CL}$ for a context W_c those classifications which hold globally, that is, which hold with respect to the global context W. We obtain, in this way, the set of what we call *proper classificatory counts-as* statements, or *proper contextual classifications*, holding in context c in a Cxt model $\mathcal{M}$.

Definition 4.5. (Set of proper classificatory counts-as in c)
The set $\mathbb{CL}^{+}(W_c)$ of proper classificatory counts-as statements of a context c in a Cxt *model $\mathcal{M}$ is defined as follows:*

$$\mathbb{CL}^{+}(W_c) := \mathbb{T}^{\rightarrow}(W_c) \setminus \mathbb{T}(W). \tag{4.10}$$

Intuitively, the set of proper classificatory count-as holding in c corresponds to the set of implications between objective formulae which hold in c, minus those implications which hold universally. Or, to put it otherwise, the set of proper classificatory count-as holding in c corresponds to the set of classificatory counts as of c, minus those implications which hold universally: $\mathbb{CL}^+(W_c) := \mathbb{CL}(W_c) \setminus \mathbb{T}(W)$. This can be seen as the most natural amendment of the classificatory view toward the specification of a stronger notion of contextual classification along the lines of Observation 4.1. Section 4.4 is devoted to a detailed analysis of this interpretation of counts-as statements, which, we argue in Section 4.8.2, highly overlaps with the view of counts-as maintained in [Jones and Sergot, 1996].

4.3.4 Getting formal: from proper classification to constitution

Let us now focus on Observation 4.2. What comes to play a role is the notion of the *definition* of the context of a counts-as statement. A definition of a context c, in a Cxt model $\mathcal{M}$, is a set of objective formulae Γ such that $\forall w \in W$:

$$\mathcal{M}, w \models \Gamma \text{ iff } w \in W_c \tag{4.11}$$

that is, the set of formulae Γ such that all and only the worlds in W_c satisfy Γ in $\mathcal{M}$. Notice that we limit definitions to set of objective formulae. This is no arbitrary choice and it is grounded on Proposition 4.2 and Corollary 4.1: contextual formulae are irrelevant for the definition of sets of worlds W_i such that $\emptyset \subset W_i \subset W$, that is, those sets which denote neither the empty nor the universal contexts.

Observation 4.2 can now get a formal formulation. Given the set of formulae Γ, we say that any formula $\gamma_1 \to \gamma_2 \in \Gamma$ is a constitutive counts-as statement w.r.t. context c iff Γ defines context c and $\gamma_1 \to \gamma_2$ belongs to the set of proper contextual classifications of c.

Definition 4.6. (Set of constitutive counts-as in c w.r.t. definition Γ)
The set $\mathbb{CO}(\Gamma, W_c)$ of constitutive counts-as statements of a context c defined by Γ in a Cxt *model $\mathcal{M}$ is:*

$$\mathbb{CO}(\Gamma, W_c) := \{\gamma_1 \to \gamma_2 \in \Gamma \mid \gamma_1 \to \gamma_2 \in \mathbb{CL}^+(W_c) \textit{ and } \forall w(\mathcal{M}, w \models \Gamma \textit{ iff } w \in W_c)\} \tag{4.12}$$

Set $\mathbb{CO}(\Gamma, W_c)$ is defined taking as domain the set of implicative statements of Γ. Notice also that as a result of this definition if Γ does not define context W_c then $\mathbb{CO}(\Gamma, W_c) = \emptyset$. In fact, Formula 4.12 can be restated as follows:

$$\mathbb{CO}(\Gamma, W_c) = \begin{cases} \mathbb{CL}^+(W_c) \cap \Gamma, & \text{if } \Gamma \text{ defines } W_c \\ \emptyset, & \text{otherwise.} \end{cases}$$

Section 4.5 is devoted to the development of a modal logic based on this definition and to a detailed analysis of this interpretation of counts-as statements.

The definitions discussed are summarized in the table below.

Cxt Classification	$\mathbb{CL}(W_c) \;=\; \mathbb{T}^{\rightarrow}(W_c)$
Proper Cxt Classification	$\mathbb{CL}^{+}(W_c) \;=\; \mathbb{CL}(W_c) \setminus \mathbb{T}(W)$
Constitution	$\mathbb{CO}(\Gamma, W_c) \;=\; \begin{cases} \mathbb{CL}^{+}(W_c) \cap \Gamma, & \text{if } \Gamma \text{ defines } W_c \\ \emptyset, & \text{otherwise.} \end{cases}$

The table pinpoints the dependencies between the formal characterizations of the three different senses of counts-as which has been taken into consideration: the notion of constitution builds on the notion of proper contextual classification which, in turn, builds on the notion of contextual classification. The modal logic analysis of contextual classification developed in Section 4.2 can thus be used as a sound starting point for the modal logic analysis of the two notions introduced in this section.

4.3.5 A methodological note

Before rendering the insights of Sections 4.3.3 and 4.3.4 in modal logic, it is worth making a methodological remark. We are here concerned with a term, "counts-as", which appears to have different meanings. At this point we had two main ways to pursue the formal characterization of counts-as we were aiming at. We could proceed axiomatically by trying to single out intuitive syntactic properties of counts-as statements. Or rather semantically, by trying to enrich the semantic characterization of classificatory counts-as exposed in the previous sections in order to capture further semantic nuances. The latter has been our choice, which is inspired by how some fundamental work in philosophical logic has been done in the past.

In particular, it is instructive to recall that this exact same concern lies also at the ground of the Tarskian characterization of the notion of truth. Because of the inherent polysemy of the predicate "to be true", Tarski found it unconvincing to proceed introducing the predicate as a primitive and then axiomatizing it:

> "[. . .] the choice of axioms always has rather accidental character, depending on inessential factors (such as e.g. the actual state of our knowledge). [. . .] a method of constructing a theory does not seem to be very natural [. . .] if in this method the role of primitive concepts —thus of concepts whose meaning should appear evident— is played by concepts which have led to various misunderstanding in the past" ([Tarski, 1983], pag. 405-406).

Instead, he preferred to first isolate a precise sense of the predicate, i.e., truth as correspondence to reality, and then to define it in terms of a better understood notion, i.e., the notion of satisfaction of a formula by a model.

The work presented in this chapter can be viewed as an application of this method to the notion of counts-as: in Sections 4.1 and 4.2 we first committed to one precise sense of the term (classificatory counts-as), which has then been characterized in terms of a better-understood notion (strict implication within a context); in this section we have isolated two more meanings of the term "counts-as", as they appear in some work in social and legal theory, and we have formally characterized them by making use of better-understood logical notions: the negation of global statements (proper classificatory counts-as) and the definition of a context (constitutive counts-as).

4.4 Counts-as as Proper Contextual Classification

In the following section a modal logic for proper contextual classification is developed which is based on Definition 4.5. By doing this we will capture the intuitions discussed in Section 4.3 concerning the intuitive reading of counts-as statements in proper classificatory terms. At the same time we will maintain the possible worlds semantics of context exposed in Section 4.2 and developed in order to account for the purely classificatory view of counts-as. It will therefore be possible to represent both the investigated senses of counts-as within the same framework, and to account for their logical relations.

4.4.1 Expansion of $\mathcal{L}_n$ and semantics

Language $\mathcal{L}_n$ is expanded as follows. The set of context indexes C is such that it always contains the special context index u denoting the universal (or global) context. We call this language $\mathcal{L}_n^u$.

Languages $\mathcal{L}_n^u$ are given a semantics via a special class of Cxt frames, namely the class of Cxt frames $\mathcal{F} = \langle W, \{W_i\}_{i \in C} \rangle$ such that $W \in \{W_i\}_{i \in C}$. That is, the frames in this class, which we call $\text{Cxt}^\top$, always contain the global context among their contexts. The definition of the satisfaction relation for language $\mathcal{L}_n^u$ follows.

Definition 4.7. (Satisfaction based on $\text{Cxt}^\top$ frames)
Let $\mathcal{M}$ be a model built on a $\text{Cxt}^\top$ *frame.*

$$\begin{array}{lll} \mathcal{M}, w \models [u]\phi & \textit{iff} & \forall\, w' \in W : \mathcal{M}, w' \models \phi \\ \mathcal{M}, w \models [c]\phi & \textit{iff} & \forall\, w' \in W_c : \mathcal{M}, w' \models \phi \end{array}$$

where u is the universal context index and c ranges on the context indexes in C. The obvious boolean clauses and the clauses for the dual modal operators are omitted.

The new clause states that the $[u]$ operator is interpreted on the universal frame contained in each $\text{Cxt}^\top$ frame. It is therefore nothing but an **S5** necessity operator.

4.4.2 Axiomatics

We call $\mathbf{Cxt^u}$ the logic characterizing the class of $\text{Cxt}^\top$ frames. Logic $\mathbf{Cxt^u}$ results from the union $\mathbf{K45_n^{ij}} \cup \mathbf{S5_u} \cup \{$ axiom $(\subseteq .ui)\}$, that is, from the union of $\mathbf{K45_n^{ij}}$ with the $\mathbf{S5_u}$ logic for the $[u]$ operator together with the interaction axiom $\subseteq .ui$ below. The axiomatics runs thus as follows:

(P)	all tautologies of propositional calculus
($\mathtt{K}^i$)	$[i](\phi_1 \rightarrow \phi_2) \rightarrow ([i]\phi_1 \rightarrow [i]\phi_2)$
(4^{ij})	$[i]\phi \rightarrow [j][i]\phi$
(5^{ij})	$\neg[i]\phi \rightarrow [j]\neg[i]\phi$
($\mathtt{T}^u$)	$[u]\phi \rightarrow \phi$
($\subseteq .ui$)	$[u]\phi \rightarrow [i]\phi$
(Dual)	$\langle i \rangle \phi \leftrightarrow \neg[i]\neg\phi$
(MP)	IF $\vdash \phi_1$ AND $\vdash \phi_1 \rightarrow \phi_2$ THEN $\vdash \phi_2$
($\mathtt{N}^i$)	IF $\vdash \phi$ THEN $\vdash [i]\phi$

where i, j denote elements of the set of indexes C and u denotes the universal context index in C. The interaction axiom $\subseteq .ui$ states something quite intuitive concerning the interaction of the $[u]$ operator with all other context operators: what holds in the global context, holds in every context. Soundness and completeness of this axiomatization w.r.t. $\textsc{Cxt}^{\top}$ frames are proven in Appendix A.

4.4.3 Proper classificatory counts-as formalized

Using a multi-modal logic $\mathbf{Cxt^u}$ on a language $\mathcal{L}_n^u$, the proper classificatory reading of counts-as statements can be formalized as follows.

Definition 4.8. (Proper classificatory counts-as: $\Rightarrow_c^{cl+}$)
"γ_1 counts as γ_2 in context c", with γ_1 and γ_2 objective formulae, is formalized in the logic $\mathbf{Cxt^u}$ *on a multi-modal language $\mathcal{L}_n^u$ as:*

$$\gamma_1 \Rightarrow_c^{cl+} \gamma_2 \ := \ [c](\gamma_1 \rightarrow \gamma_2) \wedge \neg[u](\gamma_1 \rightarrow \gamma_2)$$

Notice that this definition is nothing but the translation in the $\mathcal{L}_n^u$ language of Formula 4.10 in Definition 4.5.

What properties of counts-as are lost interpreting it as proper contextual classification? And what properties are instead still valid? The following two propositions answer these questions.

Proposition 4.6. (Properties of $\Rightarrow_c^{cl+}$: invalidities)
The $\Rightarrow_c^{cl+}$ versions of reflexivity, strengthening of the antecedent, weakening of the consequent, transitivity and cautious monotonicity are not valid, that is, the following formulae

are invalid in $\mathrm{CXT}^{\top}$ *frames:*

$$\gamma \Rightarrow^{cl+}_{c} \gamma \tag{4.13}$$

$$(\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \Rightarrow^{cl+}_{c} \gamma_2) \tag{4.14}$$

$$(\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \rightarrow (\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2 \vee \gamma_3) \tag{4.15}$$

$$((\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \wedge (\gamma_2 \Rightarrow^{cl+}_{c} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow^{cl+}_{c} \gamma_3) \tag{4.16}$$

$$((\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \wedge (\gamma_1 \Rightarrow^{cl+}_{c} \gamma_3)) \rightarrow ((\gamma_1 \wedge \gamma_2) \Rightarrow^{cl+}_{c} \gamma_3) \tag{4.17}$$

Proof. The invalidity of Formula 4.13 can be proven via reductio ad absurdum. Assume $\gamma \Rightarrow^{cl+}_{c} \gamma$. Via Definition 4.8 it follows that: $[c](\gamma \rightarrow \gamma) \wedge \neg[u](\gamma \rightarrow \gamma)$, which is impossible, $\gamma \rightarrow \gamma$ being a tautology.
That Formula 4.14 is not valid, can be proven showing a countermodel $\mathcal{M}$ such that: there exists a w such that $\mathcal{M}, w \models [c](\gamma_1 \rightarrow \gamma_2)$ and $\mathcal{M}, w \not\models [u](\gamma_1 \rightarrow \gamma_2)$ (antecedent true) and $\mathcal{M}, w \models [c](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)$ and $\mathcal{M}, w \models [u](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)$ (consequent false). For the semantics exposed in Section 4.4.1, the desired countermodel is thus a model $\mathcal{M}$ such that: $\forall w \in W$: $\mathcal{M}, w \models \gamma_1 \wedge \gamma_3 \rightarrow \gamma_2$, $\exists w' \in W$: $\mathcal{M}, w' \models \gamma_1 \wedge \neg\gamma_2$, and $\forall w \in W_c$: $\mathcal{M}, w \models \gamma_1 \rightarrow \gamma_2$. Formula 4.15 has a similar countermodel.
As to transitivity (Formula 4.16), a countermodel can be found which looks like this: $\forall w \in W$, $\mathcal{M}, w \models \gamma_1 \rightarrow \gamma_3$; $\forall w \in W_c$, $\mathcal{M}, w \models \gamma_1 \rightarrow \gamma_2$ and $\mathcal{M}, w \models \gamma_2 \rightarrow \gamma_3$; and $\exists w', w''$ s.t. $\mathcal{M}, w' \models \gamma_1 \wedge \neg\gamma_2 \wedge \gamma_3$ and $\mathcal{M}, w'' \models \neg\gamma_1 \wedge \gamma_2 \wedge \neg\gamma_3$.
Cautious monotonicity (Formula 4.17) has a similar countermodel: $\forall w \in W$, $\mathcal{M}, w \models \gamma_1 \wedge \gamma_2 \rightarrow \gamma_3$; $\forall w \in W_c$, $\mathcal{M}, w \models \gamma_1 \rightarrow \gamma_2$ and $\mathcal{M}, w \models \gamma_1 \rightarrow \gamma_3$; and $\exists w', w''$ s.t. $\mathcal{M}, w' \models \gamma_1 \wedge \neg\gamma_2 \wedge \gamma_3$ and $\mathcal{M}, w'' \models \gamma_1 \wedge \neg\gamma_2 \wedge \neg\gamma_3$. □

We have given in Section 4.3.1 an intuitive example showing why the strengthening of the antecedent fails for the notion of proper contextual classification. It might be instructive to provide, at this point, also an intuitive example for the failure of transitivity. We will provide two.

The first is very simple. Consider a public park regulation stating that self-propelled conveyances counts as (in the proper classificatory sense) vehicles, and that vehicles count as (in the proper classificatory sense) self-propelled conveyances. It follows that self-propelled conveyances counts as self-propelled conveyances, but this time, the counts-as can only be read in the classificatory sense. In fact, that being a self-propelled conveyances implies being a self-propelled conveyances is a logical, and therefore global, truth.

More interesting examples of countermodels of transitivity can be found in the legal domain. They typically arise by considering the changes in a legal system that the promulgation of new laws brings about. Up to 1982 the Italian legal code classified associations of a mafia kind under the type "criminal association" ("associazione a delinquere") and therefore under the type "crime against the public order" ("delitti contro l'ordine pubblico"). As a consequence, it was a global truth of the Italian legal code before 1982 that associations of a mafia kind were classified as crime against the public order. In 1982 (Article 416bis), the new type "mafia association" ("associazione mafiosa") was introduced in order to directly classify associations of a mafia kind and distinguish them from the more generic type of "criminal association". On the other hand, "mafia association" was also introduced as a subtype of "crime against the public order". The introduction of this article determines that, in the context of the legal code of 1982, "associations of a mafia kind count as mafia associations" and "mafia associations count as crimes against the legal order" are both proper classificatory counts-as statements since they both introduce something new with respect to what considered to be already the case. On the contrary, "associations of a mafia kind count as crimes against the legal order" is just a classificatory counts-as since that statement is already a global true of the code. To put it shortly, transitivity fails any time local middle terms are constituted within classifications which are taken to hold globally in the model.

Proposition 4.7. (Properties of $\Rightarrow_c^{cl+}$: validities)
In logic $\mathbf{Cxt^u}$ *the* $\Rightarrow_c^{cl+}$ *variants of Formulae 4.1-4.4 of Proposition 4.5 are valid in* $\mathrm{CXT}^{\top}$ *frames:*

$$\gamma_2 \leftrightarrow \gamma_3 \;/\; (\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \leftrightarrow (\gamma_1 \Rightarrow_c^{cl+} \gamma_3) \tag{4.18}$$

$$\gamma_1 \leftrightarrow \gamma_3 \;/\; (\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \leftrightarrow (\gamma_3 \Rightarrow_c^{cl+} \gamma_2) \tag{4.19}$$

$$((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \wedge (\gamma_1 \Rightarrow_c^{cl+} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow_c^{cl+} (\gamma_2 \wedge \gamma_3)) \tag{4.20}$$

$$((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \wedge (\gamma_3 \Rightarrow_c^{cl+} \gamma_2)) \rightarrow ((\gamma_1 \vee \gamma_3) \Rightarrow_c^{cl+} \gamma_2) \tag{4.21}$$

Contextualized antisymmetry, i.e., is valid in the following form:

$$(\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \wedge (\gamma_2 \Rightarrow_c^{cl+} \gamma_1) \rightarrow [c](\gamma_1 \leftrightarrow \gamma_2) \wedge \neg[u](\gamma_1 \leftrightarrow \gamma_2) \tag{4.22}$$

Cumulative transitivity (alias cut) is also valid:

$$((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \wedge ((\gamma_1 \wedge \gamma_2) \Rightarrow_c^{cl+} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow_c^{cl+} \gamma_3) \tag{4.23}$$

Conditional versions of antecedent strengthening, consequent weakening and transitivity are valid:

$$\neg[u](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2) \rightarrow ((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \Rightarrow_c^{cl+} \gamma_2)) \tag{4.24}$$

$$\neg[u](\gamma_1 \rightarrow \gamma_2 \vee \gamma_3) \rightarrow ((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \rightarrow (\gamma_1 \Rightarrow_c^{cl+} \gamma_2 \vee \gamma_3)) \tag{4.25}$$

$$\neg[u](\gamma_1 \rightarrow \gamma_3) \rightarrow (((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \wedge (\gamma_2 \Rightarrow_c^{cl+} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow_c^{cl+} \gamma_3)) \tag{4.26}$$

Proof. The validity of Formulae 4.18-4.21 is easily proven. The validity of Formula 4.22 is easily shown via application of Definition 4.8. We provide a deduction of Formula 4.23 and Formula 4.24. Let us start with Formula 4.23.

1.	(P)	$(\gamma_1 \rightarrow \gamma_2) \wedge (\gamma_1 \wedge \gamma_2 \rightarrow \gamma_3) \rightarrow (\gamma_1 \rightarrow \gamma_3)$
2.	(N), (K), (MP), 1	$([c](\gamma_1 \rightarrow \gamma_2) \wedge [c](\gamma_1 \wedge \gamma_2 \rightarrow \gamma_3)) \rightarrow [c](\gamma_1 \rightarrow \gamma_3)$
3.	(P)	$\gamma_1 \wedge \gamma_2 \wedge \neg\gamma_3 \rightarrow \gamma_1 \wedge \neg\gamma_3$
4.	(N), (K), (MP), (P), 3	$\neg[u]\neg(\gamma_1 \wedge \gamma_2 \wedge \neg\gamma_3) \rightarrow \neg[u]\neg(\gamma_1 \wedge \neg\gamma_3)$
5.	(P), (MP), 4	$\neg[u](\gamma_1 \rightarrow \gamma_2) \wedge \neg[u](\gamma_1 \wedge \gamma_2 \rightarrow \gamma_3) \rightarrow \neg[u](\gamma_1 \rightarrow \gamma_3)$
6.	(P), (MP), (Def. 4.8), 2, 5	$((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \wedge ((\gamma_1 \wedge \gamma_2) \Rightarrow_c^{cl+} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow_c^{cl+} \gamma_3)$

The deduction of Formula 4.24 follows.

1.	(P)	$(\gamma_1 \rightarrow \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)$
2.	(N), (K), (MP), 1	$[c](\gamma_1 \rightarrow \gamma_2) \rightarrow [c](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)$
3.	(P)	$\neg[u](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)$ $\rightarrow (\neg[u](\gamma_1 \rightarrow \gamma_2) \rightarrow \neg[u](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2))$
4.	(P), (MP), (Def. 4.8), 2, 3	$\neg[u](\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2)$ $\rightarrow ((\gamma_1 \Rightarrow_c^{cl+} \gamma_2) \rightarrow (\gamma_1 \wedge \gamma_3 \Rightarrow_c^{cl+} \gamma_2))$

Deductions of Formulae 4.25 and 4.26 are just as straightforward. □

Propositions 4.6 and 4.7, though very simple, are of key importance for putting our characterization of counts-as as proper contextual classification in perspective with other proposals. Such a comparison is elaborated in detail in Section 4.8.2.

Formulae 4.24-4.26 are also of interest since they show that some quite standard properties of contextual classifications are inherited by proper contextual classification, but only in a conditionalized form, the condition being an assertion of invalidity ($\neg[u]$). Proper classificatory counts-as statements are still monotonic, provided that the strengthened version of the antecedent does not universally imply the consequent. Similarly, they are still transitive provided that the implication between γ_1 and γ_3 is not a validity of the model. In other words, the monotonicity and the transitivity of proper classificatory counts-as are sensitive to the global context. The choice of models with suitable global contexts would make them valid (in those models).

It is worth emphasizing the importance of these results from a conceptual analysis perspective and their clarifying power. Suppose somebody is providing an alleged example of transitivity for proper classificatory counts-as by stating that: if X counts as Y (in a proper classificatory sense) and Y counts as Z (in a proper classificatory sense) then we can infer that X counts as Z (in a proper classificatory sense). Our analysis shows that such inference is legitimate only thanks to a hidden premise, that is, only if it is also assumed that "X implies Z" is not a global truth of the model. Similar considerations hold also for the conditionalized version of antecedent strengthening. This property will be further discussed in Section 4.8.5.

4.5 Counts-as as Constitution

In this section a modal logic is developed which is based on Definition 4.6. Again, the possible world semantics developed in order to account for the classificatory view of counts-as lies at the ground of the proposed framework.

4.5.1 Expanding $\mathcal{L}_n^u$

Language $\mathcal{L}_n^u$, which has been used in the previous section to deal with proper contextual classification, needs now further expansion to enable the necessary expressivity. The language is expanded along two lines.

First, the set of context indexes C contains now a set K of m atomic indexes c among which the universal context index u, and the set of the negations $-c$ of the atomic contexts, i.e., of the elements of K: $C = K \cup \{-c \mid c \in K\}$. The cardinality n of C is therefore equal to $2m$.

Second, the language needs also to contain an at most countable set $\mathbb{N}$ of nominals s disjoint from the set $\mathbb{P}$ of propositional atoms. Nominals are names for states in the model or, in other words, formulae that can be satisfied by only one state in the model. They can be freely combined with propositions to form well-formed formulae. The BNF is therefore extended as follows:

$$\phi ::= \top \mid p \mid s \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid \phi_1 \rightarrow \phi_2 \mid [i]\phi \mid \langle i \rangle \phi.$$

Metavariables for nominals are written as $v_1, v_2, \ldots$. Modal languages containing nominals have recently been object of thorough study and are known as hybrid languages ([Blackburn et al., 2001]). The language obtained is called $\mathcal{L}_n^{u,-}$.

Nominals are chosen here in order to provide a sound and complete axiomatization of the logic based on the semantics presupposed by Definition 4.6. To be more precise, they are necessary in order to axiomatize the notion of complement of a context[5]. This will become evident by exposing the axiomatics (Section 4.5.3) and especially, from a technical point of view, in proving its completeness (Appendix A).

4.5.2 Semantics

A semantics to language $\mathcal{L}_n^{u,-}$ is given via a special class of Cxt frames, namely the class of Cxt frames $\mathcal{F} = \langle W, \{W_i\}_{i \in C} \rangle$ such that there always exists a $u \in C$ s.t. $W_u = W$; and such that for any atomic index $c \in K$ there exists $-c \in C$ such that: $W_{-c} = W_u \backslash W_c$ That is, the frames in this class, which we call $\text{Cxt}^{\top,\backslash}$, always contain the global context among their contexts and the complement of every atomic context.

The semantics for $\mathcal{L}_n^{u,-}$ is thus obtained interpreting the formulae on models built on $\text{Cxt}^{\top,\backslash}$ frames. However, because of the introduction of nominals, the evaluation function $\mathcal{I}$ should be redefined as a function $\mathcal{I} : \mathbb{P} \cup \mathbb{N} \longrightarrow \mathcal{P}(W)$ satisfying the following constraints:

- For all nominals $s \in \mathbb{N}$, $\mathcal{I}(s)$ is a singleton set, that is, nominals always denote one and only one state in the model.

[5]For this purpose nominals were first introduced by the so-called "Sofia school" of modal logic ([Passy and Tinchev, 1985, 1991; Gargov and Goranko, 1993]) in order to axiomatize the complement and the intersection of accessibility relations, especially in a dynamic logic setting. In fact, the axiomatics we present in Section 4.5.3 is strictly related with the systems studied in their works.

- For all states $w \in W$, there exists a nominal $s \in \mathbb{N}$ such that $\mathcal{I}(s) = w$, that is, each state has a name. In other words, the restriction of the interpretation function $\mathcal{I}$ on the set of nominals ($\mathbb{N}\rceil\mathcal{I}$) is a surjection on the set of all singletons of W.

Following [Gargov and Goranko, 1993], models with valuations satisfying the conditions above are called *surjective models*. The definition of the satisfaction relation for language $\mathcal{L}_n^{u,-}$ runs as follows.

Definition 4.9. (Satisfaction based on $\text{Cxt}^{\top,\backslash}$ frames)
Let $\mathcal{M}$ be a surjective model built on a $\text{Cxt}^{\top,\backslash}$ *frame.*

$$\begin{array}{rll}
\mathcal{M}, w \models s & \textit{iff} & \mathcal{I}(s) = \{w\} \\
\mathcal{M}, w \models [u]\phi & \textit{iff} & \forall\, w' \in W_u : \mathcal{M}, w' \models \phi \\
\mathcal{M}, w \models [c]\phi & \textit{iff} & \forall\, w' \in W_c : \mathcal{M}, w' \models \phi \\
\mathcal{M}, w \models [-c]\phi & \textit{iff} & \forall\, w' \in W\backslash W_c : \mathcal{M}, w' \models \phi.
\end{array}$$

where u is the universal context index and c ranges on the context indexes in C, and s is a nominal. The obvious boolean clauses and the clauses for the dual modal operators are omitted.

Surjective models on $\text{Cxt}^{\top,\backslash}$ frames will be referred to as $\text{Cxt}^{\top,\backslash}$ models. The first clause states the satisfaction relation for nominals: a nominal s is true in a state w in model $\mathcal{M}$ iff the evaluation function associates w to s. Nominals are therefore objective formulae which are true in at most one world. The second clause, which was already introduced in Definition 4.7, states that the $[u]$ operator is interpreted on the universal frame contained in each $\text{Cxt}^{\top,\backslash}$ frame. The third one is just the standard clause for contextual truth introduced in Definition 4.3. Finally, the last and new clause states that the $[-c]$ operators range over the complements of the sets W_c on which $[c]$ operators range instead.

Some observations are in order. First of all, let us comment upon the semantics of the $[-c]$-operators. In fact, the $[c]$ operator specifies a lower bound on what holds in context c ('something more may hold in c'), that is, a formula $[c]\phi$ means that ϕ *at least* holds in context c. The $[-c]$ operator, instead, specifies an upper bound on what holds in c ('nothing more holds in c'), and a $[-c]\neg\phi$ formula means therefore that ϕ *at most* holds in c, i.e., $\neg\phi$ *at least* holds in the complement of c. It becomes thus possible in $\text{Cxt}^{\top,\backslash}$ frames to express context definitions by means of modal $\mathcal{L}_n^{u,-}$ formulae interpreted on $\text{Cxt}^{\top,\backslash}$ models. A set of objective formulae Γ defines context c in a $\text{Cxt}^{\top,\backslash}$ model $\mathcal{M}$ iff:

$$\mathcal{M} \models [c]\Gamma \wedge [-c]\neg\Gamma \tag{4.27}$$

where $\neg\Gamma$ has to be intended in the obvious sense of the disjunction of the negations of all formulae in Γ. Formula 4.27 is an object language modal translation of the property stated in Formula 4.11.

Proposition 4.8. *(Equivalence of Formulae 4.11 and 4.27)*
Let $\mathcal{M}$ be a Cxt *model and $\mathcal{M}'$ be a model on a* $\text{Cxt}^{\top,\setminus}$ *frame such that: $\mathcal{M}'$ is based on a frame having the same domain of the frame on which $\mathcal{M}$ is based, and which contains all its contexts; propositional atoms get the same evaluation in $\mathcal{M}'$ and $\mathcal{M}$. It is the case that, given a set of objective formulae Γ and a context W_c:*

$$\mathcal{M}, w \models \Gamma \textit{ iff } w \in W_c$$

is equivalent to

$$\mathcal{M}' \models [c]\Gamma \wedge [-c]\neg\Gamma.$$

Proof. The proof is based on the semantics provided in Definition 4.9. By construction of $\mathcal{M}'$, the clause "if $w \in W_c$ then $\mathcal{M}, w \models \Gamma$" is equivalent to "if $w \in W_c$ then $\mathcal{M}', w \models \Gamma$", and therefore equivalent to $\mathcal{M}' \models [c]\Gamma$. Analogously, the clause "if $w \notin W_c$ then $\mathcal{M}, w \not\models \Gamma$" is equivalent to "if $w \in W \setminus W_c$ then $\mathcal{M}', w \models \neg\Gamma$", and therefore equivalent to $\mathcal{M}' \models [-c]\neg\Gamma$. □

In practice, we are making use, in a different setting but with similar purposes, of a well-known technique developed in the modal logic of knowledge, i.e., the interpretation of modal operators on 'inaccessible states' typical, for instance, of the "all that I know" epistemic logic ([Levesque, 1990]). In our case, the set of inaccessible states is nothing but the complement of a context.

4.5.3 Axiomatics

To axiomatize the above semantics an extension of logic $\mathbf{K45_n^{ij}}$ is needed which can characterize nominals as names for modal states and, consequently, context complementation. The extension, which we call logic $\mathbf{Cxt^{u,-}}$, results by adding to $\mathbf{Cxt^{u}}$ a group of two axioms (`Least` and `Most`) and one rule (`Name`) which axiomatize nominals, and a group of two axioms (`Covering` and `Packing`) which axiomatize context complementation. The axiomatics runs as follows:

(P)	all tautologies of propositional calculus
($\mathtt{K}^i$)	$[i](\phi_1 \to \phi_2) \to ([i]\phi_1 \to [i]\phi_2)$
(4^{ij})	$[i]\phi \to [j][i]\phi$
(5^{ij})	$\neg[i]\phi \to [j]\neg[i]\phi$
($\mathtt{T}^u$)	$[u]\phi \to \phi$
($\subseteq .ui$)	$[u]\phi \to [i]\phi$
(Least)	$\langle u \rangle\, \nu$
(Most)	$\langle u \rangle\, (\nu \wedge \phi) \to [u](\nu \to \phi)$
(Covering)	$[c]\phi \wedge [-c]\phi \to [u]\phi$
(Packing)	$\langle -c \rangle\, \nu \to \neg \langle c \rangle\, \nu$
(Dual)	$\langle i \rangle\, \phi \leftrightarrow \neg[i]\neg\phi$

$$\begin{array}{ll} (\texttt{Name}) & \text{IF } \vdash \nu \to \theta \text{ THEN } \vdash \theta, \quad \text{for } \nu \text{ not occurring in } \theta \\ (\texttt{MP}) & \text{IF } \vdash \phi_1 \text{ AND } \vdash \phi_1 \to \phi_2 \text{ THEN } \vdash \phi_2 \\ (\texttt{N}^i) & \text{IF } \vdash \phi \text{ THEN } \vdash [i]\phi \end{array}$$

where i, j are metavariables for the elements of K, c denotes elements of the set of atomic context indexes C, u is the universal context index, ν ranges over nominals, and θ in rule Name denotes a formula in which the nominal denoted by ν does not occur.

The proofs of soundness and completeness of the axiomatization w.r.t. $\text{Cxt}^{\top,\setminus}$ frames are provided in Appendix A.

The new axioms and rules deserve some comments. Let us start with the axiomatization of nominals. Axiom Least states just that every nominal denotes *at least* one state. Vice versa, axiom Most states that nominals denote *at most* one state. Intuitively it says that, if there is a state named ν where ϕ holds, then ϕ holds if ν is the case. Finally, rule Name is a rule with side conditions borrowed from standard hybrid logic ([Blackburn et al., 2001]). It forces all states to be nominated. It does that by saying that if it is provable that a formula θ holds at an arbitrary state ν —the state is arbitrary since the rule requires ν not to occur in θ— then θ itself is provable since there is no world that falsifies it. From a technical point of view, as observed in [Passy and Tinchev, 1985; Gargov and Goranko, 1993], this rule ensures that in any definable set of the model, i.e., set of states in which some modal formula is true, at least one state can be picked which is named by $\mathbb{N} \rceil \mathcal{I}$. This guarantees function $\mathbb{N} \rceil \mathcal{I}$ to be a surjection on the set of all definable singletons of W[6]. To sum up, axioms Least and Most with rule Name axiomatize the conditions holding on the interpretation function $\mathcal{I}$ as exposed in Section 4.5.2.

Let us now discuss the axioms that are more central to the modeling aim we are pursuing: axioms Covering and Packing. They characterize context complementation. Axiom Covering states that if some formula holds in both c and $-c$, than it holds globally. To put it otherwise, it states that the universal context is *covered* by the contexts denoted by c and, respectively, $-c$. Axiom Packing states then that the contexts denoted by c and $-c$ are strongly disjoint, in the sense that they do not contain the same states. They *pack* the universal context in two disjoint subcontexts. Axioms Covering and Packing are therefore just modal formulations of the two properties characterizing the bipartition of a given set. Notice that nominals are necessary in the formulation of the Packing axiom. It is easy to see that, without the possibility of naming individual states, it would be impossible to axiomatize disjointness[7].

[6]Rule Name plays a central role in the completeness proof for $\mathbf{Cxt}^{\mathbf{u},-}$ (see the proof of Lemma A.6 in Appendix A).

[7]However, it is not our claim that nominals are the only viable way to achieve this aim. Another possible and probably more elegant solution might consist in using the *difference operator*, by means of which it is possible to represent both the universal modality and nominals (see [Gargov and Goranko, 1993; Blackburn et al., 2001]).

4.5.4 A remark: $\mathbf{Cxt}^{\mathbf{u},-}$ as hybrid logic

Before putting the formalism at work it might be instructive to make one last technical remark. In logic $\mathbf{Cxt}^{\mathbf{u},-}$ a family $\{@_\nu\}_{\nu \in \mathbb{N}}$ of operators is definable, by means of which it is possible to express that a formula ϕ holds in the state named ν: $@_\nu \phi$. This operator is known in hybrid logics ([Blackburn et al., 2001]) as the *satisfaction operator*. Its semantics is given in terms of the following clause:

$$\mathcal{M}, w \models @_\nu \phi \text{ iff } \mathcal{M}, \mathcal{I}(\nu) \models \phi.$$

The property of "holding in a state" is thus a global property, that is, it is independent of the point of evaluation. The clause states more precisely that, whatever the state of evaluation is, it is the case that if s holds then ϕ also holds. In fact, the satisfaction operator can be defined in any logic enabling nominals and a universal modality (see for instance [Goranko and Passy, 1992; Areces et al., 2000]) as follows:

$$@_\nu \phi := [u](\nu \rightarrow \phi) \tag{4.28}$$

where $@_\nu$ is a nominal and ϕ a formula. Leaving technicalities aside, this means that logic $\mathbf{Cxt}^{\mathbf{u},-}$ has sufficient expressive means to represent statements of the type "in situation (or state) ν state-of-affairs ϕ holds". This expressive capability of logic $\mathbf{Cxt}^{\mathbf{u},-}$ will turn out useful to represent intuitive reasoning patterns involving constitutive counts-as statements (see Example 4.10).

4.5.5 Constitutive counts-as formalized

Using a multi-modal logic $\mathbf{Cxt}^{\mathbf{u},-}$ on a language $\mathcal{L}_n^{u,-}$, the constitutive reading of counts-as statements can now be formalized.

Definition 4.10. (Constitutive counts-as: $\Rightarrow^{co}_{c,\Gamma}$)
Given a set of formulae Γ such that $\gamma_1 \rightarrow \gamma_2 \in \Gamma$, the constitutive counts-as statement "γ_1 counts as γ_2 in the context c defined by Γ" is formalized in a multi-modal logic $\mathbf{Cxt}^{\mathbf{u},\backslash}$ on language $\mathcal{L}_n^{u,-}$ as follows:

$$\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2 \quad := \quad [c]\Gamma \wedge [-c]\neg\Gamma \wedge \neg[u](\gamma_1 \rightarrow \gamma_2)$$

with γ_1 and γ_2 objective formulae.

The definition implements in modal logic the intuition summarized in Observation 4.2, and formalized in Definition 4.6: constitutive counts-as statements correspond to those non trivial classifications which are stated by the definition Γ of the context c. In fact the following can be proven.

Proposition 4.9. *(Equivalence of Definitions 4.10 and 4.6)*
Let $\mathcal{M}$ be a $\text{Cxt}^{\top,\backslash}$ frame and Γ a set of objective formulae. It is the case that: $\gamma_1 \rightarrow \gamma_2 \in \mathbb{CO}(\Gamma, W_c)$ iff $\gamma_1 \rightarrow \gamma_2 \in \{\gamma_1 \rightarrow \gamma_2 \in \Gamma \mid \mathcal{M} \models \gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2$ }. To put it otherwise:

$$\mathbb{CO}(\Gamma, W_c) = \{\gamma_1 \rightarrow \gamma_2 \in \Gamma \mid \mathcal{M} \models \gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2\}$$

Proof. The proof follows from Proposition 4.8 and Definition 4.10. □

A detailed comment of Definition 4.10 is in order. Its most important consequence is that it is possible to talk about constitutive counts-as only once a set Γ is given. As already stressed in Section 4.3.4, there is no formula that is constitutive in isolation from a set of rules.

Secondly, notice that a constitutive counts-as is false if either Γ does not define the context denoted by c, or if it expresses a classification which is valid in the model. This is the distinctive feature of constitutive counts-as with respect to its two classificatory relatives. While the classificatory versions of counts-as express what at least holds in a context (contextual classification) and, respectively, what at least hold in a context which is not globally true (proper contextual classification), the constitutive version expresses also what at most holds in a context, thereby making explicit what the context actually is in terms of a set of formulae of the language. We can have a constitutive counts-as statement only if it is known what the definition is of the context at issue. In the classificatory versions of counts-as this knowledge is absent since it is only partially known what the context explicitly is. Classificatory and proper classificatory counts-as statements presuppose the existence of a context of which only some information is available. This issue is discussed in some more detail in Section 4.7.1.

From a technical point of view, this linguistic dependence corresponds to the fact that $\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2$ formulae are defined only for pairs of formulae (γ_1, γ_2) s.t. $\gamma_1 \rightarrow \gamma_2 \in \Gamma$. To put it another way, symbols $\Rightarrow^{co}_{c,\Gamma}$ are not genuine connectives. As a consequence, it is not possible to study $\Rightarrow^{co}_{c,\Gamma}$ conditionals from a structural perspective like it has been done for the other forms of counts-as in Propositions 4.5, 4.6 and 4.7.

How awkward this might sound it is perfectly aligned with the intuitions on the notion of constitution which backed Definition 4.10: constitutive counts-as are those classifications which are explicitly stated in the specification of the normative system. In a sense, constitutive statements are just given, and that is it. This does not mean, however, that constitutive statements cannot be used to perform reasoning. The following example depicts the most typical form of reasoning involving constitutive counts-as statements.

Proposition 4.10. *($\Rightarrow^{co}_{c,\Gamma}$ and $@_v$)*
The following formula is valid in $\textsc{Cxt}^{\top,\setminus}$ *frames for any* Γ *containing* $\gamma_1 \rightarrow \gamma_2$*:*

$$\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2 \rightarrow ((@_v\Gamma \wedge @_v\gamma_1) \rightarrow @_v\gamma_2) \tag{4.29}$$

Proof. Follows from Definition 4.6, Formula 4.28 and propositional logic. □

This property shows how constitutive rules work in providing grounds for inferring the occurrence of new states-of-affairs: it is a rule of the normative system of Utrecht University that if the promotor pronounces the PhD student to be a doctor then this counts as the PhD student to be a doctor ($\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2$); the current situation v falls under the rules of Utrecht University ($@_v\Gamma$) and in the current situation the promotor pronounces a PhD student to be a doctor ($@_v\gamma_1$), hence in the current situation the PhD student is a doctor ($@_v\gamma_2$).

It is remarkable that Formula 4.29 perfectly depicts the notion of "conventional generation" as described in [Goldman, 1976]:

> "Act-token A of agent X conventionally generates act-token B [. . .] only if the performance of A [. . .], together with a rule R saying that A [. . .] counts as B, guarantees the performance of B" ([Goldman, 1976], p. 25).

Notice also that, besides formula $\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2$, what plays an essential role here is formula $@_v\Gamma$ (i.e., $[u](v \to \Gamma)$), which states that situation v is one of the situations in context c. Without the notion of context definition and the availability of nominals, this could not be expressed.

Complex reasoning patterns involving constitutive counts-as statements arise also in relation with the other two notions of counts-as. The following section investigates the logical relationships between the three different senses of counts-as.

4.6 Relating the many faces of counts-as

This section is devoted to pursuing the last goal mentioned in the quote from [Tarski, 1944] mentioned at the beginning of the chapter: "and then we may proceed to a quiet and systematic study of all concepts involved, which will exhibit their main properties and mutual relations."

The logical relations between $\Rightarrow^{co}_{c,\Gamma}$, $\Rightarrow^{cl+}_{c}$ and $\Rightarrow^{cl}_{c}$ can be studied in logic $\mathbf{Cxt^{u,\setminus}}$ which extends both $\mathbf{K45^{ij}_n}$, i.e., the logic in which $\Rightarrow^{cl}_{c}$ has been defined, and $\mathbf{Cxt^{u}}$, i.e., the logic in which $\Rightarrow^{cl+}_{c}$ has been defined.

Proposition 4.11. ($\Rightarrow^{cl}_{c}$ vs $\Rightarrow^{cl+}_{c}$ vs $\Rightarrow^{co}_{c,\Gamma}$)
In logic $\mathbf{Cxt^{u,\setminus}}$ *the following formulae are valid:*

$$(\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \to (\gamma_1 \Rightarrow^{cl}_{c} \gamma_2) \tag{4.30}$$

$$(\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \to (\gamma_1 \wedge \gamma_3 \Rightarrow^{cl}_{c} \gamma_2) \tag{4.31}$$

$$((\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \wedge (\gamma_2 \Rightarrow^{cl+}_{c} \gamma_3)) \to (\gamma_1 \Rightarrow^{cl}_{c} \gamma_3) \tag{4.32}$$

$$(\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2) \to (\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \tag{4.33}$$

provided that $\gamma_1 \to \gamma_2 \in \Gamma$.

Proof. The validity of Formula 4.30 follows directly from Definitions 4.4 and 4.8: $(\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \leftrightarrow (\gamma_1 \Rightarrow^{cl}_{c} \gamma_2 \wedge \neg[u](\gamma_1 \to \gamma_2))$.
The validity of Formula 4.31 follows from the validity of Formula 4.30, the validity of Formula 4.8 for $\Rightarrow^{cl}_{c}$ (Proposition 4.5) and MP.
Finally, the validity of Formula 4.32 follows also from the validity of Formula 4.30, the validity of Forula 4.6 of $\Rightarrow^{cl}_{c}$ (Proposition 4.5) and MP. Formula 4.33 follows straightforwardly from Definition 4.10. □

Let us have a look at the intuitive meaning of the formulae just proven. Formula 4.30 states something very simple: proper contextual classification implies contextual classification. This corresponds, in the model-theoretic notation used in Section 4.3, to the following inclusion relation: $\mathbb{CL}^{+}(W_c) \subseteq \mathbb{CL}(W_c)$.

Formulae 4.31 and 4.32 are particularly interesting. If we forget that the two operators $\Rightarrow^{cl+}_{c}$ and $\Rightarrow^{cl}_{c}$ denote two different notions and we read both expressions $\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2$ and $\gamma_1 \Rightarrow^{cl}_{c} \gamma_2$ just as "γ_1 counts as γ_2", these formulae would sound as statements of the property of antecedent strengthening and of the transitivity of "counts-as". However, our formal analysis based on the acknowledgment of the polisemy of counts-as has shown that transitivity and antecedent strengthening hold for $\Rightarrow^{cl}_{c}$ but not for $\Rightarrow^{cl+}_{c}$. On the other hand, and this is what Proposition 4.11 shows, their logical interactions display patterns clearly reminiscent of those properties. In a sense, it has been shown that questions such as "is transitivity an intuitive property for a characterization of counts-as?" are flawed by the possibility of confusing under the label counts-as different notions which enjoy different logical properties.

More specifically, Formula 4.31 expresses that given a counts-as statement interpreted as a proper classification, a contextual classification can be inferred having as antecedent a strengthened version of the antecedent of the first statement, and this although proper contextual classification does not enjoy antecedent strengthening. In other words, although $\Rightarrow^{cl+}_{c}$ does not enjoy antecedent strengthening, it is nonetheless grounds for performing monotonic reasoning via $\Rightarrow^{cl}_{c}$. Analogous considerations apply to Formula 4.32. Proper contextual classification does not enjoy transitivity but reasoning via transitivity remains valid shifting from $\Rightarrow^{cl+}_{c}$ to $\Rightarrow^{cl}_{c}$.

Finally, Formula 4.33 translates the following intuitive fact: the promulgation of a constitutive rule *guarantees*, to say it with [Jones and Sergot, 1996], the possibility of applying specific classificatory rules. If it is a rule of Γ that self-propelled conveyances count as vehicles (constitutive sense) then self-propelled conveyances count as vehicles (proper classificatory sense) in the context c defined by Γ.

The following two propositions display further interesting consequences of Definition 4.8 concerning the relation between constitution and classification.

Proposition 4.12. (Impossibility of $\Rightarrow^{cl+}_{u}$ and $\Rightarrow^{co}_{u,\Gamma}$)
Proper classificatory counts-as statements and constitutive counts-as statements are impossible with respect to the universal context u. In symbols, the following formulae are valid in $\text{Cxt}^{\top,\setminus}$ *frames:*

$$(\gamma_1 \Rightarrow^{cl+}_{u} \gamma_2) \to \bot \tag{4.34}$$

$$(\gamma_1 \Rightarrow^{co}_{u,\Gamma} \gamma_2) \to \bot \tag{4.35}$$

for $\gamma_1 \to \gamma_2 \in \Gamma$.

Proof. The proposition is easily proven considering that Definition 4.8 yields that Formula 4.34 is equivalent to: $[u](\gamma_1 \to \gamma_2) \wedge \neg[u](\gamma_1 \to \gamma_2)$ which is, for Definition 4.10, implied also by Formula 4.35. $\square$

Intuitively, Formula 4.34 expresses that what is considered to hold in general is not the product of constitution, it is just, so to say, what is taken to be necessarily the case. Formula 4.35 states something slightly different, although much related: the global context u is not constituted by any set of rules Γ. To put it another way, what Formulae 4.34 and 4.35 say is that the global context u is what sets the boundaries of the possible constitutions. Notice that contextual classificatory statements are instead perfectly sound also with respect to the universal context. In fact, formula $\gamma_1 \Rightarrow^{cl}_{u} \gamma_2$ is satisfiable in $\mathrm{C}\textsc{xt}^{\top,\backslash}$ models.

Proposition 4.13. (Impossibility of $\Rightarrow^{cl+}_{c}$ and $\Rightarrow^{co}_{c,\Gamma}$)
Global truths cannot be the content of proper classificatory counts-as or constitutive counts-as statements. In symbols, the following formulae are valid in $\mathrm{C}\textsc{xt}^{\top,\backslash}$ *frames:*

$$[u](\gamma_1 \rightarrow \gamma_2) \rightarrow ((\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \rightarrow \bot) \tag{4.36}$$

$$[u](\gamma_1 \rightarrow \gamma_2) \rightarrow ((\gamma_1 \Rightarrow^{co}_{c,\Gamma} \gamma_2) \rightarrow \bot) \tag{4.37}$$

for $\gamma_1 \rightarrow \gamma_2 \in \Gamma$.

Proof. The proposition follow directly from Definitions 4.8 and 4.10. From Definition 4.8 it follows that Formula 4.36 implies: $[u](\gamma_1 \rightarrow \gamma_2) \wedge \neg[u](\gamma_1 \rightarrow \gamma_2)$. The same follows from Definition 4.10, which proves Formula 4.37. □

Formulae 4.36 and 4.37 express that what is taken to be globally the case cannot be a proper contextual classification and cannot be used to constitute a context. The reason for this is that global truths hold in all contexts, and therefore, they cannot be specific of any one. To put it in yet another way, if something is considered to be a proper contextual counts-as or a constitutive one, then it is also presupposed that what stated by the counts-as can possibly not be the case. For instance, if we take "apples are fruits" to be a global truth of our reality, then "apples count as fruits" cannot be a constitutive rule since it adds nothing to what is already the case. On the contrary, if we take "apples count as fruits" to be one of the constitutive rules of a system Γ then we are assuming that in some cases apples are not classified as fruits.

Let us now take into consideration properties displaying more complex reasoning patterns.

Proposition 4.14. (From $\Rightarrow^{co}_{c,\Gamma}$ to $\Rightarrow^{cl}_{c}$ and $\Rightarrow^{cl+}_{c}$ via $\Rightarrow^{cl}_{u}$)
The following formulae are valid in $\mathrm{C}\textsc{xt}^{\top,\backslash}$ *frames:*

$$(\gamma_2 \Rightarrow^{co}_{c,\Gamma} \gamma_3) \rightarrow ((\gamma_1 \Rightarrow^{cl}_{u} \gamma_2) \rightarrow (\gamma_1 \Rightarrow^{cl}_{c} \gamma_3)) \tag{4.38}$$

$$(\gamma_2 \Rightarrow^{co}_{c,\Gamma} \gamma_3) \rightarrow (((\gamma_1 \Rightarrow^{cl}_{u} \gamma_2) \wedge \neg[u](\gamma_1 \rightarrow \gamma_3)) \rightarrow (\gamma_1 \Rightarrow^{cl+}_{c} \gamma_3)) \tag{4.39}$$

provided that $\gamma_1 \rightarrow \gamma_2 \in \Gamma$.

Proof. The proof of Formula 4.38 is straightforward from Definition 4.4, Definition 4.10, Proposition 4.7 and the transitivity of classificatory counts-as (Proposition 4.5). Formula 4.39 is proven by just adding the application of Definition 4.8 to the the proof of Formula 4.38. □

These properties represent typical forms of reasoning patterns involving constitutive rules. Formula 4.38: if it is a rule of Γ that $\gamma_2 \rightarrow \gamma_3$ ("self-propelled conveyances count as vehicles") and it is always the case that $\gamma_1 \rightarrow \gamma_2$ ("cars count as self-propelled conveyances"), then $\gamma_1 \rightarrow \gamma_3$ ("cars count as vehicles") holds in the context c defined by normative system Γ.

Formula 4.39: if it is a rule of Γ that $\gamma_2 \rightarrow \gamma_3$ ("conveyances transporting people or goods count as vehicles") and it is always the case that $\gamma_1 \rightarrow \gamma_2$ ("bikes count as conveyances transporting people or goods") but it is not always the case that $\gamma_1 \rightarrow \gamma_3$ ("bikes count as vehicles"), then $\gamma_1 \rightarrow \gamma_3$ ("bikes count as vehicles") holds as a constituted classification in the context c defined by normative system Γ. Notice that while "cars count as self-propelled conveyances" is a classificatory counts-as, since it might still be the case that cars are globally classified as vehicles, "bikes count as vehicles" is instead a proper classificatory counts-as since it is explicitly stated that such classification is not a validity. Formula 4.39 represents nothing but the form of the reasoning pattern that has been used as example in Section 4.3.2 to introduce the notion of constitution.

The very remarkable aspect about these properties is that they neatly show how the three senses of counts-as all play a role in the kind of reasoning we perform with constitutive rules. In particular, they show that the constitutive sense, though not enjoying any structural property, grounds in fact all the rich reasoning patterns proper of classificatory reasoning.

4.7 Discussion: Putting Counts-as in Perspective

Aim of this section is to place the formal analysis presented in a broader context, by showing how it relates to a number of issues that have been raised in the applied logic literature. This will clearly suggest that the logical nature of counts-as is, in effect, nothing unique or exotic, and that it is, on the contrary, strictly related to logical issues of a general kind.

More specifically: Section 4.7.1 relates classificatory and proper classificatory counts-as to the notion of enthymeme; Section 4.7.2 addresses the relationships that the logics presented here bear with formalisms developed in epistemic logic; and Sections 4.7.3 and 4.7.4 discuss counts-as in the light of deontic logic.

4.7.1 Counts-as as enthymeme

We introduced the notion of contextual classification for capturing a precise sense in which counts-as statements can be interpreted. Looking just at the formal machinery used, this section aims at showing the close relation that the notion of contextual classification enjoys with other logical notions. Contextual classification has been defined as an implicative statement holding with respect to a context, i.e., a set of valuations or, in the modal logic terminology, possible worlds or states (Definition 4.4). We have seen that the role of a context is to limit the set of states with respect to which the implicative statement is evaluated in order for it to represent a classification holding "locally" (see Section 4.1). As such, we can thus consider contexts to play the role of hidden collections of premises.

Inferences with hidden premises have a long history in logic. The ancient Greeks used to call them *enthymemes* from *en*, in, and *thymos*, mind, so as to mean some knowledge that is left implicit and kept *in the mind*. Counts-as statements state enthymemes in a very precise sense, the hidden premises being the constitutive rules of the normative system to which they pertain. A statement "X counts as Y in context C", interpreted as contextual classification, can therefore be rephrased as "it follows (classically) from the rules of the normative system specifying the context (i.e., set of models) C that X implies Y". Once the statement is considered abstracting from the set of rules of the relevant normative system —and this is the case in the Searlean analysis ([Searle, 1969, 1995])— what remains is just a general notion of context, whose specification is left open, and whose function is just to localize the truth of the statement "X implies Y".

Enthymemes have been studied as special consequence operations in [Makinson, 2005], where they are shown to provide a bridge between classical logic and non-monotonic logics. In that work the notion of enthymeme is captured by a specific logical consequence operation called *pivotal-valuation consequence*. The definition of this consequence operation runs as follows.

Definition 4.11. *(Pivotal-valuation consequence)*
Let W be the set of valuations of a propositional language $\mathcal{L}$. A formula γ_2 follows from γ_1 modulo the set of valuations $W_c \subseteq W$ iff there is no valuation v in W_c s.t. $v(\gamma_1) = 1$ and $v(\gamma_2) = 0$.

It is easy to see that Definition 4.4 and Definition 4.11 are very similar. In fact, we can restate Definition 4.11 in terms of the validity of $\mathbf{Cxt^u}$ formulae on logically universal $\text{C}\textsc{xt}^\top$ models, i.e., those models containing all possible valuations of a propositional language $\mathcal{L}$ and all possible contexts on that domain. Let $\mathcal{M}$ be such a model. A formula γ_2 follows from γ_1 modulo the set of valuations $W_c \subseteq W$ iff:

$$\mathcal{M} \models [c](\gamma_1 \rightarrow \gamma_2), \text{ i.e., } \mathcal{M} \models \gamma_1 \Rightarrow^{cl}_c \gamma_2$$

Contextual classification and pivotal-valuation consequence are, formally speaking, strict relatives, and modal logics such as $\mathbf{Cxt^u}$ are logics in which the notion of pivotal-valuation can be studied at an object-language level.

It is in fact no surprise to see (Proposition 4.5) that all the properties, which following [Makinson, 2005] characterize pivotal-valuation consequences, are enjoyed by our operator $\Rightarrow^{cl}_c$ for contextual classification: reflexivity (Formula 4.5), antecedent strengthening or monotonicity (Formula 4.8), transitivity (Formula 4.6), the property of disjunction of the premises (Formula 4.4), and supraclassicality. Supraclassicality means that via pivotal-valuation consequence more can be inferred than what can be classically inferred, which is what guaranteed by axiom schema ($\subseteq .ui$) in logic $\mathbf{Cxt^u}$ ($[u]\phi \rightarrow [i]\phi$).

Counts-as statements, when interpreted as contextual classifications, can be sensibly viewed as enthymemes: "it follows (classically) from the rules of the normative system specifying context C that X implies Y". The very same analogy can be drawn for those counts-as statements denoting proper classifications. In this case the reading would run like this: "it follows (classically) from the rules of the normative system specifying context C that X implies Y, but it does not hold in general (i.e. in all states) that X implies Y". Thus, proper contextual classification represents an enthymeme where what is inferred from the implicit premises is something that cannot be logically inferred from the empty set of premises.

A subclass of pivotal-valuation consequences is the class of pivotal-assumption consequences, which correspond to those pivotal-valuation consequences where the set W_c of valuations is definable by a set Γ of formulae.

Definition 4.12. *(Pivotal-assumption consequence)*
Let Γ be a set of formulae on a propositional language $\mathcal{L}$ and W be the set of valuations of a propositional language $\mathcal{L}$. A formula γ_2 follows from γ_1 modulo the set of assumptions Γ iff there is no valuation v s.t. $v(\Gamma \cup \{\gamma_1\}) = 1$ and $v(\gamma_2) = 0$.

Now, context definability is one of the ingredients lying at the ground of the formal characterization of constitutive counts-as (see Section 4.3.4). In fact, Definition 4.12 can also be restated in modal logic. Let $\mathcal{M}$ be a logically universal model on a $\textsc{Cxt}^{\top,\setminus}$ frame, containing all propositional valuations of language $\mathcal{L}$ and all possible contexts: a formula γ_2 follows from γ_1 modulo the set of assumptions Γ iff there is no $w \in \{w \in W \mid \mathcal{M}, w \models \Gamma\}$ s.t. $\mathcal{M}, w \models \gamma_1$ and $\mathcal{M}, w \not\models \gamma_2$. If we set $W_c = \{w \in W \mid \mathcal{M}, w \models \Gamma\}$, then we can characterize pivotal-assumption consequences via validity in $\mathcal{M}$ as follows: a formula γ_2 follows from γ_1 modulo the set of assumptions Γ iff:

$$\mathcal{M} \models [c]\Gamma \wedge [-c]\neg\Gamma \wedge [c](\gamma_1 \rightarrow \gamma_2)$$

that is to say, iff formula γ_2 follows from γ_1 modulo the set of valuations $W_c \subseteq W$ (pivotal-valuation consequence) and W_c is defined by Γ. In other words, pivotal assumption consequences correspond, in our framework, to those contextual classifications ($\Rightarrow_c^{cl}$) which follow from the constitution of a context by a set of rules Γ, that is to say, which hold in some context W_c defined by some set of formulae Γ.

4.7.2 Counts-as and epistemic logic

Readers acquainted with epistemic logic have probably noticed striking similarities of some of the logics presented with logics usually used to represent epistemic states of different agents. These similarities are here explicitly pointed out and discussed.

Logic $\mathbf{K45_n^{ij}}$ and epistemic logic

In fact, logic $\mathbf{K45_n^{ij}}$ presented in Section 4.2.5 is close to logic $\mathbf{KD45_n}$, i.e., the multi-modal version of logic **KD45**, which can be used to represent the beliefs of a number of agents ([Hintikka, 1962]). In a doxastic perspective, the distinguishing features of logic $\mathbf{K45_n^{ij}}$, with respect to $\mathbf{KD45_n}$, are two. First, $\mathbf{K45_n^{ij}}$ allows for inconsistent doxastic states. Second, interaction schemata 4^{ij} and 5^{ij} are valid. From a doxastic point of view, this means that agents' beliefs are always the same no matter by which agent they are considered. If agent i believes ϕ (respectively, does not believe ϕ), then agent j believes that agent i believes ϕ (respectively, that agent i does not believe ϕ). In other words, agents are transparent to each other. This is obviously a quite unrealistic assumption since, for instance, agent j can believe that agent i believes ϕ, while indeed it does not[8].

However, a convincing epistemic interpretation of logic $\mathbf{K45_n^{ij}}$ can be found once the notion of context is not interpreted as the epistemic state of one agent in a group of agents, but rather as one of the epistemic states one single agent *can* assume. What we called here context can be doxastically interpreted as 'set of hypotheses', or 'set of presumable beliefs with respect to a given situation', or 'context of a given theory' etc.:

> "One may think of a researcher in physics who may consider an electron alternately as a particle or as a wave, depending on whether s/he thinks classically or quantum-physically" ([Meyer and van der Hoek, 1995], p. 79).

More precisely, formulae $[i]\phi$ would intuitively mean: "the agent believes ϕ with respect to context (or situation, or opinion, or body of hypothesis) i". This perspective underpins various work carried out in the area of epistemic logic, especially when the core issue is the innocuous representation of inconsistencies[9].

This is the case, for instance, of the model theoretic analysis of *local reasoning* proposed in [Fagin et al., 1988] which is in many respects very similar to the semantics we presented in Section 4.2.2. In fact, the structures grounding the semantics in [Fagin et al., 1988] are called 'cluster models' and are a slightly more complex version of Cxt models. Clusters are in fact contexts, but the set of clusters is not invariant in the model, but varies from state to state. In fact, cluster models can be seen as models relativizing the set of available contexts to each state in the model. In a way, every state in a cluster model can be thought of a Cxt model. Modal operators are then interpreted quantifying on the set of available clusters, and not on the set of states within a cluster: $[i]$ modalities express that something holds in all clusters (contexts) within set i; $\langle i \rangle$ modalities express that something holds in at least one cluster (context) within set i.

Logic $\mathbf{Cxt^u}$ and epistemic logic

In [Meyer and van der Hoek, 1995], a variant of logic $\mathbf{Cxt^u}$, called **EDL** (*epistemic default logic*)[10], has been developed in order to provide a monotonic epistemic account of defeasible reasoning. As observed in Section 2.6.3, that work proposes a

[8] Cf. Footnote 3.

[9] As it has been stressed in Section 4.1, the possibility of representing conflicting classifications holding in different normative systems is also one of the motivating features of our work.

[10] See, for details, Chapter 4 of [Meyer and van der Hoek, 1995].

notion of context as "monotonic means for analyzing non-monotonic reasoning" in Section 2.6.3.

Logic **EDL** contains a standard **S5** knowledge operator and a family of $[i]$ operators referring to different possible sets of beliefs ('working beliefs') that one agent can assume given what it knows. Formula $[i]\phi$ means that ϕ holds given the set i of working beliefs, or also that the agent presumes ϕ given i. In other words, **EDL** represents at the same time what an agent knows to be true ($\Box$ formulae) and what it considers to be possible w.r.t what it knows ($[i]$ formulae). The models of **EDL** are in fact identical with $\textsc{Cxt}^{\top}$ models, except for the fact that the $\Box$ modality, i.e., the knowledge operator, is interpreted as an equivalence relation, while the global modality $[u]$ of $\mathbf{Cxt^u}$ is interpreted on universal relations. However, it is well-known that these two different interpretations are modally indistinguishable, and they deliver in fact the same logic[11].

Logic $\mathbf{Cxt^{u,-}}$ and epistemic logic

Finally, as it has been mentioned in Section 4.5.2, logic $\mathbf{Cxt^{u,-}}$ incorporates the logical machinery that was first introduced in [Levesque, 1990] in order to express "all that it is known" by an agent. What is at least believed by an agent is modeled by a set of states (i.e., the context making true what is at least believed), and what it is at most believed is modeled by the complement of that set (i.e., the context making false what is at most believed). The logic proposed in [Levesque, 1990] is therefore essentially mono-modal, so to say, and it extends logic **KD45**. Logic $\mathbf{Cxt^{u,-}}$ is a multi-modal version thereof, except for the absence of axiom `D` and for the presence of the interaction axioms 4^{ij} and 5^{ij}. A system which is similar to the one proposed in [Levesque, 1990] is system **S5O** studied in [Meyer and van der Hoek, 1995]. Logic **S5O** extends the "all that it is known" logic with a strongly universal context, that is, the context consisting of all logically possible valuations, i.e., the context of logical truths. The presence of a universal context makes logic **S5O** very similar to $\mathbf{Cxt^{u,-}}$, although the universal context of the latter is not necessarily the logically universal one. This difference is reflected by the axiomatizations of those logics: the axiomatization of logic **S5O** makes use of the notion of uniquely satisfiable formulae, i.e., formulae for which there exists one unique propositional valuation among the set of all logically possible valuations, while in the axiomatization of $\mathbf{Cxt^{u,-}}$ we make use of the notion of nominal. Like uniquely satisfiable formulae, mominals only denote one state each, but unlike them, they do not necessarily denote all possible propositional valuations of a language.

4.7.3 A counts-as reduction of deontic logic

In this section, we will concisely focus on those approaches to deontic logic[12] which make use of the well-known reduction strategy first presented by Anderson in [Anderson, 1957, 1958], and relate them to the notion of counts-as.

[11] See Appendix A, Lemma A.5.

[12] See Section 2.5 for some introductory words.

The reduction strategy is based on the intuition according to which ϕ is ideally the case means that $\neg\phi$ "necessarily" implies a violation (of the relevant set of norms or deontic constraints), in symbols: $\Box(\neg\phi \rightarrow V)$, where V is a specific atom for which it is valid that $\Diamond\neg V$, i.e., that the violation is not "necessary". The nature of the reduction lies in how this reference to a "necessity" is formally modeled. In the original proposal of Anderson the system chosen for the reduction was **KT**, i.e., system **K** plus the axiom schema T for reflexivity ($\Box\phi \rightarrow \phi$). Here we consider alternative reductions based instead on system **S5** such as the ones studied in [d'Altan et al., 1993; Krabbendam and Meyer, 1999; Lomuscio and Sergot, 2003]. Interpreting the $\Box$ operator occurring in the reduction expression as an **S5** necessity yields expressions formally identical to the ones discussed at the beginning of our analysis in Section 4.1.2 where the analogy was stressed between subsumption statements $X \sqsubseteq Y$ and **S5** strict implications. In this view, formulae $\Box(\neg\phi \rightarrow V)$ could therefore be rephrased as: *the negation of ϕ is unconditionally classified as a violation*. According to this approach, deontic notions are reduced to universal classification statements.

Taking this reduction strategy as a starting point, our analysis of counts-as statements as contextual classifications is then readily applicable and delivers a straightforward and intuitive way of treating contextual forms of obligations via a reduction based on $\mathbf{K45^{ij}_n}$ logic. The fact that ϕ is obligatory in context i can be formalized as $[i](\neg\phi \rightarrow V)$ and read as: *the negation of ϕ counts as a violation in context i*. It becomes thus possible to express that ϕ is obligatory in context i while $\neg\phi$ is permitted in context j: $[i](\neg\phi \rightarrow V) \wedge \langle j \rangle (\neg\phi \wedge \neg V)$.

In this perspective, constitutive counts-as statements concerning a violation atom are of particular interest from a theoretical point of view. Formulae such as $\neg\gamma \Rightarrow^{co}_{c,\Gamma} V$ (with $\neg\gamma \rightarrow V \in \Gamma$) denote constitutive rules of normative system Γ defining context c. Such rules constitute the notion of violation for context c, that is, they define what precisely counts as a violation in c. Now, if we consider deontic notions to be sensibly representable via appropriate modal reductions, and if we consider our analysis of counts-as to be a viable basis for such a reduction, then it naturally follows that regulative rules are, in our perspective, nothing but constitutive rules. They are constitutive rules of a special kind since they constitute the most crucial institutional notion, i.e., the notion of violation. This answers the second part of the third research question.

It is worth stressing that this is an original result. To our knowledge, only reduction approaches of constitutive norms to regulative ones have been proposed and not vice versa[13]. We will come back again on this issue in Section 5.7 in the next chapter.

4.7.4 Deontics and contextuality

Interestingly and curiously, a very similar logic to $\mathbf{K45^{ij}_n}$ is used also in [Lomuscio and Sergot, 2000, 2003] in order to express deontic constraints directed to agents.

[13] See [Bulygin, 1992] for a survey of these attempts.

That logic is logic $\mathbf{KD45^{ij}_n}$, i.e., logic $\mathbf{K45^{ij}_n}$ extended with axiom D to rule out the possibility of empty contexts.

In the aforementioned work, the fact that ϕ is obligatory for agent i is expressed via formulae such as $[i]\phi$, that is, logic $\mathbf{KD45^{ij}_n}$ is directly used as a form of deontic logic without any reduction strategy. Indexes refer, in that case, to agents and not to contexts as in our case, and the system is then intended as a multi-agent deontic logic. Nevertheless, this conception of obligations presupposes the understanding of the set of ideal states for some agent as a context, and more precisely, a kind of "ideal context". This makes the nesting of deontic operators uninformative since all nested modalities can be reduced to single modalities (see the discussion axioms 4^{ij} and 5^{ij} in Section 4.2.5). In fact, the authors explicitly mention these features:

> "It is worth pointing out that the criterion for what *counts as* a green state [that is, a state in which no violation occurs] is absolute, that is to say, the set of green states for an agent is independent of the state in which it currently is" ([Lomuscio and Sergot, 2003], p. 15).

It might be interesting to point out the technical and theoretical differences between representing deontic notions directly, like in [Lomuscio and Sergot, 2000, 2003] (i.e., by using $[i]$ as an 'ought-to-be' operator), or by means of a reduction like we did in the previous section (i.e., by using $[i]$ as the operator of an Anderson-like reduction), but using at the same time about the same logical background ($\mathbf{KD45^{ij}_n}$ or $\mathbf{K45^{ij}_n}$). In other words, what is the difference in representing the fact that ϕ is obligatory as $[i]\phi$ or as $[i](\neg\phi \rightarrow V)$[14]?

The difference becomes evident at a semantic level. In [Lomuscio and Sergot, 2000, 2003] it is the the set W_i (see Section 4.2.2), which clusters the set of all ideal states, that is, the states in which $\neg V$ holds with respect to agent i. In our reduction approach, instead, ideal states are clustered inside the set W_i by means of the atom V. In other words, while in [Lomuscio and Sergot, 2000, 2003] contexts are sets of ideal states, in our counts-as based reduction contexts just classify what counts as a violation, and it can therefore be possible that violations occur in some of the states constituting a context. This difference becomes evident in the representation of violation: if $\mathbf{KD45^{ij}_n}$ is directly used as a deontic logic, violations are representable as $[i]\phi \wedge \neg\phi$; while if it is used as basis for a reduction it becomes possible to distinguish between a kind of factual violation V, analogous to the previous one, and a kind of local or contextual violation $\langle i \rangle V$. This is of course related also with the intuitive reading attached to the modal indexes: in [Lomuscio and Sergot, 2000, 2003] each index i corresponds to an agent, while in our work it corresponds to the context defined by a normative system.

[14]The reader should notice the strict analogy between these two representations of deontics and the ones discussed in relation with contextual terminologies in Section 2.5.1.

4.8 Related Work

This section aims at putting our approach in perspective with the other formal approaches to counts-as available in the literature, and to present our results in a more general fashion. For doing this we address four points.

- Firstly, we provide a detailed comparison of our semantics of proper contextual classification with the semantics of the counts-as conditional studied in [Jones and Sergot, 1996] (Section 4.8.1).
- Secondly, we compare the characterizations of classificatory and proper classificatory counts-as with those proposals in the literature which are based on conditional logic: in particular [Jones and Sergot, 1996] and [Governatori et al., 2002; Gelati et al., 2004]. These proposals address counts-as mainly from an axiomatic perspective, identifying a number of syntactic properties that appear to be intuitive for characterizing a counts-as conditional. Because of this, the comparison will be carried out from a structural point of view, that is, by analyzing what kind of properties are enjoyed by the different conditionals (Section 4.8.2).
- Thirdly, some remarks about counts-as and defeasibility are provided (Section 4.8.3), and yet another proposal for the analysis of counts-as, which is advanced in [Boella and van der Torre, 2003; Boella and Van der Torre, 2004; Boella and van der Torre, 2005], is briefly discussed (Section 4.8.4).
- Finally, we analyze how our three versions of counts-as behave with respect to the so-called 'transfer problem' (Section 4.8.5). This provides the last ingredient for a full comparison of our proposal with the one advanced in [Jones and Sergot, 1996].

The section ends by pointing at some issues which we consider worthy of further research.

4.8.1 A model-theoretic comparison

We start by providing a comparison between our analysis and the formal characterization of counts-as proposed in [Jones and Sergot, 1996]. We proceed by comparing our semantics for proper contextual classificatory statements (Section 4.4) with the minimal model semantics of counts-as proposed in [Jones and Sergot, 1996], which is based on minimal conditional models (Mc-models).

Definition 4.13. (Mc-models for counts-as conditionals)
An Mc*-model for counts-as conditionals in the fashion of* [Jones and Sergot, 1996] *is a structure* $\mathcal{M}'$ *such that:* $\mathcal{M}' = \langle W, f_i, \mathcal{I}' \rangle$ *where* $f_i : W \times Pow(W) \longrightarrow Pow(Pow(W))$*, that is, given a world* w *and a set of worlds* X *it assigns a finite set* $\{Y_1, ..., Y_n\}$ *of sets of worlds, and* f_i *is such that, for all* $X, Y, Z \subseteq W$ *and* $w \in W$*:*

1. *if* $Y \in f_i(w, X)$ *and* $Z \in f_i(w, X)$ *then* $Y \cap Z \in f_i(w, X)$*;*

2. *if* $X \in f_i(w, Y)$ *and* $X \in f_i(w, Z)$ *then* $X \in f_i(w, Y \cup Z)$*;*

3. *if* $Y \in f_i(w, X)$ *and* $Z \in f_i(w, Y)$ *then* $Z \in f_i(w, X)$*.*

The satisfaction relation for a generic counts-as operator $\Rightarrow_c$ based on this semantics would run as follows:

$$\mathcal{M}', w \models \gamma_1 \Rightarrow_c \gamma_2 \text{ iff } \mathcal{I}'(\gamma_2) \in f_c(w, \mathcal{I}'(\gamma_1)). \tag{4.40}$$

The semantics of counts-as conditionals just sketched consists therefore in a function assigning, for each world, sets of sets of worlds (i.e., sets of propositions) to sets of worlds (i.e., propositions). It is important to notice though, that the specification of this function is left completely abstract in the sense that nothing is said about what kind of set $\{Y_1, ..., Y_n\}$ is to be expected given a set X (and a world w), i.e, about the kind of relation holding between the arguments and the values of f_c. In fact, only abstract formal constraints are imposed on f_c and there can be a number of different concrete functions obeying those constraints.

We have already seen that the Mc-models semantics validates different principles for the counts-as operator (Section 4.4). In particular, it validates transitivity (third item in Definition 4.13) while the $\text{Cxt}^{\top}$-models semantics for proper contextual classification does not (Proposition 4.6). On the other hand, the Mc-models semantics does not validate cumulative transitivity (cut), which is instead valid in our semantics, and it does not validate cautious monotonicity either, which is invalid also in the $\text{Cxt}^{\top}$-models semantics (Proposition 4.6)[15].

Proposition 4.15. (Mc-models, Cut and Cautious Monotonicity)
Cumulative Transitivity (Cut) and Cautious Monotonicity of $\Rightarrow_c$*, i.e.:*

$$((\gamma_1 \Rightarrow_c \gamma_2) \wedge ((\gamma_1 \wedge \gamma_2) \Rightarrow_c \gamma_3)) \rightarrow (\gamma_1 \Rightarrow_c \gamma_3)$$
$$((\gamma_1 \Rightarrow_c \gamma_2) \wedge (\gamma_1 \Rightarrow_c \gamma_3)) \rightarrow ((\gamma_1 \wedge \gamma_2) \Rightarrow_c \gamma_3)$$

are invalid in the Mc*-models semantics.*

Proof. It is easy to build the desired countermodels. A countermodel for Cut is provided by a model $\mathcal{M}'$ and a world w s.t.: $\mathcal{I}'(\gamma_2) \in f_i(w, \mathcal{I}'(\gamma_1))$ and $\mathcal{I}'(\gamma_3) \in f_i(w, \mathcal{I}'(\gamma_2) \cap \mathcal{I}'(\gamma_1))$ and $\mathcal{I}'(\gamma_3) \notin f_i(w, \mathcal{I}'(\gamma_1))$. An analogous countermodel can be found for Cautious Monotonicity. □

Besides validating different principles the two semantics differ also in other more fine-grained respects. Let us recall Definition 4.8 and spell it out semantically:

$$\begin{aligned}
\mathcal{M}, w \models \gamma_1 \Rightarrow_c^{cl+} \gamma_2 \quad & \text{iff} \quad \mathcal{M}, w \models [c](\gamma_1 \rightarrow \gamma_2) \wedge \neg[u](\gamma_1 \rightarrow \gamma_2) \\
& \text{iff} \quad (\forall w' \in W_c : w' \in \mathcal{I}(\gamma_1) \text{ IMPLIES } w' \in \mathcal{I}(\gamma_2)) \\
& \qquad \text{AND } (\exists w'' \in W : w'' \in \mathcal{I}(\gamma_1) \text{ AND } w'' \notin \mathcal{I}(\gamma_2)) \qquad (4.41) \\
& \text{iff} \quad W_c \cap \mathcal{I}(\gamma_1) \subseteq \mathcal{I}(\gamma_2) \text{ AND } \mathcal{I}(\gamma_1) \not\subseteq \mathcal{I}(\gamma_2) \qquad (4.42)
\end{aligned}$$

[15] All these facts will be of use in Section 4.8.2.

where $\mathcal{M}$ is a $\text{Cxt}^{\top}$-model: $\mathcal{M} = \langle W, W, W_c, \mathcal{I} \rangle$. These equivalences point to a couple of technical differences which nicely show where the most essential theoretical differences between the two approaches lie.

First of all, they show (Equivalence 4.41) that the truth of a proper contextual classification ($\Rightarrow_c^{cl+}$) does not depend on the point of evaluation. In other words, the set of counts-as statements holding in a context does not depend on the state of evaluation in the model. This is not surprising since we have already seen, in Section 4.2, that our semantics presupposes the notion of truth in a context to be of a global kind. Instead, whether the truth of a counts-as in a Mc-model depends or not on the point of evaluation is an issue which is left unaddressed in [Jones and Sergot, 1996]. In fact, although the function f_i in a Mc-model takes the evaluation point as one of its two arguments, that work does not discuss whether the fact that $f_i(w, X) \neq f_i(w', X)$ (i.e., the set of counts-as statements of normative system i in w is different from the set of those in w') would model something meaningful at all. In our view it would not, since what determines the set of counts-as statements of a context is only the context itself and not the world. If $f_i(w, X) \neq f_i(w', X)$, then i denotes, in fact, in w and w' two different normative systems. Noticeably, this means that the approach held in [Jones and Sergot, 1996] allows for a given index to denote different normative systems in different worlds. However, the rationale for this choice is not discussed.

Second, they show that the truth of a proper contextual classification is a function of the context W_c and of the interpretation function $\mathcal{I}$. In fact, the truth of $\Rightarrow_c^{cl+}$-formulae is determined by a set-theoretical relation between the evaluation of the antecedent ($\mathcal{I}(\gamma_1)$), the evaluation of the consequent ($\mathcal{I}(\gamma_2)$), the context (W_i) of the counts-as and the universal context W (Equivalence 4.42). As a consequence, the truth of a $\Rightarrow_c^{cl+}$-formula is logically related to the truth of its antecedent and consequent. In fact, it is easy to see that the following holds on the grounds of Formula 4.42:

$$(w \in W_c \text{ AND } \mathcal{M}, w \models \gamma_1 \Rightarrow_c^{cl+} \gamma_2 \text{ AND } \mathcal{M}, w \models \gamma_1) \text{ IMPLIES } \mathcal{M}, w \models \gamma_2 \tag{4.43}$$

where $\mathcal{M}$ is a $\text{Cxt}^{\top}$-model: $\mathcal{M} = \langle W, W, W_c, \mathcal{I} \rangle$. That is to say, if we are in context W_c, and that context properly classifies γ_1 as γ_2, and it is the case that γ_1, then it is also the case that γ_2. To put it another way, this shows how counts-as statements, interpreted as proper contextual classifications, have an influence on what holds in a world[16].

In Mc-models, on the contrary, function f_i is in no way related to the evaluation function $\mathcal{I}$ and the index i does not get a concrete denotation in the model like in $\text{Cxt}^{\top}$-models. Therefore, no formal relation such as the one in Formula 4.43 between the counts-as statement, the context, the antecedent and the consequent can be inferred. Intuitively, this means that the counts-as statements holding in a world, and the truth of their antecedents and consequents are completely independent from each other. We find this a quite counter-intuitive idea, since the role of counts-as

[16]Notice that this is a proper classificatory version of what stated in Proposition 4.10.

statements is exactly to allow the logical connection of formulae which are otherwise logically unrelated.

This section has shown that what our semantics based on $\text{CXT}^{\top}$-models adds to the proposal in [Jones and Sergot, 1996] amounts to three essential aspects: first, counts-as statements are of a global kind, i.e., their truth is independent of the point of evaluation in a model; second, indexes in counts-as statements have a precise semantics, i.e., they are contexts (sets of situations); third, counts-as statements correspond to precise set-theoretical relations between their context, their antecedent, their consequent and the universal context. All these features are nothing but the formal translation of the intuition from which our whole approach moves: counts-as statements represent the way a normative system classifies situations.

4.8.2 A structural comparison with other approaches

Building on the results stated in Propositions 4.5, 4.6 and 4.7, this section and the two following ones discuss our approach from a merely structural perspective, showing which properties of $\Rightarrow_c^{cl}$ and $\Rightarrow_c^{cl+}$ are accepted or rejected in the approaches developed in [Jones and Sergot, 1996] and [Governatori et al., 2002; Gelati et al., 2004].

Constitutive conditionals are not included in this comparison since, as observed in Section 4.5.5, they cannot be properly studied from a structural perspective. In a way, they are therefore radically different from the proposals available in the literature.

An overview of the main properties enjoyed by each characterization is provided in Table 4.8.2. With "1" we denote that the notion of counts-as in the column enjoys the property in the row, with "0" vice versa.

Proposition 4.5 showed that $\Rightarrow_c^{cl}$ enjoys strong properties (in particular reflexivity, antecedent strengthening, and transitivity) and displays, therefore, a very classical behavior. As shown by Proposition 4.6 and 4.7, the logic of $\Rightarrow_c^{cl+}$ behaves instead much less classically rejecting reflexivity, strengthening of the antecedent, even in the weaker version of cautious monotonicity, and transitivity. On the other hand, it still retains a weaker form of transitivity, namely cumulative transitivity.

The approach proposed in [Jones and Sergot, 1996], which has already been discussed from a semantic point of view in Section 4.8.1, develops a logic for counts-as conditionals (denoted by the operator $\Rightarrow_c$) obeying the following principles: left logical equivalence ($\Rightarrow_c$-version of Formula 4.19), right logical equivalence ($\Rightarrow_c$-version of Formula 4.18), disjunction of antecedents ($\Rightarrow_c$-version of Formula 4.21), conjunction of the consequents ($\Rightarrow_c$-version of Formula 4.20) and transitivity ($\Rightarrow_c$-version of Formula 4.16). Recall, though, that it does not enjoy cumulative transitivity and cautious monotonicity (Proposition 4.15).

In [Governatori et al., 2002; Gelati et al., 2004] it is argued instead that the logic of counts-as conditionals, which they denote via the operator $\Rrightarrow$, amounts to the logic of preferential reasoning ([Kraus et al., 1990]), preferential reasoning being characterized by the following properties: reflexivity ($\Rrightarrow$-version of Formula 4.13), left logical equivalence ($\Rrightarrow$-version of Formula 4.19), weakening of the consequent

		$\Rightarrow^{cl}_{c}$	$\Rightarrow^{cl+}_{c}$	$\Rightarrow_{c}$	$\Rrightarrow$
A	Reflexivity	1	0	0	1
B	Antecedent Strengthening	1	0	0	0
C	Transitivity	1	0	1	0
D	Disjunction of the Antecedents	1	1	1	1
E	Conjunction of the Consequents	1	1	1	1
F	Left Logical Equivalence	1	1	1	1
G	Right Logical Equivalence	1	1	1	1
H	Consequent Weakening	1	0	0	1
I	Cumulative Transitivity	1	1	0	1
L	Cautious Monotonicity	1	0	0	1

Table 4.1: Properties of counts-as operators

($\Rrightarrow$-version of Formula 4.15), conjunction of the consequents ($\Rrightarrow$-version of Formula 4.20), cut ($\Rrightarrow$-version of Formula 4.23), cautious monotonicity ($\Rrightarrow$-version of Formula 4.17) and disjunction of the antecedents ($\Rrightarrow$-version of Formula 4.21)[17].

This overview provides grounds for a number of interesting observations. First of all, notice that there seems to be a structural hard core of all characterizations of counts-as including ours, which corresponds to properties from D to G. These properties are exactly the ones recognized as a sort of minimal characterization of counts-as in [Jones and Sergot, 1996]. There are then two remarkable facts to be noticed, which concern the relation between our notions of contextual and proper contextual classification and the notions of counts-as axiomatically characterized in [Governatori et al., 2002; Gelati et al., 2004] and [Jones and Sergot, 1996]. We discuss them separately in the following two sections.

Operator $\Rrightarrow$ corresponds to a defeasible $\Rightarrow^{cl}_{c}$

The notion of counts-as statements as conditional counterparts of preferential reasoning ([Governatori et al., 2002; Gelati et al., 2004]) represents a defeasible form of our notion of contextual classification, since the only properties distinguishing the two notions are strengthening of the antecedent (B) and transitivity (C), which in presence of reflexivity (A) and cut (I) are actually equivalent (see [Kraus et al., 1990]).

[17]To be precise, in [Governatori et al., 2002] it is argued that the logic of counts-as corresponds to preferential reasoning, while in [Gelati et al., 2004] it is considered to correspond *at least* to cumulative reasoning, i.e., preferential reasoning without the property of disjunction of the antecedents (see [Kraus et al., 1990]).

In the light of our semantics-driven analysis of counts-as, this constitutes a very interesting fact. In a way, it allows us to attach a precise meaning to the notion of counts-as axiomatized in [Governatori et al., 2002; Gelati et al., 2004] deriving it from the notion of contextual classification or enthymeme (see Section 4.7.1): if the statement "X counts-as Y in context C", intended as contextual classification, means "X is classified as Y in C", then the same statement read in the fashion of [Governatori et al., 2002; Gelati et al., 2004] would mean "X is classified as Y in C, *modulo exceptions*", or "it *normally* follows from C that X is classified as Y". Yet another possible meaning of counts-as statements is therefore disentangled which we might call *defeasible contextual classification*. The logic of this notion was already studied, from an axiomatic perspective, in [Governatori et al., 2002; Gelati et al., 2004] but it can now get a precise place within the map of the many senses of the term "counts-as" we are sketching here.

Operator $\Rightarrow_c$ as an axiomatic approximation of $\Rightarrow_c^{cl+}$

The notion of proper contextual classification appears to correspond to a slightly weaker version of the counts-as conditional studied in [Jones and Sergot, 1996] where transitivity (C) is substituted by the weaker property of cumulative transitivity (I).

In fact, $\Rightarrow_c^{cl+}$ does not validate transitivity while $\Rightarrow_c$ does. However, in [Jones and Sergot, 1996] the transitivity of counts-as is not accepted with strong conviction:

> "[. . .] we have been unable to produce any convincing counter-instances [of transitivity] and are inclined to accept it" ([Jones and Sergot, 1996], p.436).

Our analysis shows instead that once we first proceed to the isolation of the exact sense of the term "counts-as" we are aiming at formalizing, no room for uncertainty is then left about the syntactic properties enjoyed by the formalized notion: if we intend counts-as statements as proper contextual classifications, then transitivity must be rejected on the grounds of mere logical reasons.

The notion of proper contextual classification stemmed from the need to express the idea of new classifications which are brought about by contexts (Section 4.3). The question remains whether the approach developed in [Jones and Sergot, 1996] aimed at formalizing yet a different meaning of counts-as statements, like the approach in [Governatori et al., 2002; Gelati et al., 2004] did (see the previous section), or whether it was actually aiming at axiomatizing proper contextual classification. In this case the acceptance of transitivity would have been led by the sort of misunderstandings at which we pointed in Section 4.6 discussing Proposition 4.11. We favor indeed the second hypothesis, on the grounds of the following observations. We read in [Jones and Sergot, 1996]:

> "Even if it were to transpire that convincing counter-examples to S [read transitivity] could be found, a weakened form of transitivity:
>
> $$(A \Rightarrow_c B) \rightarrow ((A \Rightarrow_c C) \rightarrow D_c(A \rightarrow B))$$

will nevertheless be a truth of the logic" ([Jones and Sergot, 1996], p.436).

In that work the operator D_c is the operator of a multi-modal $\mathbf{KD_n}$ logic and it aims at capturing a notion of "general institutional constraints". Now, if we interpret the D_c operator as our $[c]$ operator in logic $\mathbf{Cxt^u}$, and the $\Rightarrow_c$ operator as our $\Rightarrow_c^{cl+}$ operator, then the formula above is nothing but Formula 4.32, which was proven to be valid in logic $\mathbf{Cxt^u}$ (Proposition 4.11). The strict implication under a D_c operator would correspond to contextual classification in context c. Such a move is not arbitrary since, in that work, logic $\mathbf{KD_n}$ was explicitly considered to be a "provisional proposal" ([Jones and Sergot, 1996], p.437). We hope to have shown that logic $\mathbf{Cxt^u}$ could then be seen as a natural strengthening of $\mathbf{KD_n}$ to model contexts[18].

Remarkably, the very same observation can be made for another crucial constraint on $\Rightarrow_c$, which was intended in [Jones and Sergot, 1996] to relate the notion of counts-as with the notion of "general institutional constraint":

$$A \Rightarrow_c B \rightarrow D_c(A \rightarrow B).$$

Again, substituting $\Rightarrow_c$ and D_c with $\Rightarrow_c^{cl+}$ and respectively $[c]$, another validity of our framework is obtained, i.e., Formula 4.30 (Proposition 4.11).

To recapitulate, [Jones and Sergot, 1996] did not consider transitivity to be ultimately established as an essential constraint for a characterization of counts-as. On the other hand they did consider essential two validities of our system (Formula 4.32 and Formula 4.30) expressing a logical relation between proper contextual classification and contextual classification. Furthermore, they provisionally assumed logic $\mathbf{KD_n}$ as a logic for expressing institutional constraints, suggesting that stronger logics might actually work better.

On this grounds, it becomes tempting to claim that what authors in [Jones and Sergot, 1996] tried to axiomatize was actually the notion of proper contextual classification. And what they meant under the label "general institutional constraint" was nothing but the notion of truth in a context, i.e., what we have here represented via the $[c]$ operators. The temptation is even stronger if we consider that, as a matter of fact, our modal logic analysis started by taking seriously the intuitive reading they attach to counts-as statements at the beginning of their paper, and which has been already quoted in Section 4.1.2: "certain states of affairs of a given type count as, or *are to be classified as*, states of affairs of another given type" ([Jones and Sergot, 1996], pag. 431).

4.8.3 Defeasibility and counts-as: a note

We have dealt with the structural properties enjoyed by counts-as operators, discussing various answers to the question: "what properties should a genuine formal

[18]To be precise, $\mathbf{Cxt^u}$ would be a conservative extension of $\mathbf{KD_n}$ if it contained axiom D, which we rejected in order to leave the possibility open of representing inconsistent contexts. However, as we already observed in Section 4.2.4, this does not constitute an essential feature of our approach.

characterization of counts-as obey?" One such question in particular deserves some more considerations: "is counts-as defeasible?"

In discussing Proposition 4.11, we have already noticed that questions of this type can easily mislead the formal analysis when the to-be-analyzed notion displays a high level of vagueness and when there are reasons for believing that the name commonly attributed to that notion (counts-as) can actually hide not one but more notions (at least three in our case).

There is also a second potential source of misunderstanding in this kind of questions. When we ask whether counts-as is or is not defeasible, are we asking whether the representation of counts-as via a specific operator of the object-language enjoys antecedent strengthening, or whether antecedent strengthening is enjoyed by the logical consequence relation defined on the formulae of that object language, that is, whether the consequence relation of the logic is monotonic? In other words, are we asking whether a specific sense of the term "counts-as" inherently enjoys antecedent strengthening, or whether the reasoning we perform on counts-as statements is monotonic? Without the proof of a deduction theorem linking a counts-as operator with a corresponding consequence relation, these two questions are logically independent. However, the literature on counts-as has never emphasized this difference with the necessary precision.

As a matter of fact, logics with counts-as operators rejecting antecedent strengthening, such as our $\mathbf{Cxt^u}$ for $\Rightarrow_c^{cl+}$ or the logic developed in [Jones and Sergot, 1996], feature monotonic consequence relations. In [Governatori et al., 2002; Gelati et al., 2004], instead, the counts-as operator rejects antecedent strengthening and, in addition, the consequence relation of the logic is also non-monotonic.

On the other hand, an operator enjoying antecedent strengthening can be embedded in a logic endowed with a non-monotonic consequence relation. For instance, our logic $\mathbf{K45_n^{ij}}$ for $\Rightarrow_c^{cl+}$ could easily be merged in a suitable argumentation system ([Prakken and Vreeswijk, 2002]) providing the desired defeasible inferential properties, or, as it has been done in [Meyer and van der Hoek, 1995], it can be extended including mechanisms for representing default reasoning. This is also what happens in [Boella and van der Torre, 2003; Boella and Van der Torre, 2004; Boella and van der Torre, 2005], where the defeasibility of counts-as is intended as the non-monotonicity of the operation which extracts consequences from a given input plus the constitutive rules of a given system (see Section 4.8.4). To say it with [Sergot, 2004]:

> "Viewed as a kind of conditional, X counts as Y is not a *defeasible* conditional, though in practice the laws defining when X counts as Y holds will often be naturally formulated as defeasible general rules subject to exceptions" ([Sergot, 2004], p.72).

All these choices can have a precise rationale. However, what we want to stress here is that the two issues are radically different in nature. The first concerns the set of validities involving counts-as statements (in the various possible senses of the term "counts-as"), i.e., the logical properties of counts-as statements as they have been studied in Propositions 4.5 and 4.7. The second concerns instead the way statements can soundly (in the various possible senses of the term "soundly") be inferred from other statements, and in this case from counts-as statements, i.e., the possible reasoning patterns involving counts-as. To use a philosophical terminology, the first issue is of an *ontological* kind (what are the validities concerning counts-as statements?), while the second is of an *epistemological* one (what can be inferred on the grounds of counts-as statements?). In this work we have addressed the first issue.

4.8.4 Some words on yet another approach

In the structural comparison exposed in Section 4.8.2 we did not take into consideration the formal approach to counts-as proposed in [Boella and van der Torre, 2003; Boella and Van der Torre, 2004; Boella and van der Torre, 2005]. In that work, counts-as is investigated as an ingredient within a broader attempt to formalize normative systems as wholes, and especially the interaction between constitutive and regulative rules. There are two essential aspects of that proposal which make it difficult to compare it with ours in the fashion followed above for [Jones and Sergot, 1996] and [Governatori et al., 2002; Gelati et al., 2004].

First of all, counts-as statements represent in that work forms of consequence relation statements: X counts-as Y in context C iff from the input X and the context C it can be inferred, via the rules of the normative system, that Y. In effect, they do not propose a framework for studying the properties of counts-as statements —there is nothing such as a counts-as operator in their language— but rather a framework for drawing conclusions via the constitutive rules of a normative system. To put it another way, they are not interested in expressing that counts-as statements enjoy reflexivity, transitivity etc., but just that given a set of constitutive rules and an input a certain output follows.

Thus, in their view, counts-as statements represent the results of a reasoning process based on the constitutive rules of a given normative system. However — and this is the second aspect— no precise logic is chosen for specifying the reasoning process giving rise to the counts-as statements. Instead, a general framework based on input/output logics ([Makinson and van der Torre, 2000]) is proposed, within which a number of different ways of drawing conclusions from the rules of the system can be specified. In the end, neither an actual proposal for the reasoning style grounding counts-as statements is set forth, apart from the rejection of reflexivity, nor the issue is addressed about what intuitive notion of counts-as would correspond to each of the possible inference styles which are specifiable in the framework of input/output logics.

4.8.5 The *transfer problem* in the light of $\Rightarrow^{cl}_{c}$, $\Rightarrow^{cl+}_{c}$ and $\Rightarrow^{co}_{c,\Gamma}$

The 'transfer problem' has been introduced in [Jones and Sergot, 1996] as a landmark for testing the intuitive adequacy of formalizations of counts-as. It can be exemplified as follows: suppose that somebody brings it about —for instance by coercion— that a priest effectuates a marriage, does this count as the creation of a state of marriage? Does anything implying that a priest effectuates a marriage count as the creation of a state of marriage? In other words, is the possibility to create a marriage transferable to anybody who brings it about that the priest effectuates the ceremony? In our framework, these questions get a triple formulation, one for each of the different senses of counts-as.

The transfer problem and $\Rightarrow^{cl}_{c}$

In [Jones and Sergot, 1996], the transfer problem has been used as grounds for the rejection of the property of antecedent strengthening for counts-as conditionals. It is beyond doubt that a characterization of counts-as which enjoys the strengthening of the antecedent also exhibits the transfer problem: if that property holds, then the fact that the performance of the ceremony counts as the creation of a state of marriage implies that also a coerced performance does. As already noticed in [Grossi et al., 2005d], contextual classification ($\Rightarrow^{cl}_{c}$), which enjoys the strengthening of the antecedent (Proposition 4.5), does exhibit the transfer problem: whatever situation in which a priest performs a marriage ceremony is classified as a situation in which a marriage state comes to be. And this is precisely what we intuitively expect given the notion of contextual classification as informally introduced in Section 4.1. In other words, contextual classification *should* exhibit the transfer problem or, to put it another way, it should display a *transfer property*: the determining of a state of marriage should be transferable to any state in which a priest performs the ceremony.

The transfer problem and $\Rightarrow^{cl+}_{c}$

It has been shown that the characterization of proper contextual classification ($\Rightarrow^{cl+}_{c}$) does not enjoy the strengthening of the antecedent (Proposition 4.6). From a mere conditional logic perspective, such as the one assumed in [Jones and Sergot, 1996], this would be enough to rule out the occurrence of the transfer problem.

However, it seems this is quite not the case, the reason being that the transfer problem has manifestations which go beyond the structural rule of antecedent strengthening. The follwing formula, proven valid in Proposition 4.7, also expresses an instance of the transfer problem:

$$\neg[u](\gamma_1 \rightarrow \gamma_3) \rightarrow ((\gamma_1 \Rightarrow^{cl+}_{c} \gamma_2) \wedge (\gamma_2 \Rightarrow^{cl+}_{c} \gamma_3)) \rightarrow (\gamma_1 \Rightarrow^{cl+}_{c} \gamma_3)$$

Intuitively, this formula expresses what follows. If the fact that a priest effectuates a marriage (γ_1) under coercion of a third party (γ_3) is not globally classified as giving rise to a state of marriage (γ_2) —which is the case, given the intuitive reading of the scenario at issue— then it is safe to say that if the priest's performance of the marriage counts as (in a proper classificatory sense) a marriage, then a coerced performance of the marriage counts also as a marriage.

Notice that this is again something perfectly intuitive given the assumptions about proper contextual classification exposed in Section 4.3: if a context c makes a classification $\gamma_1 \rightarrow \gamma_2$ true, which does not hold in general, then also the strengthened version of it $\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2$ is true in that context. Besides, if the strengthened version is also not true in general, it then follows that $\gamma_1 \wedge \gamma_3 \rightarrow \gamma_2$ is also a novel classification which is brought about by context c. Exhibiting the transfer problem is also for proper contextual classification not problematic.

From a technical point of view, Proposition 4.7 shows that a characterization of counts-as, which does not enjoy the strengthening of the antecedent, can still exhibit the transfer problem. This is a point worth stressing because, by assuming a purely conditional perspective like in [Jones and Sergot, 1996], instances of the transfer problem such as the one represented in the above formula could simply not be expressed.

To conclude, proper contextual classification does not exhibit the transfer problem, if by "transfer problem" we just mean the rejection of antecedent strengthening, like it was proposed in [Jones and Sergot, 1996]. On the other hand, if we consider broader forms of the problem which did not get a formulation in [Jones and Sergot, 1996], then proper contextual classification does exhibit them.

The transfer problem and $\Rightarrow^{co}_{c,\Gamma}$

The constitutive reading of counts-as statements does not exhibit any of the considered forms of the transfer problem. Counts-as statements represent the rules specifying a normative system. So, all that it is explicitly stated by the 'institution of marriage' is that if the priest performs the ceremony then the couple is married. No rule belongs to that institution which states that the action of a third party bringing it about that the priest performs the ceremony also counts as a marriage. Our formalization fully captures this feature. Let the 'marriage institution' c be represented by the set of rules $\Gamma = \{p \rightarrow m\}$, i.e., by the rule "if the priest performs the ceremony, then the couple is married". Let then t represent the fact that a third party brings it about that p. For Definition 4.10 the counts-as $(t \wedge p) \Rightarrow^{co}_{c,\Gamma} m$ is just an undefined expression, because $((t \wedge p) \rightarrow m) \notin \Gamma$, that is, because the 'marriage institution' does not state such a classification.

4.8.6 Future work: towards a logic of (legislative) rulings

Even though the chapter has analyzed the notion of constitutive rule from a static perspective, it is undeniable that the idea of constitution hides also a dynamic flavour. To constitute means, somehow, to *bring about* something new. We have captured this idea of novelty in the notion of proper classification by comparing what is considered to be necessarily the case (truth in the global context u) with what is instead the case within a given context, possibly defined by a set of rules (truth in a context c defined by a set of forumlae Γ).

Now, this idea could be used as grounds for modeling also the dynamic flavor of constitution. This could be achieved by introducing a dynamics in the framework obtained by 'context definition' which we could denote as: $c := \Gamma$. These actions, which we can intuitively view as (legislative) rulings, could be read as: "let context c

be defined by the (finite) set of formulae Γ" or, using the terminology of [Smith, 2001], "fiat c (by Γ)". Syntactically, such actions could be thought of as promulgations of the form $[c]\Gamma \wedge [-c]\neg\Gamma$, that is, 'all that it holds'-statements (e.g., "all that it holds in c is Γ"). The effects of such actions would then be the creation of a context c consisting of all and only the states satisfying Γ.

In models based on $\mathrm{Cxt}^{\top,\setminus}$ frames, the semantics of such actions would amount to the modification of the cluster W_c or, equivalently, of the accessibility relation R_c in the corresponding secondarily universal frame. Typically, it would happen that a non-existent context is brought to life, so to say, by carving it out the realm of (logical) possibilities: from $W_c = W$ to $W_c = \{w \mid w \models \Gamma\}$. Similarly, a context can be modified by extending or restricting its definition. Update logics (see for instance [van Benthem et al., 2006]) are the natural formal environment for such investigations.

We are convinced that this research line would complete the formal picture of how constitution, and therefore normative systems, work.

4.9 Conclusions

Moving from hints provided by the literature on legal and social theory concerning constitutive rules, the paper has analyzed counts-as statements as forms of contextual classifications. This analytical option, which we have studied from a formal semantics perspective, has delivered three semantically precise senses (Definitions 4.4, 4.8 and 4.10) in which counts-as statements can be interpreted, which we called *classificatory*, *proper classificatory* and *constitutive* readings. The three readings have then been formally analyzed in modal logic.

The classificatory reading resulted in a strong logic of counts-as conditionals enabling many properties which are typical of reasoning with concept subsumptions such as, in particular, reflexivity, strengthening of the antecedent and weakening of the consequent (Proposition 4.5). In fact, the logic obtained could be thought of as a modal logic version of the logic of contextual subsumptions which has been investigated in Chapter 2. It has been shown (Section 4.8.2), that this notion is a close relative of the counts-as studied in [Governatori et al., 2002; Gelati et al., 2004] which constitutes, from a structural point of view, the defeasible version of contextual classification. This is not surprising if we consider, as shown in Section 4.7.1, that contextual classification corresponds to a specific notion of logical consequence relation (Definition 4.11) which constitutes a well-known bridge between monotonic and non-monotonic logics, and which is related to the notion of enthymeme.

The characterization of proper contextual classification resulted, instead, in a much weaker logic rejecting reflexivity, transitivity and antecedent strengthening (Proposition 4.6), but retaining cumulative transitivity (Proposition 4.7). Noticeably, it has been shown (Section 4.8.2) that this notion corresponds to the counts-as characterized in [Jones and Sergot, 1996] once transitivity is substituted with cumulative transitivity. We claimed indeed that the axiomatization proposed in [Jones and Sergot, 1996] was aiming at capturing precisely the notion of proper contextual classification. Also the semantics of counts-as conditionals proposed in [Jones and Sergot, 1996] has been subject to thorough investigation and its theoretical shortcomings emphasized (Section 4.8.1). Finally, the notion of proper contextual classification has offered some new insights on the transfer problem (Section 4.8.5) showing that it cannot be genuinely avoided just by means of rejecting the strengthening of the antecedent in a conditional logic setting.

The formal analysis of constitutive counts-as has provided a formal characterization of the notion of constitution as the definition of a context by a set of rules. This formal characterization has made explicit how constitutive rules provide logical grounds for attributing institutional properties to situations (Proposition 4.10). The constitutive reading of counts-as has also been shown to imply the two classificatory readings (Proposition 4.11). Other logical interrelationships between the three notions of counts-as have also been studied (Propositions 4.12-4.14) showing that the logical relations between them could actually be grounds for fallacies in the formal characterization of counts-as once the polysemy of the term "counts-as" is overlooked.

All in all the main contribution of the chapter consists, in our view, in showing how the formal systematization of the notion of counts-as can be grounded on a very simple intuition about what counts-as statements actually mean, i.e., subsumptions of state types.

Part II

Institutions and Organizations in MAS

Chapter 5

Institutions as TBoxes

"One way is to make it so simple that there are obviously no deficiencies and the other way is to make it so complicated that there are no obvious deficiencies."

C. A. R. Hoare, The 1980 ACM Turing Award Lecture: "The Emperor's Old Clothes", p. 81

In MASs the application of the organizational and institutional metaphors to system design has proven to be useful for the development of methodologies and tools ([Vázquez-Salceda, 2004; Dignum, 2003]). In many cases, however, the application of these conceptual apparatuses amounts to mere heuristics guiding the high level design of the systems. It is our thesis that the application of those apparatuses can be pushed further once their key concepts are treated formally, that is, once notions such as norm, role, structure, etc. obtain a formal semantics. This has been the case for agent programming languages after the relevant concepts borrowed from folk psychology (belief, intention, desire, knowledge, etc.) have been addressed in comprehensive formal logical theories such as, for instance, BDI^{CTL} ([Rao and Georgeff, 1991]) and KARO ([Meyer et al., 2001]). As a matter of fact, those theories have fostered the production of architectures and programming languages.

What is lacking at the moment for the design and development of MASs is, in our opinion, something that can play the role that BDI-like formalisms have played for the design and development of single-agent architectures. Aim of the chapter is to fill this gap with respect to the notion of institution, providing formal foundations for the application of the institutional metaphor to the design of MASs.

The aim of the chapter is to show that a number of key institutional notions can be formally analyzed by means of relatively simple logical languages —description logics— interpreted on labeled transition systems. In a way, the study presented in this chapter can be viewed as an attempt to squeeze a number of relevant institutional notions within description logic constructs. The analytical and formal results presented in Chapter 2 and Chapter 4 provide the starting point of the results presented here.

The chapter extends work presented in [Grossi et al., 2006a] and [Grossi et al., 2007]. It is structured according to the following outline. Section 5.1 provides some analytical and formal preliminaries motivating the key thesis of the chapter: institutions = terminologies. Section 5.2 provides an account of the issue of abstract norms and of the notion of role. In Section 5.3 we provide a funny intermezzo about the discrepancies that can arise between different concrete versions of the same abstract norms. Section 5.4 addresses some computational issues of the formalism presented. Sections 5.5 and 5.6 deal with the notion of institutional infrastructure and, respectively, the issue of norm implementation. Section 5.7 provides a recapitulation of the general view on norms which we hold in this work and which finds broad application in this chapter. A section on related work follows and some conclusive remarks on the results of the chapter are provided in Section 5.9.

5.1 Preliminaries

The section introduces the notion of institution as terminological box.

5.1.1 Institutions

In Chapter 1, it has been made clear that the present work presupposes the normative system perspective on institutions and that normative systems are thought of as the imposition of social or institutional terminologies over 'brute' ones. Let us quote [Pufendorf, 1688] again:

> "Now, as the original manner of producing physical entities is creation, there is hardly a better way to describe the production of moral entities than by the word *'imposition'* [impositio]. For moral entities do not arise from the intrinsic substantial principles of things but *are superadded to things already existent and physically complete*" ([Pufendorf, 1688], pp. 100-101).

At this point, the step toward eInstitutions is natural. eInstitutions impose properties on the possible states of a MAS: they specify what are the states in which an agent *i* enacts a role *r*; what are the states in which a certain agent is violating the norms of the institution, etc. They do this via linking some institutional properties of the possible states and transitions of the system (e.g., agent *i* enacts role *r*) to some brute properties of those states and transitions (e.g., agent *i* performs protocol No.56). An institutional property is therefore a property of system states or system transitions (i.e., a state type or a transition type) that does not belong to a merely technical, or factual, description of the system.

It is worth stressing that, although the notion of "imposition" might suggest that a system should already be known in order to design an e-Institution, nothing like this is presupposed by our perspective. In fact, the design of an e-Institution (see for instance [Vázquez-Salceda, 2004; Vázquez-Salceda et al., 2004; Aldewereld et al., 2006a]) would most probably move from the specification of a set of norms, which would only in a second time be "imposed" on a system. At the time of the normative specification of the institution the system is yet to be designed and the 'brute' vocabulary for its description might yet be unknown. Nevertheless, the final specification of the institution will connect a 'brute' system reality to an 'institutional' one.

To sum up, institutions are viewed as sets of norms (normative system perspective), and norms are thought of as the imposition of an institutional description of the system upon its description in terms of brute properties. In a nutshell, *institutions are impositions of institutional terminologies upon brute ones*. Section 5.2 will provide a formal analysis of this thesis and show its explanatory power in delivering a rigorous understanding of key features of institutions. The next section introduces the DL we are going to work with.

5.1.2 A very expressive DL

The description logic language enabling the necessary expressivity expands the standard description logic language $\mathcal{ALC}$ (see Section 2.1.2) with relational operators ($\sqcup$,$\circ$,$\neg$,id) to express complex transition types, and relational hierarchies ($\mathcal{H}$) to express inclusion between transition types. This language extends also language $\mathcal{ALCH}^{(\neg)}$ which has been use in Chapter 2. Following a notational convention common within DL we denote this language with $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$.

Definition 5.1. *(Syntax of $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$)*
transition types and state type constructs are defined by the following BNF:

$$\alpha := a \mid \alpha \circ \alpha \mid \alpha \sqcup \alpha \mid \neg\alpha \mid id(\gamma)$$
$$\gamma := c \mid \bot \mid \neg\gamma \mid \gamma \sqcap \gamma \mid \forall\alpha.\gamma$$

where a and c are atomic transition types and, respectively, atomic state types.

It is worth providing the intuitive reading of a couple of the operators and the constructs just introduced. In particular $\forall\alpha.\gamma$ has to be read as: "after all executions of transitions of type α, states of type γ are reached". The operator $\circ$ denotes the concatenation of transition types. The operator *id* applies to a state description γ and yields a transition description, namely, the transition ending in states of type γ. It is the description logic variant of the test operator in Dynamic Logic ([D. Harel amd Kozen and Tiuryn, 1984]). Notice that we use the same symbols $\sqcup$ and $\neg$ for denoting the boolean operators of disjunction and negation of both state and transition types.

In our formalizations, we will work with non-logical alphabets which exhibit some structure. These alphabets contain state types and transition types. Atomic state types are obtained by indexing elements of a set *c_Form* (state type forms) with elements of *Agents*, and atomic transition types are obtained by indexing elements of *a_Form* (transition type forms) with elements of $(Agents \times Agents) \backslash \{(i, j) \mid i = j\}$, i.e., pairs of different agents. As a result, we have that atomic state types c are indexed by an agent identifier i in order to express agent properties (e.g., dutch(i)), and atomic transition types a are indexed by a pair of agent identifiers (i, j) (e.g., PAY(i, j)) denoting the actor and, respectively, the recipient of the transition. Obviously, other more complex forms of indexing are straightforwardly definable. By removing the agent identifiers from state types and transition types we obtain state type forms (e.g., dutch) and transition type forms (e.g., PAY).

Definition 5.2. *(Non-logical alphabets)*
The non-logical alphabet is built from three sets: a finite set Agents of agents, a finite set c_Form of state type forms, and a finite set a_Form of transition type forms. Atomic state and transition types are built as follows:

$$c := \underline{c}(i)$$
$$a := \underline{a}(i, j)$$

where $i, j \in Agents$, $\underline{c} \in c\_Form$ *and* $\underline{a} \in a\_Form$.

A terminological box (henceforth TBox) $T = \langle \Gamma, A \rangle$ consists of a finite set Γ of state type inclusion assertions ($\gamma_1 \sqsubseteq \gamma_2$), and of a finite set A of transition type inclusion assertions ($\alpha_1 \sqsubseteq \alpha_2$).

The semantics of $\mathcal{ALCH}^{(\sqcup, \circ, \neg, id)}$ is given in terms of interpreted transition systems ([van Benthem et al., 1994]). As usual, state types are interpreted as sets of states and transition types as sets of state pairs.

Definition 5.3. *(Semantics of* $\mathcal{ALCH}^{(\sqcup, \circ, \neg, id)}$*)*
An interpreted transition system (or model) m for $\mathcal{ALCH}^{(\sqcup, \circ, \neg, id)}$ *is a structure* $\langle S, \mathcal{I} \rangle$ *where S is a non-empty set of states and* $\mathcal{I}$ *is a function such that:*

$$\begin{array}{rcl}
\mathcal{I}(c) & \subseteq & S \\
\mathcal{I}(a) & \subseteq & S \times S \\
\mathcal{I}(\bot) & = & \emptyset \\
\mathcal{I}(\neg\gamma) & = & S \backslash \mathcal{I}(\gamma) \\
\mathcal{I}(\gamma_1 \sqcap \gamma_2) & = & \mathcal{I}(\gamma_1) \cap \mathcal{I}(\gamma_2) \\
\mathcal{I}(\forall \alpha.\gamma) & = & \{s \in S \mid \forall t, (s, t) \in \mathcal{I}(\alpha) \Rightarrow t \in \mathcal{I}(\gamma)\} \\
\mathcal{I}(\alpha_1 \sqcup \alpha_2) & = & \mathcal{I}(\alpha_1) \cup \mathcal{I}(\alpha_2) \\
\mathcal{I}(\neg\alpha) & = & S \times S \setminus \mathcal{I}(\alpha) \\
\mathcal{I}(\alpha_1 \circ \alpha_2) & = & \{(s, s'') \mid \exists s', (s, s') \in \mathcal{I}(\alpha_1) \;\&\; (s', s'') \in \mathcal{I}(\alpha_2)\} \\
\mathcal{I}(id(\gamma)) & = & \{(s, s) \mid s \in \mathcal{I}(\gamma)\}
\end{array}$$

An interpreted transition system m is a model of a state type inclusion assertion $\gamma_1 \sqsubseteq \gamma_2$ *if* $\mathcal{I}(\gamma_1) \subseteq \mathcal{I}(\gamma_2)$. *It is a model of a transition type inclusion assertion* $\alpha_1 \sqsubseteq \alpha_2$ *if* $\mathcal{I}(\alpha_1) \subseteq \mathcal{I}(\alpha_2)$. *An interpreted transition system m is a model of a TBox* $T = \langle \Gamma, A \rangle$ *if m is a model of each inclusion assertion in* Γ *and A.*

Boolean operators $\sqcup$ on state types and $\exists$ are defined as usual. We will discuss the complexity of this logic in Section 5.4.

Remark 5.1. *(Derived constructs) The correspondence between description logic and dynamic logic is well-known ([Baader et al., 2002]). In fact, the language presented in Definitions 5.1 and 5.3 is a notational variant of the language of Dynamic Logic ([D. Harel amd Kozen and Tiuryn, 1984]) without the iteration operator on transition types. As a consequence, some key constructs are still definable in $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$. In particular we will make use of the following definition of the if-then-else transition type:*

$$\mathtt{if}\ \gamma\ \mathtt{then}\ \alpha_1 \mathtt{else}\ \alpha_2 = (id(\gamma) \circ \alpha_1) \sqcup (id(\neg\gamma) \circ \alpha_2).$$

5.1.3 Institutional TBoxes

We have upheld that institutions "impose" new system descriptions which are formulated in terms of sets of norms. The step toward a formal grounding of this view of institutions is now short: norms can be thought of as terminological axioms, and institutions as sets of terminological axioms, i.e., terminological boxes.

An institution can be specified as a terminological box $\mathfrak{Ins} = \langle \Gamma_{ins}, A_{ins} \rangle$, where each inclusion statement in Γ_{ins} and A_{ins} models a norm of the institution. Obviously, not every TBox can be considered to be an institution specification. In particular, an institution specification $\mathfrak{Ins}$ must have some precise linguistic relationship with the 'brute' descriptions on the top of which the institution is supposed to be specified. We denote by $\mathcal{L}_{ins}$ language built from a non-logical alphabet containing only institutional state and transition types, and by $\mathcal{L}_{brute}$ the language built from a non-logical alphabet containing those types taken to talk about, instead, 'brute' states and transitions. Languages $\mathcal{L}_{ins}$ and $\mathcal{L}_{brute}$ are disjoint. The language on which an institution is specified should always consist of these two parts.

Definition 5.4. *(Institutions as TBoxes)*
A TBox $\mathfrak{Ins} = \langle \Gamma_{ins}, A_{ins} \rangle$ is an institution specification if:

1. *The non-logical alphabet on which $\mathfrak{Ins}$ is specified contains elements of both $\mathcal{L}_{ins}$ and $\mathcal{L}_{brute}$.*

2. *There exist sets of terminological axioms $\Gamma_{bridge} \subseteq \Gamma_{ins}$ and $A_{bridge} \subseteq A_{ins}$ such that either the left-hand side of these axioms is always a description expressed in $\mathcal{L}_{brute}$ and the right-hand side a description expressed in $\mathcal{L}_{ins}$, or those axioms are definitions. In symbols: if $\gamma_1 \sqsubseteq \gamma_2 \in \Gamma_{bridge}$ then either $\gamma_1 \in \mathcal{L}_{brute}$ and $\gamma_2 \in \mathcal{L}_{ins}$ or it is the case that also $\gamma_2 \sqsubseteq \gamma_1 \in \Gamma_{bridge}$. The clause for A_{bridge} is analogous.*

3. *The remaining sets of terminological axioms $\Gamma_{ins} \backslash \Gamma_{bridge}$ and $A_{ins} \backslash A_{bridge}$ are all expressed in $\mathcal{L}_{ins}$.*

The definition states that an institution specification needs to be expressed on a language including institutional as well as brute terms (1); that a part of the specification concerns a description of mere institutional terms (3); and that there needs to be a part of the specification which connects institutional terms to brute ones (2).

From a design perspective language $\mathcal{L}_{brute}$ has to be thought of as the language on which a designer would specify a system instantiating a given institution[1]. Definition 5.4 shows that for such a design task it is needed to formally specify an explicit bridge between the concepts used in the description of the actual system and the institutional 'abstract' concepts.

Example 5.1. *(A simple institution specification) Suppose language $\mathcal{L}_{brute}$ to be obtained from the set of agents Agents and a set of transition type forms containing at least the subset $\{SEND(msg9), SEND(msg11)\}$. Suppose then language $\mathcal{L}_{ins}$ to be obtained from the same set of agents and from a set of transition type forms which includes at least $\{REQUEST\_ACCESS, DENY\_ACCESS\}$, and a set of state type forms containing at least $\{authorized\}$. An institution $\mathfrak{Ins} = \langle \Gamma_{ins}, A_{ins} \rangle$ specifying how access (for instance to some web service) can be requested by an agent i and denied by an agent j would include the following terminological axioms for all $i, j \in Agents$:*

$$SEND(msg9, i, j) \sqsubseteq REQUEST\_ACCESS(i, j) \tag{5.1}$$

$$SEND(msg11, j, i) \sqsubseteq DENY\_ACCESS(j, i) \tag{5.2}$$

$$\top \sqsubseteq \forall DENY\_ACCESS(j, i).\neg authorized(i) \tag{5.3}$$

where Formulae 5.1 and 5.2 are transition type bridge axioms in A_{bridge}, and Formula 5.3 is a state type inclusion axiom on $\mathcal{L}_{ins}$ specifying the effects of the transitions of form DENY_ACCESS.

Notice that the TBox specified in Example 5.1 is preceded by the locution "for all $i, j \in Agents$". In fact, in the rest of the chapter, terminological axioms have to be read as schemata determining a finite number of subsumption expressions depending on the cardinality of the set *Agents* considered.

5.1.4 TBoxes and contextual terminologies

By thinking of an institution specification as the specification of a TBox, we directly establish a link to what investigated in Chapter 2, i.e., contextual terminologies. A TBox is always expressed on an alphabet —we have seen that this can be institutional or brute (Definition 5.4)— and it always defines a set of models, i.e., the set of models validating its terminological axioms. Once the set of system's states that need an institutional description is selected, a TBox $\mathfrak{Ins}$ defines a context in the sense of Chapter 2, a set of possible institutional interpretations of those states according to the norms in $\mathfrak{Ins}$.

[1]To make a concrete example, the AMELI middleware [Esteva et al., 2004] can be viewed as a specification tool at a $\mathcal{L}_{brute}$ level. In fact, AMELI does not support the formal specification of the regulations —the $\mathcal{L}_{ins}$-level specification— which the to-be-designed e-Institution is supposed to implement. This is a limit of the AMELI approach which has been noticed in the literature on e-Institutions (see, for instance, [Vázquez-Salceda, 2004]).

In Chapter 4 it has been analyzed, in modal logic, how classificatory statements holding in a context (classificatory counts-as) relate to classificatory statements defining the context with respect to which they are also considered to hold (constitutive counts-as). The best way to look at the present chapter is to think of it as making, in DL, the same step from what holds in a context to what defines the context: from contextual terminologies, i.e., sets of terminological axioms holding with respect to a context (Chapter 2), to TBoxes, i.e., sets of terminological axioms defining a context. The axioms contained in a TBox $\mathfrak{Ins} = \langle \Gamma_{ins}, A_{ins} \rangle$ are all *constitutive* statements in the sense made precise in Section 4.3.2. As such, they can all be thought of as counts-as statements of a constitutive kind, provided of course that they do not express logical truths (see Section 4.3).

Notice, finally, that a TBox $\mathfrak{Ins}$ defines a context which is a concreter version —in the sense made precise by formula 2.3 of Chapter 2— of the context defined by the same TBox after removing the bridge axioms, that is to say, the 'abstract' TBox $\left\langle \Gamma_{ins} \backslash \Gamma_{bridge}, A_{ins} \backslash A_{bridge} \right\rangle$. To get back to the Examples 2.1 and 2.2 discussed in Chapter 2, such an 'abstract' TBox would define the context c_{Reg} of the regional regulations on public park access, while TBoxes $\mathfrak{Ins}$ would define the contexts of the municipal regulations (c_{M1}, c_{M2} and c_{M3}) interpreting the regional one in concrete terms. In this chapter only such concrete regulations are treated as institutions.

5.2 Explaining Institutions

This section illustrates Definition 5.4, and shows its explanatory power in accounting for some essential aspects of institutions.

5.2.1 From abstract to concrete norms

Abstract norms in an institution specification $\mathfrak{Ins}$ are subsumptions between descriptions stated in $\mathcal{L}_{ins}$, that is, elements of $\Gamma_{ins} \backslash \Gamma_{bridge}$ or $A_{ins} \backslash A_{bridge}$ (see Definition 5.4). Concrete norms are, in contrast, subsumtpions between descriptions stated in $\mathcal{L}_{brute}$. The connections between the two is provided by the subsumptions of $\mathfrak{Ins}$ to be found in Γ_{bridge} or A_{bridge}[2]. An example follows which clarifies the interaction between abstract and concrete norms within institutions.

Example 5.2. *(From abstract to concrete norms) Consider an institution supposed to regulate access to a set of public web services. It may contain the following norm: "it is forbidden to discriminate access on the basis of citizenship". Suppose now a system has to be built which complies with this norm. The first question is: what does it mean, in concrete, "to discriminate on the basis of citizenship"? The system designer should make some concrete choices for interpreting the norm and these choices should be kept track of in*

[2]It is worth stressing that more than just two abstractness levels could in principle be represented, depending on how many sublanguages are considered. As shown in Chapter 2, abstractness and concreteness are, in the first instance, attributes of contexts w.r.t. other contexts. In general, a norm is abstract/concrete if it pertains to an abstract context and, respectively, to a concrete one. In this chapter, by considering just two sublanguages ($\mathcal{L}_{ins}$ and $\mathcal{L}_{brute}$), we will work with only two levels.

order to explicitly link the abstract *norm to its* concrete *interpretation. The problem can be represented as follows. The abstract norm is formalized as described in Section 2.5.1: the statement "it is forbidden to discriminate on the basis of citizenship" amounts to the statement "after every execution of a transition of type $DISCR(i, j)$ the system always ends up in a violation state". Together with the norm also some intuitive background knowledge about the discrimination action needs to be formalized. Here, as well as in the rest of the examples in the chapter, we provide just that part of the formalization which is strictly functional to show how the formalism works in practice. Formulae 5.5 and 5.6 express two effect laws: if the requester j is Dutch, then after all executions of transitions of type $DISCR(i, j)$ j is accepted by i. If it is not, then all the executions of the transitions of the same type have as effect that it is not accepted.*

$$\forall DISCR(i, j).viol \equiv \top \tag{5.4}$$

$$dutch(j) \sqsubseteq \forall DISCR(i, j).accepted(j) \tag{5.5}$$

$$\neg dutch(j) \sqsubseteq \forall DISCR(i, j).\neg accepted(j) \tag{5.6}$$

The rest of the axioms concern the translation of the abstract type $DISCR(i, j)$ to concrete transition types. Formula 5.7 refines it by making explicit that a precise if-then-else procedure counts as *a discriminatory act of agent i. Formulae 5.8 and 5.9 specify which messages of i to j count as acceptance and rejection. If the designer uses transition types $SEND(msg33, i, j)$ and $SEND(msg38, i, j)$ for the concrete system specification, then Formulae 5.8 and 5.9 are bridge axioms connecting notions belonging to the institutional alphabet (to accept, and to reject) to concrete ones (to send specific messages). Finally, Formulae 5.10 and 5.11 state two intuitive effect laws concerning $ACCEPT(i, j)$ and $REJECT(i, j)$ by tuning the labeling of the states reachable via those transition types.*

$$\begin{aligned} &\texttt{if}\ dutch(j)\texttt{then}\ ACCEPT(i, j)\\ &\texttt{else}\ REJECT(i, j) \sqsubseteq DISCR(i, j) \end{aligned} \tag{5.7}$$

$$SEND(msg33, i, j) \sqsubseteq ACCEPT(i, j) \tag{5.8}$$

$$SEND(msg38, i, j) \sqsubseteq REJECT(i, j) \tag{5.9}$$

$$\forall ACCEPT(i, j).accepted(j) \equiv \top \tag{5.10}$$

$$\forall REJECT(i, j).\neg accepted(j) \equiv \top \tag{5.11}$$

It is easy to see, on the grounds of the semantics exposed in Definition 5.3, that the following concrete inclusion statement holds w.r.t. the specified institution:

$$\begin{aligned} &\texttt{if}\ dutch(j)\ \texttt{then}\ SEND(msg33, i, j)\\ &\texttt{else}\ SEND(msg38, i, j) \sqsubseteq DISCR(i, j) \end{aligned} \tag{5.12}$$

Notice also that this translation is aligned with the constraints stated in Formulae 5.5 and 5.6.

This scenario exemplifies a pervasive feature of human institutions which, as extensively argued in [Grossi et al., 2006b], should be incorporated by electronic ones. Current formal approaches to institutions, such as ISLANDER [Esteva et al., 2002], do not allow for the formal specification of explicit translations of abstract norms into concrete ones, and focus only on norms that can be specified at the concrete system specification level. What Example 5.2 shows is that the problem of the abstractness of norms in institutions can be formally addressed and can be given a precise formal semantics.

5.2.2 Non-arbitrariness of institutional specifications

The scenario depicted in Example 5.2 suggests that, just by modifying an appropriate set of terminological axioms, it is possible for the designer to obtain a different institution by just modifying the sets of bridge axioms without touching the terminological axioms expressed only in the institutional language $\mathcal{L}_{ins}$. In fact, it is the case that a same set of abstract norms can be translated to different and even incompatible sets of concrete norms. Is such translation completely arbitrary? The answer is no.

Example 5.3. *(Acceptable and unacceptable translations of abstract norms) Reconsider again the scenario sketched in Example 5.2. The transition type $DISCR(i, j)$ has been translated to a complex procedure composed by concrete transition types. Would any translation do? Consider an alternative institution specification $\mathfrak{Ins}'$ containing Formulae 5.4-5.6 and 5.10, and the following translation rule:*

$$ACCEPT(i, j) \sqsubseteq DISCR(i, j) \tag{5.13}$$

Would this formula be an acceptable translation for the abstract norm expressed in Formula 5.4? The axiom states that transitions where i accepts j count as transitions of type $DISCR(i, j)$. In fact, this is not intuitive because the abstract transition type $DISCR(i, j)$ obeys some intuitive conceptual constraints (Formulae 5.5 and 5.6) that all its translations should also obey. In fact, the following inclusions hold in $\mathfrak{Ins}'$ as consequences of Formula 5.13:

$$dutch(j) \sqsubseteq \forall ACCEPT(i, j).accepted(j) \tag{5.14}$$

$$\neg dutch(j) \sqsubseteq \forall ACCEPT(i, j).\neg accepted(j) \tag{5.15}$$

These properties of the transition type $ACCEPT(i, j)$ conflict with what follows from Formula 5.10:

$$dutch(j) \sqsubseteq \forall ACCEPT(i, j).accepted(j) \tag{5.16}$$

$$\neg dutch(j) \sqsubseteq \forall ACCEPT(i, j).accepted(j) \tag{5.17}$$

Transitions of type $ACCEPT(i, j)$ always bring about states of type $accepted(j)$. Now, from Formula 5.15 and 5.17 it follows:

$$\neg dutch(j) \sqsubseteq \forall DISCR(i, j).\bot \tag{5.18}$$

which would be quite odd.

The moral of the story is that when abstract transition or state types are translated, via appropriate inclusion axioms, to concrete ones, these concrete types should be compatible with the inheritance of the properties of the abstract types. This compatibility marks the boundaries within which translations are possible, and sets therefore precise logical limitations to the choice of the translation which cannot be fully arbitrary. To say it with Searle:

> "the selection of the X term [in the X counts as Y rule] is *more or less* arbitrary" ([Searle, 1995], p.49).

The choice of the translation is only "more or less arbitrary" and not merely arbitrary in virtue of the properties of the concrete term which should be compatible with the properties of the abstract one.

It is important to stress that this very same issue was already addressed, although from a slightly different perspective, in Section 2.4.3 of Chapter 2 where the notion of open-texture has been formally analyzed. Translations have boundaries because the to-be-translated terms are open-textured and not arbitrary.

5.2.3 Institutional modules and roles

Viewing institutions as the impositions of institutional descriptions on systems' states and transitions allows for analyzing the normative system perspective itself (i.e., institutions are sets of norms) at a finer granularity. We have seen that the terminological axioms specifying an institution concern complex descriptions of new institutional notions. Some of the institutional state types occurring in the institution specification play a key role in structuring the specification of the institution itself. The paradigmatic example in this sense are facts such as "agent i enacts role r" which will be denoted by state types $rea(i, r)$ ([Dignum, 2003]). By stating how an agent can enact and 'de-act' a role r, and what normative consequences follow from the enactment of r, an institution describes expected forms of agents' behavior while at the same time abstracting from the concrete agents taking part of the system.

The sets of norms specifying an institution can be clustered on the grounds of the *rea* state types. For each relevant institutional state type (e.g., $rea(i, r)$), the terminological axioms which define an institution, i.e., its norms, can be clustered in (possibly overlapping) sets of three different types: the axioms specifying how states of that institutional type can be reached (e.g., how an agent i can enact the role r); how states of that type can be left (e.g., how an agent i can 'de-act' the a role r); and what kind of institutional consequences do those states bear (e.g., what rights and power does agent i acquire by enacting role r). Borrowing the terminology from work in legal and institutional theory ([Ruiter, 1997; Searle, 1995; Sartor, 2006]), these clusters of norms can be called, respectively, institutive, terminative and status modules.

Remark 5.2. *(Refraining from executing transition types) In what follows we will need to represent a form of negation of atomic transition types crudely corresponding to some notion of refraining. It is well-known that this is a hard issue to solve in dynamic logic-like* *formalisms like ours (see [Broersen, 2003]) and the readily available solution of using the negation $\neg$ of transition types is obviously too strong, since such negation is interpreted as the complement w.r.t. the whole state space $S \times S$. We choose for a low-profile solution, which suits our needs without introducing heavy logical machinery that would not be used in our analysis. The non-logical alphabet of our language (see Definition 5.2) needs to be extended as follows: for every atomic transition type a, non_a is also an atomic transition type. Obviously, if $a \in \mathcal{L}_{ins}$ then also $non\_a \in \mathcal{L}_{ins}$ and, respectively, if $a \in \mathcal{L}_{brute}$ then also $non\_a \in \mathcal{L}_{brute}$. In addition, any occurrence of an atomic transition type non_a in a TBox needs to be accompanied by the following role inclusion axiom: $non\_a \sqsubseteq \neg a$*[3]*. More elegant but complex solutions to the problem can be found in [Meyer, 1988] and in the comprehensive survey [Broersen, 2003].*

Status modules

We call status modules those sets of terminological axioms which specify the institutional consequences of the occurrence of a given institutional state-of-affairs, for instance, the fact that agent i enacts role r.

Example 5.4. *(A status module for roles) Enacting a role within an institution bears some institutional consequences that are grouped under the notion of* status*: by playing a role an agent acquires a specific status. Some of these consequences are deontic and concern the obligations, rights, permissions under which the agent puts itself once it enacts the role. An example which pertains to the normative description of the status of both a "buyer" and a "seller" roles is the following:*

$$rea(i, buyer) \sqcap rea(j, seller) \sqcap bought(i, j, b) \sqsubseteq \forall non\_PAY(i, j, b).viol(i) \qquad (5.19)$$

If agent i enacts the buyer role and j the seller role and i wins bid b, then if i does not perform a transition of type $PAY(i, j, b)$, i.e., does not pay to j the price corresponding to bid b, then the system ends up in a state that the institution classifies as a violation state with i being the violator.

Of particular interest are those consequences that attribute powers to agents enacting specific roles. These powers concern the bringing about of institutional states of affairs, such as the fact that an object has been bought.

$$rea(i, buyer) \sqcap rea(j, seller) \sqsubseteq \forall BID(i, j, b) \circ BID\_OK(j, i, b).bought(i, j, b) \qquad (5.20)$$

$$SEND(i, j, msg49) \sqsubseteq BID(i, j, b) \qquad (5.21)$$

$$SEND(j, i, msg50) \sqsubseteq BID\_OK(j, i, b) \qquad (5.22)$$

If agent i enacts the buyer role and j the seller role, every time agent i bids b to j and j accepts the bid, then this action results in an institutional state testifying that the corresponding bid has been placed by i (Formula 5.20). Formulae 5.21 and 5.22 state how the bidding action can be executed by sending a specific message to j ($SEND(i, j, msg49)$) and, respectively, how a bid b from i can be accepted by j ($SEND(j, i, msg50)$).

[3] Notice that these kind of axioms cannot be bridge axioms.

The example shows how roles can be formally related to deontic notions and institutionalized power. Formula 5.19 formalizes at the same time an obligation pertaining to the role 'buyer', and a right of type claim ([Hohfeld, 1911]) whose bearer is the role 'seller'.

Formulae 5.20-5.22 show how the notion of institutional power ([Jones and Sergot, 1996]) can be formalized in the language introduced. Institutional power is modeled by means of two types of rules: one specifying the institutional effects of an institutional action (Formula 5.20), and one translating institutional transition types in brute ones (Formulae 5.21 and 5.22). Formulae 5.21 and 5.22, as we have already seen in Example 5.4, can be properly thought of as counts-as statements. This is an essential aspect since it shows how counts-as statements ground the specification of institutional power in normative systems. Such systems of rules *empower* the agents enacting some relevant role by establishing a connection between the brute actions of the agents and some institutional effect.

These considerations align our perspective with the main theses of the seminal work on institutionalized power presented in [Jones and Sergot, 1996] and already thoroughly discussed in Chapter 4. In that work counts-as was primarily studied in order to provide grounds for a formal account of the notion of power within institutions. From the technical point of view, our perspective is also very close to what maintained in [Sergot, 2004], where counts-as statements are studied exactly as transition types subsumptions in interpreted transition systems. We will come back to this work in Section 5.8.

Whether the agents are actually able to execute the required 'brute' actions is a different issue, since agent i can be in some states (or even all states) unable to effectuate a SEND$(i, j, msg49)$ transition. This is the case also in human societies: priests are empowered to give rise to marriages but, if a priest is not in state of performing the required speech acts he is actually unable to marry anybody. There is a difference between "being entitled" to make a bid and "being in state of" making a bid ([Castelfranchi, 2003]). In other words, Formulae 5.20 and 5.21 express only that agents playing the buyer role are entitled to make bids. The actual possibility of performing the required 'brute' actions is not an institutional issue, but rather an issue concerning the implementation of an institution in a concrete system. We address this issue extensively in Section 5.5[4].

Institutive modules

We call institutive modules those sets of terminological axioms of an institution specification describing how states with certain institutional properties can be reached, for instance, how an agent i can reach a state in which it enacts role r. They can be seen as procedures that the institution define in order for the agents to bring about institutional states of affairs.

Example 5.5. *(An institutive module for roles) The fact that an agent i enacts a role r (rea(i, r)) is the effect of a corresponding enactment action ENACT(i, r) performed under*

[4]See in particular Example 5.7 and Definition 5.6

certain circumstances (Formula 5.23), namely that the agent does not already enact the role, and that the agent satisfies given conditions (cond(i, r)), which might for instance pertain the computational capabilities required for an agent to play the chosen role, or its capability to interact with some specific system's infrastructures. Formula 5.24 specifies the procedure counting as an action of type ENACT(i, r). Such a procedure is performed through a system mediator s, which notifies to i that it has been registered as enacting role r after sending the necessary piece of data d (SEND(i, s, d)), e.g., a valid credit card number.

$$\neg rea(i, r) \sqcap cond(i, r) \sqsubseteq \forall ENACT(i, r).rea(i, r) \tag{5.23}$$

$$SEND(i, s, d) \circ NOTIFY(s, i) \sqsubseteq \forall ENACT(i, r) \tag{5.24}$$

Terminative modules

Analogously, we call terminative modules those sets of terminological axioms stating how a state with certain institutional properties can be left. Rules of this kind state for instance how an agent can stop enacting a certain role. Thus, they can be thought of as procedures that the institution defines in order for the agent to see to it that certain institutional states stop holding.

Example 5.6. *(A terminative module for roles) Terminative modules for roles specify, for instance, how a transition type DEACT(i, r) can be executed which has as consequence the reaching of a state of type $\neg$rea(i, r):*

$$rea(i, r) \sqsubseteq \forall DEACT(i, r).\neg rea(i, r) \tag{5.25}$$

$$SEND(i, s, msg9) \sqsubseteq \forall DEACT(i, r) \tag{5.26}$$

That is to say, i de-acting a role r always leads to a state where i does not enact role r; and i sending message No.9 to a specific interface infrastructure s count as i de-acting role r. Notice that the role-de-actment activity might be subjected to precise norms. For instance, an institution might require that a certain role cannot be de-acted before certain objectives have been achieved. Such norms would then all be part of the status module of the role at issue.

Examples 5.4-5.6 have shown how roles can be formalized in our framework. Roles are sets of terminological axioms concerning state types of the sort *rea*(i, r). It is worth noticing that this modeling strategy is aligned with work on social theory addressing the concept of role such as [Pörn, 1977], and it offers the possibility to finally provide role specification, as it occurs in a number of methodologies for MASs such as GAIA ([Zambonelli et al., 2003]) or OPERA ([Dignum, 2003]), with a formal semantics in terms of interpreted transition systems.

5.3 Intermezzo: Caligula's Horse

This section is inspired by some considerations advanced in [Azzoni, 2003]. It tries to provide an enjoyable recapitulation of some of the aspects of institutions discussed in the previous sections. We take as imput this quote about the Roman emperor Caligula.

> "He used to send his soldiers on the day before the games and order silence in the neighborhood, to prevent the horse Incitatus from being disturbed. Besides a stall of marble, a manger of ivory, purple blankets and a collar of precious stones, he even gave this horse a house, a troop of slaves and furniture, for the more elegant entertainment of the guests invited in his name; and *it is also said that he planned to make him consul*" ([Svetonius, 110], Caligula, LV, 8).

Caligula passed to history for his foolishness. Among all the oddities historians attributed him one of the most famous concerns his horse Incitatus, which he pronounced, or planned to pronounce, consul. We will focus on the emphasis in the quoted excerpt: "it is said that he planned to make him consul".

At the time of the empire, a long republican tradition concerning the role 'consul' had established some precise constraints about how it was possible for a Roman to enact that role. If a Roman was eligible for that position and was not already a consul, then by following a precise procedure he could become consul:

$$\neg rea(cons, j) \sqcap \text{eligible}(cons, j) \sqsubseteq \forall \text{ENACT}(j, cons).rea(cons, j).$$

Needless to say, being eligible meant at least to be able to communicate with the fellow Romans:

$$\text{eligible}(cons, j) \sqsubseteq \text{speak\_latin}(j).$$

If those requirements were met, then the enactment of the role 'consul' consisted in winning a specific election:

$$id(rea(emp, i) \sqcap \neg rea(cons, j) \sqcap \text{eligible}(cons, j)) \circ \text{ELECT}(j, cons) \equiv \text{ENACT}(j, cons).$$

At the time of the republic, it was therefore the case that:

$$\exists \text{ENACT}(i, cons).\top \sqsubseteq \text{speak\_latin}(i).$$

that is, it was necessarily the case that if somebody could engage in the standard procedure for becoming consul then he could speak Latin.

With the introduction of the imperial institutions things chanced a bit. Emperors used to enjoy a certain status which guaranteed unlimited institutional power. It seems it was enough to pronounce somebody consul for him to enact the role 'consul'. The procedure for enacting the role 'consul' was thereby broadened:

$$\begin{aligned}&(id(\neg rea(cons, j) \sqcap \text{eligible}(cons, j)) \circ \text{ELECT}(j, cons)\\ &\cup id(rea(emp, i)) \circ \text{MAKE}(i, j, cons)) \equiv \text{ENACT}(j, cons).\end{aligned}$$

It followed that although something was not eligible to become a consul in the republican sense, he could still become consul by the emperor's overruling power:

$$\neg \text{speak\_latin}(j) \sqcap rea(emp, i) \sqcap \exists \text{MAKE}(i, j, cons).\top \sqsubseteq \exists \text{MAKE}(i, j, cons).rea(cons, j).$$

This state-type subsumption rule covers exactly the case of Caligula's horse which, although not in the condition to engage in the procedure for becoming a consul, it became consul nevertheless.

Remark 5.3. *(Caligula's horse and the range of concepts in contexts) The scenario obtains a natural representation in terms of contextual terminologies (Chapter 2). Once we denote with Rep the context of republican institutions and with Emp the context of the imperial ones, the shift from the former to the latter can be appreciated as the shift from:*

$$Incitatus \notin \mathfrak{Range}_{\mathbb{M}}(\exists ENACT(i, cons).\top, Rep)$$

to

$$Incitatus \in \mathfrak{Range}_{\mathbb{M}}(\exists ENACT(i, cons).\top, Emp)$$

that is, from Incitatus lying in, to Incitatus lying out of the range of $\exists ENACT(i, cons).\top$.

The example highlights the aspect of the constitution of institutional facts, which we touched upon in Section 5.2.2, concerning the boundaries delimiting the possible translations of abstract institutional terms. If we consider the abilities of agents j to be invariant in the model, such as for instance $(\neg)$speak_latin(j) then Incitatus found itself in the awkward situation of being a 'consul' without being able to exercise the powers related to that status, such as, for instance, propose new laws:

$$rea(cons, j) \sqcap \neg \text{speak\_latin}(j) \sqsubseteq \neg \exists \text{PROPOSE\_LAW}.\top$$

What imperial institutions allowed to, was to make a consul of something which could not reasonably count as a consul. Caligula exploited this possibility concretely showing the logical difficulties of unlimited institutional power, and he has been remembered for this.

5.4 Tractable specifications of institutions

In the previous sections we fully deployed the expressivity of the language introduced in Section 5.1.2 and used its semantics to provide a formal understanding of many essential aspects of institutions in terms of transition systems. This section spends a few words about the viability of performing reasoning in the logic presented.

5.4.1 Reasoning in TBoxes: a sketch

The standard TBox reasoning tasks in DL are essentially two ([Baader et al., 2002]): satisfiability and subsumption.

The satisfiability problem amounts to check whether a state description γ is satisfiable w.r.t. a given TBox $\mathfrak{T}$, i.e., to check if there exists a model $m = \langle S, \mathcal{I} \rangle$ of $\mathfrak{T}$ such that $\emptyset \subset \mathcal{I}(\gamma)$. The subsumption problem amounts instead to check whether a given subsumption relation $\gamma_1 \sqsubseteq \gamma_2$ is modeled by all models of a given a TBox $\mathfrak{T}$, i.e., if it logically follows from $\mathfrak{T}$.

If negation and intersection of arbitrary state types are enabled in the language —and this is the case for $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$— then the subsumption problem can be reduced to the satisfiability one: $\gamma_1 \sqsubseteq \gamma_2$ iff $\gamma_1 \sqcap \neg\gamma_2 \sqsubseteq \bot$ ([Baader et al., 2002]).

5.4.2 Reasoning in fragments of $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$

The satisfiability problem for logic $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$ is undecidable since transition-type inclusion axioms correspond to a version of what in DL are known as "role-value maps". Logics extending $\mathcal{ALC}$ with role-value maps are known to be undecidable ([Baader et al., 2002]).

Tractable (i.e., polynomial time decidable) fragments of $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$ can however be isolated which still exhibit some key expressive features. In the following we mention one in particular, which we call logic $\mathcal{ELH}^{(\circ)}$. It is obtained from description logic $\mathcal{EL}$, which contains only state types intersection $\sqcap$, existential restriction $\exists$ and $\top$[5], but extended with the $\bot$ state type and with transition type inclusion axioms of a complex form: $a_1 \circ \ldots \circ a_n \sqsubseteq a$ (with n finite). Logic $\mathcal{ELH}^{(\circ)}$ is also a fragment of the well investigated description logic $\mathcal{EL}^{++}$ whose satisfiability and subsumption problems have been shown, in [Baader et al., 2005], to be decidable in polynomial time.

Despite the very limited expressivity of this fragment, some rudimentary institutional specifications can still be successfully represented. Specifically, institutive and terminative modules can be represented which contain transition types inclusion axioms expressing, for instance, that a given sequential procedure counts as a specific institutional action: "sending message n.13 by i to j followed by the sending of message n.31 by j to i counts as a registration action of i":

$$\mathrm{SEND}(msg13, i, j) \circ \mathrm{SEND}(msg33, j, i) \sqsubseteq \mathrm{REG}(i) \tag{5.27}$$

The only serious limit for representing institutive and terminative modules in $\mathcal{ELH}^{(\circ)}$ consists in the impossibility of expressing effect laws since the $\forall$ operator is not available. However, we can always express that there exists some transition of a certain type that, if executed, leads to a given state: "if agent i has the necessary capabilities for enacting role r then there always exists at least one transaction of type $\mathrm{ENACT}(i, r)$ leading to a state where agent i actually enacts role r:

$$\mathrm{cond}(i, r) \sqsubseteq \exists \mathrm{ENACT}(i, r).\mathrm{rea}(i, r) \tag{5.28}$$

Notice that this formula can be viewed as an alternative representation of what is stated in Formulae 5.23 and 5.24. Notice also that it implies the executability law: $\mathrm{cond}(i, r) \sqsubseteq \exists \mathrm{ENACT}(i, r).\top$.

Status modules can also be represented. However, to represent deontic notions we need to expand the language $\mathcal{L}_{ins}$ with a set of state types $\{\mathrm{legal}(i)\}_{0 \leq i \leq n}$ whose

[5] Notice therefore that $\mathcal{EL}$ is a seriously restricted fragment of $\mathcal{ALC}$ since it does not contain the negation operator for state types (operators $\sqcup$ and $\forall$ remains thus undefinable).

intuitive meaning is to denote legal states as opposed to states of type viol(i). In $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$ this role is played by $\neg$viol(i) descriptions, but since $\mathcal{ELH}^{(\circ)}$ does not allow negation the expansion of $\mathcal{L}_{ins}$ is needed, together with the finite set of n terminological axiom stating that legal and violation types are disjoint: viol(i) $\sqcap$ legal(i) $\sqsubseteq \bot$. The expressible deontic notions are thus reduced to two essential notions: "it is possible (respectively, impossible) to reach a violation state by performing a transition of a certain type" (e.g., Formulae 5.29 and 5.32), and "it is possible (respectively, impossible) to reach a legal state by performing a transition of a certain type" (e.g., Formulae 5.30 and 5.31).

$$rea(i, \text{buyer}) \sqcap rea(j, \text{seller}) \sqcap \text{bought}(i, j, b) \sqsubseteq \exists \text{non\_PAY}(i, j, b).\text{viol}(i) \tag{5.29}$$

$$rea(i, \text{buyer}) \sqcap rea(j, \text{seller}) \sqcap \text{bought}(i, j, b) \sqcap \exists \text{non\_PAY}(i, j, b).\text{legal}(i) \sqsubseteq \bot \tag{5.30}$$

$$rea(i, \text{buyer}) \sqcap rea(j, \text{seller}) \sqcap \text{bought}(i, j, b) \sqsubseteq \exists \text{PAY}(i, j, b).\text{legal}(i) \tag{5.31}$$

$$rea(i, \text{buyer}) \sqcap rea(j, \text{seller}) \sqcap \text{bought}(i, j, b) \sqcap \exists \text{PAY}(i, j, b).\text{viol}(i) \sqsubseteq \bot \tag{5.32}$$

These formulae approximate what is stated in Formula 5.19 by saying something less and something more at the same time. Formulae 5.29 and 5.30 state the institutional consequences of not paying after winning a bid: this transition can always end up in a violation state, and there is no state from which a legal state can be reached given those circumstances. Formulae 5.31 and 5.32 state, in similar fashion, that by performing a transition of type PAY(i, j, b) the agent can end up in a legal state, whereas it is impossible for it to end up in an illegal one. Notice that what these formulae state which is not implied by Formula 5.19 is that there is always the possibility to reach either a legal or a violation state. Status modules represented in $\mathcal{ELH}^{(\circ)}$ have thus the limit of stating only what the possible institutional effects are of given transition type.

The moral of the story is that it seems indeed possible to single out some tractable fragments of $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$ which are suitable for performing reasoning about some stretched form of institutional modules[6].

5.5 Infrastructures

In discussing Example 5.4 we observed how "being entitled" to make a bid does not imply "being in the position to" make a bid. In other words, an institution can empower agents by means of appropriate rules but this empowerment can remain dead letter. Similar observations apply also to deontic notions: agents might be allowed to perform certain transactions under some relevant conditions but they might be unable to do so under those same conditions. This is an old, though sporadically addressed issue in the formal analysis of normative notions. In the literature on deontic logic the first reference on this issue is probably [Kanger,

[6] Such fragments could also be used as target logics within theory approximation approaches ([Schaerf and Cadoli, 1995]) by aiming at compiling TBoxes expressed in $\mathcal{ALCH}^{(\sqcup,\circ,\neg,id)}$ into approximations within more tractable logics such as, for instance, $\mathcal{ELH}^{(\circ)}$.

1985], which discusses the problem of the realization of rights distinguishing it from the problem of the compliance of rules of rights:

> "I think we must distinguish realization from compliance" ([Kanger, 1985], p.75).

If those system's states where a given agent is entitled of making a bid are also states where the agent is in the position to make a bid, then the institutional status of that agent is appropriately realized.

From the point of view of agent institutions, these issues are of an infrastructural nature and essentially amount to what the interaction possibilities are of the agents taking part in the institution: what actions are available to an agent and under what conditions. This relates to the design of appropriate *coordination infrastructures* ([Castelfranchi, 2000]) and artifacts ([Omicini et al., 2004]).

5.5.1 Infrastructure specification

In our view, the formal specification of an infrastructure amounts to the formal specification of interaction requirements, that is to say, the specification of which relevant transition types are executable and under what conditions.

The transition types, as well as their executability conditions, which are addressed by an infrastructure specification are 'brute' types. Infrastructure specifications are therefore to be expressed in $\mathcal{L}_{brute}$. In a sense, an infrastructure specification, sets constraints on the possible states that the system can reach via those transitions that the designer considers to be primitive in the whole system specification. In more technical terms, an infrastructure specification sets constraints on the span of the transition system which is supposed to model the institution specification. The formal definition follows.

Definition 5.5. *(Infrastructures as TBoxes)*
An infrastructure $\mathfrak{Inf} = \left\langle \Gamma_{inf}, A_{inf} \right\rangle$ *for institution* $\mathfrak{Ins}$ *is a TBox on* $\mathcal{L}_{brute}$ *such that for all* $a \in \mathcal{L}(A_{bridge})$ *there exist terminological axioms in* Γ_{inf} *of the following forms:* $\gamma \sqsubseteq \exists a.\top$ *(a is executable in* γ *states) or* $\gamma_1 \sqsubseteq \forall a.\gamma_2$ *(a has effects of type* γ_2 *if executed in* γ_1 *states) and their possibly closed versions (i.e., with* $\equiv$ *in place of* $\sqsubseteq$*). Notice that a special case of the second form is:* $\gamma \sqsubseteq \forall a.\bot$ *(a is not executable in* γ *states).*

In other words, an infrastructure specification states conditions under which an atomic brute transition type, which occurs in the brute alphabet of the bridge axioms of $\mathfrak{Ins}$, is executable or not executable and what kind of concrete effects it bears. In other words, it states what can be *in concrete* done and under what conditions.

Example 5.7. *(Infrastructure specifications) Consider the institution specified in Example 5.2. A simple infrastructure* $\mathfrak{Inf}$ *for that institution could contain for instance the following terminological axioms for any pair of different agents i, j and message type msg:*

$$\top \sqsubseteq \exists SEND(msg33, i, j).\top \tag{5.33}$$

The formula states that it is always in the possibilities of agent i to send message No. 33 *to agent j. It follows, on the grounds of Example 5.2, that agent i can always accept agent j.*

$$\top \sqsubseteq \exists ACCEPT(i, j).\top \tag{5.34}$$

Notice that the executability condition is just $\top$.

Consider now the status module specified in Example 5.4, and suppose the following translation rule to be also part of the institution:

$$BNK(i, j, b) \sqcup CC(i, j, b) \equiv PAY(i, j, b) \tag{5.35}$$

which states how the payment can be concretely carried out (via bank transfer or credit card). An infrastructure specification for this module will have to make it possible for agents to comply with the norm specified in Formula 5.19*, by stating the executability of the obligatory action* $PAY(i, j, b)$ *under the relevant conditions:*

$$\forall(SEND(i, j, msg49) \circ SEND(j, i, msg50)).$$
$$protocol\_executed(i, j, msg49, msg50) \equiv \top \tag{5.36}$$
$$protocol\_executed(i, j, msg49, msg50) \sqsubseteq \exists(BNK(i, j, b) \sqcup CC(i, j, b)).\top \tag{5.37}$$

All states reached by $SEND(i, j, msg49) \circ SEND(j, i, msg50)$ *are labelled as states of type* $protocol\_executed(i, j, msg49, msg50)$ *and those states are such that actions* $BNK(i, j, b)$ *or* $CC(i, j, b)$ *are always executable and, therefore,* $PAY(i, j, b)$ *is also executable in those states.*

It is worth noticing that infrastructures consisting of only executability laws, such as the first one in Example 5.7 can be represented in the tractable language $\mathcal{ELH}^{(\circ)}$.

5.5.2 Concrete institution specifications

We call a *concrete institution* specification $\mathfrak{CIns}$ an institution specification $\mathfrak{Ins}$ coupled with an infrastructure specification $\mathfrak{Inf}$.

Definition 5.6. *(Concrete institution)*
A concrete institution obtained by joining the institution $\mathfrak{Ins} = \langle \Gamma_{ins}, A_{ins} \rangle$ *and the infrastructure* $\mathfrak{Inf} = \left\langle \Gamma_{inf}, A_{inf} \right\rangle$ *is a TBox* $\mathfrak{CIns} = \langle \Gamma, A \rangle$ *such that* $\Gamma = \Gamma_{ins} \cup \Gamma_{inf}$ *and* $A = A_{ins} \cup A_{inf}$.

Obviously, different infrastructures can be devised for the same institution giving rise to different concrete institutions which make precise implementation choices explicit. Of particular relevance are the implementation choices concerning abstract norms like the one represented in Formula 5.19. A designer can choose to regiment such norm ([Jones and Sergot, 1993]), i.e., make violation states unreachable, via an appropriate infrastructure.

Example 5.8. *(Regimentation via infrastructure specification) Consider the instition specification in Example 5.4 and the infrastructure specification in Example 5.7. In order to* *specify a regimentation at the infrastructural level it is enough to state that all atomic transition types of agent i except* $BNK(i, j, b)$ *and* $CC(i, j, b)$ *are not executable if it holds that* $protocol\_executed(i, j, msg49, msg50)$*:*

$$protocol\_executed(i, j, msg49, msg50) \sqsubseteq \forall\alpha.\bot \tag{5.38}$$

where α *is the complex action corresponding to the union* $\sqcup$ *of all transition types of agent i different from* $BNK(i, j, b)$ *and* $CC(i, j, b)$*. In other words, in all states reachable via* $SEND(i, j, msg49) \circ SEND(j, i, msg50)$ *the only executable brute actions are* $BNK(i, j, b)$ *or* $CC(i, j, b)$*.*

Regimentation is one of the two possible answers to the problem of norm implementation in institutions. The next section addresses this issue in more details.

5.6 Norm implementation

The purpose of agent institutions is to guarantee the overall behavior of a MAS to exhibit desired properties without compromising agents' autonomy. The viability of this purpose depends on the actual impact that the norms of the institutions have on the agents operating the system, i.e., the problem of norm implementation. It is assumed here that e-Institutions do not have access to the internal states of the agents and hence, that they cannot modify them in order to avoid any incongruence between the agents' behaviour and the institution's norms. There are systems, however, such as KAoS [Bradshaw et al., 1995], where agents' mental states can be accessed by the system and non-compliant goals can be modified or get a lower priority in the deliberation cycle. This is the strongest possible form of regimentation which drastically reduces agents' autonomy[7]. If we rule this possibility out —and this is what is done in this section following, for instance, [Vázquez-Salceda, 2004]— by assuming agents' mental states to be black boxes, the problem arises of how to let those norms have an effective influence on the activities of the agents.

The implementation problem has two sides. There is first of all the interpretation issue, which has already been broadly addressed (see Section 5.2.1), and which concerns the translation of the institutional concepts used in the formulation of the norms in terms of the brute ones used at the system level. However, once the interpretation issue is settled by means of appropriate constitutive norms, the problem is then to make the agents interact in a norm-compliant way. As observed in Section 5.5.2 norms can be trivially implemented at the $\mathfrak{Inf}$ level by making it impossible for agents to end up in violation states (Example 5.8). This section explores another implementation strategy.

[7]It might be interesting to note that this kind of cognitive regimentation reminds of the Burgess's novel "A Clockwork Orange" [Burgess, 1962] where Alex, the main character, after being treated with the "Ludovico Technique", is not able to pursue or even contemplate immoral goals any more. As observed also in Kubrick's film adaptation of the novel, by the words of the prison's chaplain, Alex ceases to be autonomous: "He ceases to be a wrong-doer. He ceases also to be a creature capable of moral choice".

5.6.1 On the notion of enforcement

With enforcement we mean the reaction that an institution specifies to respond to a violation of its norms. Enforcement presupposes, therefore, the possibility of violation. Institutions aim at regulating the behavior of agents through norms, but it is commonplace that norms are useless if the violation of those norms is ignored. Getting back to the Romans again: "*ubi culpa est, ibi poena subesse debet*", which means "where there is a violation, there must be a sanction". In other words, the enforcement of a norm by an institution requires the institution to be in the condition of recognizing the occurrence of violations of that norm in the society and to react upon them. Paradoxical as it might seem, this check-react enforcement procedure is specified by means of more norms. Enforcement is pursued by further regulating the domain, i.e., by adding norms imposing checks and norms specifying reactions to the occurrences of a given violation. To sketch a simplistic picture of the enforcement of tax regulations: tax payment is impossible to be regimented but checks which could detect possible violations are made obligatory. Once the detection takes place, precise reactions are also specified and made obligatory.

We can therefore single out two types of norms involved in the specification of institutions. There is a set of *primary norms* ([Lopez et al., 2006]) which consists of those norms which describe the society's behavior desired by the institution, and there is a set of *enforcement norms* ([Lopez et al., 2006; Grossi et al., 2006a]) consisting of norms regulating the institutional reaction on violations of other norms[8].

Notice that the violations upon which enforcement norms react might be violations of primary norms as well as of other enforcement norms. In fact, via a normatively specified enforcement of the primary norms, the enforcement issue is just lifted up to the set of enforcement norms because, if not regimented, those norms could be violated and be in need of enforcement as well. In principle, this pattern could be endlessly iterated unless there exists a final enforcement level, whose norms are all regimented, or whose violations are not punished (see Figure 5.1).

As a matter of fact, this is precisely how human institutions are structured, where several levels of enforcement regulations may be recognized. Violations on the last level are not considered. For example, the rulings of a supreme court are supposed to be final, even though they might be violating a norm. In human institutions it seems that instead of a full regimentation, the devising of a deep (i.e. structured on more enforcement levels) normative guided reaction offers an efficient norm implementation strategy, granting at the same time a certain institutional flexibility and the room for institutional change and development. It is finally important to notice that, although we have somehow drawn a neat line between the regimentation approach and the enforcement one, an institution will most likely choose for a mixed approach deciding to regiment some norms and to enforce others. We will come back on this issue in the next section.

[8] In [Lopez et al., 2006] *reward norms* are also discussed. These are perfectly analogous to the enforcement ones, except for the fact that they are triggered by a primary norm being complied with.

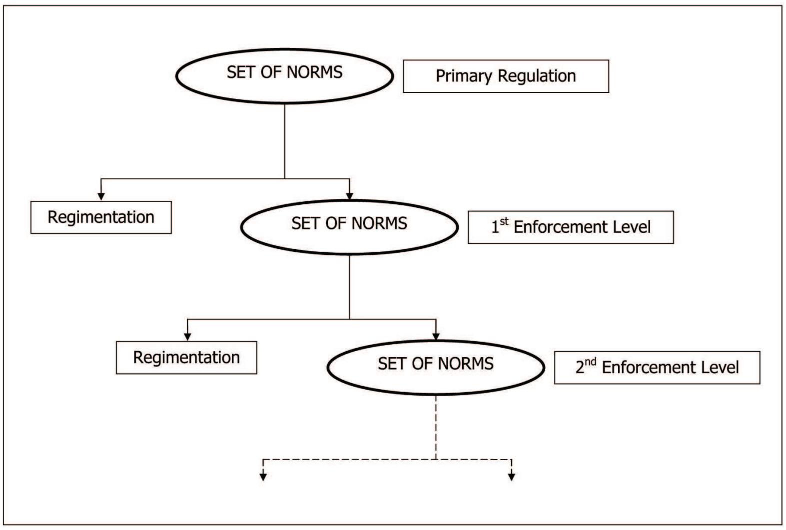

Figure 5.1: Norm implementation between regimentation and enforcement.

5.6.2 On enforcement as complex activity

As we have argued above, the enforcement activity can be divided in two sub-activities: check and reaction. Enforcement norms are not exactly triggered by the violation of a primary norm, bur rather by the perception of the violation by some enforcer.

This is the case in human institutions.

Primary norm "Citizens ought to pay taxes".

Check norm "The tax office ought to perform random checks on citizens' individual income tax declarations".

Reaction norm "If a citizen is found guilty of false individual income declaration then he ought to be punished according to the law".

Check norms play a central role. They specify the way the institution is supposed to perceive the occurrence of violations. Needless to say, this can happen in many different ways. Either directly, via random checks, like in the above example; or via constant monitoring activity, like a referee in a sport match. Or indirectly, allowing agents to denounce the occurrence of a violation and then verifying their claim. This last checking activity is of an intrinsically more complex nature, calling for the establishment of tribunal-like sub-institutions within the main institution. It would be appropriate, in this case, to talk about check sub-institutions rather than simple check norms. For eInstitutions the enforcement activity can partly abstract from the aforementioned check issues, since it can be safely assumed that the occurrence of a violation can in many cases be automatically detected by the system and this information can be stored in an appropriate data log or narrative, as it is called in [Sergot, 2004]. Enforcers would then react directly on the grounds of this information without engaging in violation-perceiving procedures. Furthermore, in eInstitutions enforcers can be considered to be programmed directly by the designer of the e-Institution himself. This means that they can be thought of as fully norm-compliant agents, thereby stopping the need for an enforcement specification regression.

5.6.3 Normative specification of norm enforcement

In terms of our framework, enforcement specification for eInstitutions amounts to the incorporation of a number of terminological axioms specifying what should happen in states of type $viol(i)$ ("agent i is in a violation state").

Like for all other normatively specified activities, enforcement needs to be addressed both at the institutional level itself ($\mathfrak{Ins}$) and at the infrastructural level ($\mathfrak{Inf}$): at the institutional level, we have to specify how an enforcer can execute a punishment, and what kind of punishment it ought to execute in case of a certain violation occurs; at the infrastructural level, we have to specify the executability conditions of the brute transitions by means of which the enforcer can execute the required punishing action.

The framework supports the formal specification of such procedures. First, the institutional language $\mathcal{L}_{ins}$ is expanded with a new set of violation constants $\{viol'(i)\}_{i \in E}$, where E is a set of enforcers. These constants denote that an institutional reaction has not properly taken place. Second, the specification $\mathfrak{Ins}$ is extended with a number of appropriate terminological axioms.

Example 5.9. *(Norm enforcement specification) A simple institutional reaction for the occurrence of a violation of agent i can for instance amount to the de-actment of a role r played by i:*

$$\forall FORCE\_DEACT(e, i, r).\neg rea(i, r) \equiv \top \tag{5.39}$$

$$rea(i, r) \sqcap viol(i) \sqsubseteq \forall \neg FORCE\_DEACT(e, i, r).viol'(e) \tag{5.40}$$

$$SEND(e, s, msg5) \sqsubseteq FORCE\_DEACT(e, i, r) \tag{5.41}$$

That is, if i is in a violation state while enacting role r then enforcer e ought to force i to de-act role r by means of sending message No.5 to a dedicated system component s which is responsible of keeping track of the active rea's in the system.

Notice that Formulae 5.39-5.41 can be in fact viewed as part of a terminative module of role r (Example 5.6).

After incorporating the specification of institutional reactions of the type described in the example, the enforcement problem for $\mathfrak{Ins}$ is shifted to the reaction level.

5.6.4 Different implementations, different societies

The way in which we have conceptualized norm implementation in institutions offers a direct way for showing the differences between different implementation strategies. In particular, what makes the difference is at what level would full regimentation occur. Consider these three cases:

1. The set of primary norms is regimented;
2. The set of first level enforcement norms is regimented;
3. The set of second level enforcement norms is regimented.

As we have seen, in Case 1 violation is impossible. In Case 2 violation is possible but the reaction is automatic. This would result in creating perfect deterrence. Agents would violate the norms only if they think they are better off by violating the regulation. Instead, only in Case 3 it is possible to violate a primary norm without any reaction to take place. This can happen because of a failure by the enforcer in complying with the first level enforcement norms. Such failure would however be automatically punished by the second level enforcers.

Implementation choices generate different societies. These differences are illustrated in the following three toy examples. concerning three possible implementation strategies for an institution that two agents can use in order to play a chess match.

Example 5.10. (Electronic chess) *Let us first consider what happens in an electronic chess match. Players cannot move pieces other than in the way prescribed by the rules of the game, that means that they cannot violate them: the set of actions they can perform within the game is limited and each of these actions is norm compliant. For example, there is no possibility for them to move the rook as if it were a bishop. For these reasons electronic versions of the game of chess constitute a clear example of regimentation of a primary regulation.*

It is instructive to notice that the only institution-based middleware for agents' interaction, which is available at the moment, fall under this category. In AMELI [Esteva et al., 2004], every agent is coupled with an institutional agent, the "governor", which acts as a filter on the agent's activities letting only allowed actions actually take place. Governors are, as such, an excellent example of norm implementation based on the full regimentation of the set of primary norms.

Example 5.11. (Chess with flawless referee agent) *A variation on the previous example would be the use of an automatic agent referee regimenting the first enforcement level norms. Such a referee would always recognize violations and react to them. What would be the* *difference of this implementation of the chess institution with respect to the one described before? In that implementation, the agents could not do anything but play chess, while here they would have a wider range of actions at their disposal such as, for instance, making illegal moves in order to distract the opponent or to signal something. The resulting games would therefore be quite different from the one implemented in the previous example, even though the set of substantial rules (the rules of chess) is the same.*

As we have suggested above (Section 5.6.2) this can be considered to be the simplest form of eInstitutions, where enforcing agents are programmed by the institution designer. To our knowledge, no existing system falls under this category. We are convinced systems of this type could constitute the first step towards complex institution-based agent systems.

Example 5.12. (Chess with referee agent) *Consider now how a chess match in a standard live contest is devised. The two players are not subjected to any regimentation. For example, they have the possibility to move rooks as bishops, thus violating the rules of chess. However, there is a further set of norms stating precisely how to react to a violation. There might be a third party involved, namely a referee, whose task is to detect violations and react to them in specific ways, or to whom suspected violations can be reported by the players. We can then think of a norm, addressed to the referee, stating that the referee ought to check what happens on the chessboard (check norm), to signal an occurring violation and to intervene in the game suspending it and ordering the faulting player to retract its move (reaction norms).*

However, as already noticed, violations can occur also at this level. What should happen if the referee does not detect a move that is not allowed, or does not sanction a player? A further set of norms siding, this time, the first enforcement regulation can provide an answer to these questions. A second enforcement level can be added. This might be a contest committee which is obliged to annul a game vitiated by referee's faults and so on. As already noticed, reactive levels can in principle be added ad infinitum, but they are, of course, de facto limited. For a chess contest, two reactive levels could be practically sufficient to grant a regular chess match, but they are not enough in an absolute sense. What are the new opportunities in this situation? Notice that in this situation players might violate the norms without being noticed (and sanctioned). Therefore the simple fact that a player does not violate the rules might already give him extra credit with his opponent. A notion like trust might suddenly become important in this setting. In general, the possible reasons for making an (illegal) move have again multiplied as well as the interpretation of them. Again, the game is enriched even though the primary rules remain the same.

By means of this example we tried to illustrate how different implementation strategies of the same primary set of norms can actually give rise to radically different institutions and therefore to considerably different systems. The natural question arising is then: what would be, given a society and a set of primary norms, the most sensible implementation strategy? And more crucially, why to allow for violations instead of choosing for a full regimentation?

5.6.5 e-Institutions: to regiment or to enforce?

The implementation of a set of primary norms can be obtained either via regimentation or via the specification of an enforcement activity to be carried out by the institution. Enforcement specification takes place normatively, i.e., via adding more norms to the prior set which, as a consequence, also require implementation. Schematically, suppose S to be the non-empty set of to-be-implemented norms, $Regiment(X)$ to denote the set of norms from X which are regimented, and $Enforce(X)$ to denote the set of norms containing X together with all the norms specifying the enforcement of X ($X \subseteq Enforce(X)$). The implementation of S is the enforcement of the norms in S which are not regimented: $Implement(S) = Enforce(S \setminus Regiment(S))$. In other words, to implement a set of norms amounts to implement the set of unregimented norms together with their enforcement. This definition clearly states that the implementation of a set of norms yields a set of norms, and this is, in a nutshell, the main thesis of the section. In some sense, it is very difficult to get rid of the normative reality. The only possibility is via full regimentation. If $Regiment(S) = S$ then there is no norm left to be implemented. Instead if $Regiment(S) \subset S$ then $\emptyset \subset Implement(S)$, which means that the implementation operation should be iterated on $Implement(S)$.

As we have observed in Sections 5.6.2 and 5.6.4, implementation choices introducing a second enforcement level can be ruled out when considering eInstitutions. There are two options:

1. All primary norms are regimented: $Regiment(S) = S$. In this case no checking and reacting activities are necessary like in Example 5.10.
2. Some (possibly all) norms are left unregimented ($Regiment(S) \subset S$), while what is regimented is just their enforcement like in Example 5.11, that is: $Regiment(Enforce(S \setminus Regiment(S)))$.

The implementation question boils down to: "when is it better to choose 1 over 2 or vice versa?" In general, the preference for 2 over 1 can be dictated by two factors.

Complexity of the regimented activities Regimentation can considerably raise the complexity of the activities that agents carry out within the institution, so that for an agent to pursue its goals it would be necessary to go through too complex procedures. This is illustrated by a simple example: consider a postal service in which the deliverer should wait for the addressee to open his/her parcels and confirm the content has been delivered in the desired state. This would rule out the possibility of deliveries of damaged parcels, but it would also make the delivery process considerably slower and inconvenient for the agents which should always be present at the delivery. In other words, regimentation can give rise to computationally demanding activities (see [Vázquez-Salceda et al., 2004]) both for the institution itself, and for the agents acting within it. This aspect has directly to do with the delicate balance between the two fundamental goals of e-Institutions, i.e, increase trust in agents' transactions and facilitate those transactions [Vázquez-Salceda, 2004]. The point is that, although via regimentation the highest level of trust can be achieved, agents' interaction can end up being not facilitated at all.

Usefulness of the violations As we have seen in Example 5.11 the possibility for agents to violate the primary regulation would allow for activities which would otherwise be impossible. The agent can choose to violate the regulation and possibly incur a sanction in order to pursue some specific goals. In Example 5.11 agents playing chess in an institution with a flawless referee would actually have the possibility to use a wider variety of strategies for winning the game by trying to distract the opponent via performing invalid moves. Alternatively, suppose a reputation value to be attached to each chess-playing agent so that the less often they violate the norms the higher reputation they get. In this case, the possibility to violate the norms enables also the possibility to introduce a reputation value system which might be useful for further purposes: for instance, a high reputation value might be required to access chess tournaments.

In the end, allowing for violations results in a higher flexibility of the e-Institutions which might happen to serve more purposes than the ones for which it was designed. As stressed in [Castelfranchi, 2004], violations can be functional for the institution as a whole in as much as they trigger institutional evolution.

Before closing the section it should be stressed, as we already did in 1.2.3, that in order to comprehensively address norm-implementation problems formal disciplines such as mechanism design and implementation theory ([Jackson, 2001]) should be taken into consideration, and possibly enhanced with logic-based verification techniques as argued in [Pauly and Wooldridge, 2003; Pauly, 2005a,b]. This is an interesting research line worth pursuing in future work.

5.7 On Norms as Terminological Axioms

The previous sections have touched upon a number of issues related with the representation of institutions and it has shown how those issues can be tackled in terminological logics. Before closing the chapter though, we want to come back to one of its read threads, namely the formal representation of norms as terminological axioms. The present section provides some final considerations on this central issue.

5.7.1 Regulative vs. constitutive: dropping logical distinctions

In Section 4.7.3 we have shown how the formal analysis of counts-as naturally leads to a reduction of regulative norms to constitutive ones. Regulative norms are those which 'constitute' the notion of violation for the institution they define. That is precisely what some terminological axioms in any institution specification $\mathfrak{Ins}$ do by stating constraints on the interpretation of the violation state type *viol*, or of an appropriately chosen family of violation state types (e.g., $\{viol(i)\}_{i \in Agents}$). In this view, regulative norms are therefore just state type terminological axioms of the following form:

$$\textbf{\textit{Prohibition:}} \quad \gamma \sqsubseteq \forall \alpha.viol \tag{5.42}$$

$$\textbf{\textit{Permission:}} \quad \gamma \sqsubseteq \exists \alpha.\neg viol \tag{5.43}$$

Formula 5.42 expresses that states of type γ are all states such that by executing a transition of type α a violation state is always reached. That is to say, transitions of type α are forbidden in γ states. When transition type $\alpha = \text{non\_}a$ for some a, i.e., when α is the negation of an atomic transition type a then Formula 5.42 expresses that transitions of type a are obligatory. Formula 5.43 expresses a notion of permission: by performing a transition of type α in γ states, violation states are not necessarily reached.

By representing norms as terminological axioms, we conceive of them as stating global properties of transition systems, i.e., properties holding in every state of the system. These are exactly the "universal criteria" mentioned in the opening quote of Chapter 2 taken from [Husserl, 1988].

5.7.2 Pushing Anderson's envelope

Terminological axioms of the form $\gamma_1 \sqsubseteq \gamma_2$ take us back to the modal intuition we discussed in Section 4.2 at the beginning of Chapter 4, which laid the ground for the formal analysis of counts-as and constitutive norms developed in that chapter: from the subsumption statement $X \sqsubseteq Y$ to the strict implications $[u](X \to Y)$ where $[u]$ is the "box" of the universal modality. In fact, viewing norms as terminological axioms takes Anderson's reduction, which we have recalled in Section 2.5.1, to its extreme theoretical consequences by conceiving the necessity of his reduction as a terminological, and thus universal, kind of necessity. As a consequence, the distinction between regulative and constitutive norms just seems to vanish. All norms acquire a definitional nature, which we have seen to be the characteristic feature of constitutive norms.

We have already touched upon Anderson's reduction of deontic logic ([Anderson, 1957, 1958]) a couple of times in the course of our analysis (Sections 2.5.1 and 4.7.3). It is time to make explicit how it precisely relates to the view of norms to which we have committed.

By interpreting the necessity of Anderson's reduction as universal necessity, i.e., the one presupposed by terminological axioms, this is what we get:

$$\textbf{Ought-to-be prohibition:} \quad [u](\gamma \to viol) \tag{5.44}$$

$$\textbf{Ought-to-be permission:} \quad \neg[u](\gamma \to viol) \tag{5.45}$$

where $[u]$ is the "box" of the universal modality. All γ states are violation states (Formula 5.44), and there are some γ states that are not violation states (Formula 5.45). Notice that Formula 5.44 can be represented as a DL subsumption statement, while Formula 5.45 represents the satisfiability of concept $\gamma \sqcap \neg viol$.

As to the reduction of ought-to-do deontics[9], notice that formulae 5.42 and 5.43 are a DL version of the dynamic logic reduction of deontic logic first advanced in [Meyer, 1988]. In that work, the deontic notions of ought-to-do prohibition and permission are represented in dynamic logic as $[\alpha]viol$ and, respectively, as $\langle\alpha\rangle\neg viol$. However, given the dynamic logic setting, such formulae were predominantly used to express *local* properties of transition systems while by specifying a TBox in DL we are only interested in stating global properties of the system. More precisely, Formulae 5.42 and 5.43 correspond thus to a 'universalization' of the reduction approach proposed in [Meyer, 1988]:

$$\textbf{Ought-to-do prohibition:} \quad [u](\gamma \rightarrow [\alpha]viol) \tag{5.46}$$

$$\textbf{Ought-to-do permission } \textit{\textbf{à la}} \textbf{ Meyer:} \quad [u](\gamma \rightarrow \langle\alpha\rangle\neg viol) \tag{5.47}$$

$$\textbf{Ought-to-do 'actual' permission:} \quad \neg[u](\gamma \rightarrow [\alpha]viol) \tag{5.48}$$

Again, notice that Formula 5.48 expresses the satisfiability of concept $\gamma \sqcap \neg[\alpha]viol$. It expresses that there is at least one γ state where executing α does not necessarily lead to a violation. This is a stronger property that the one expressed in Formula 5.47 which just states that all γ states are states where α can be executed without necessarily leading to a violation. This does not guarantee that γ states actually exist though.

It might be worth noticing that contextual versions of Formulae 5.44-5.48 can be naturally obtained, as shown in Chapter 4, by means of modalities $[c]$ interpreted on secondarily universal instead of universal accessibility relations.

5.7.3 Regulative vs. constitutive: regaining a distinction

Now the question which should be answered is, of course, why to keep the distinction between normative and constitutive rules at all? Is not our Anderson-like reduction flattening a theoretically relevant distinction?

The point we wish to stress is that the theoretical upshot of the reduction of regulative rules to constitutive ones, which was first exposed in Section 4.7.3, is just that constitutive and regulative rules are, from a formal point of view, just the same. To put this in a rather more positive perspective, this amounts to saying that the reasoning patterns concerning constitutive or regulative rules can be captured by the same logic. From the point of view of an economy of thought in Ockham's sense, this might be regarded as something desirable. After all it seems safe to claim that "logicae non sunt multiplicandae praeter necessitatem[10]" (logics should not be

[9]It is good to recall the standard deontic logic distinction between ought-to-be and ought-to-do normative statements. Ought-to-be statements concern the deontic properties of state types, while ought-to-do ones concern the deontic properties of transition types. On this distinction, see for instance [von Wright, 1983].

[10]It might be instructive to mention, in passing, that the notorious motto we are referring to, i.e., "entia non sunt multiplicanda praeter necessitatae" (entities should not be multiplied unnecessarily), does not appear in any of Ockham's writings. A near statement is "Frustra fit per plura quod potest fieri per pauciora" (it is vain to do with more what can be done with less), to be found in the Summa Totius Logicae (i.12). [Thorburn, 1918] offers an interesting historical discussion on the actual sources of the motto "entia non sunt multiplicanda praeter necessitatae".

multiplied unnecessarily).

An important difference remains, even though it is not of a formal nature. It manifests itself by bringing agents' mental states into the picture. As described in [Conte and Castelfranchi, 1995] agents that can autonomously comply with institutional norms (i.e., normative agents [Dignum, 1999]) are able to create an internal copy of norms (i.e., the *norm instance* [Conte and Castelfranchi, 1995]) that is used as a motivational attitude from which they can infer norm-compliant goals. It is our claim that, while this is the case for regulative norms, it is not for constitutive ones. The mental attitudes related to the norm instances of constitutive norms have nothing to do with motivational attitudes, such as goals, but rather with epistemic ones and specifically with beliefs. So, regulative norms can be reduced to constitutive ones when it comes to logical form, but they do keep their identity in the mind of the agents, so to say.

Let us expand on this. How can an agent distinguish between those subsumption statements which represent regulative norms from the ones that represent constitutive norms? That depends on the concepts occurring in the subsumption statements. Regulative norms are subsumption statements concerning concepts with a 'motivational meaning', so to say. In theory of norms and ethics (see for instance [Husserl, 1988; Moore, 1903; Hare, 1952]) these concepts are often referred to as evaluative. Typical examples are the terms 'good' or 'bad', and in the context of institutions the term 'violation'. In a way, the state type 'violation' could be relabeled by something like "don't get there!". If an agent internalizes a norm classifying certain states as violations, then it will believe that those states are violation states, but it will also avoid including those states among its goals. To put it in a nutshell, regulative norms are subsumption statements classifying evaluative notions.

On the other hand, subsumption statements that handle non-evaluative, or descriptive, notions influence only the epistemic attitudes of agents. By creating a norm instance of a constitutive norm, an agent just believes that some institutional state-of-affairs occur under some conditions. By creating a norm instance of the norm "bikes count as vehicles", agent i will just believe that, for instance, if it rides a bike then it actually drives a vehicle.

Interestingly enough, the view just exposed is highly reminiscent of the thesis advanced in [Boella and van der Torre, 2003; Boella and Van der Torre, 2004; Boella and van der Torre, 2005], of viewing counts-as conditionals as the conditional beliefs and regulative norms as the conditional goals of an 'agentified' normative system. We have touched upon this work, which differs considerably from ours in its formal aspects, in Section 4.8.4. In any case, the issue concerning the mechanisms governing the interaction of norms and agents' mental states has not yet been —to our knowledge— thoroughly addressed.

To recap, even though agents can reason about constitutive and regulative norms by following the very same logic, their mental attitudes are influenced by those types of norms in different ways: if internalized, regulative norms motivate, while constitutive norms determine belief updates. This completes the answer to our third research question.

5.8 Institutions and Transition Systems

The main aim backing the analysis of institutions presented in this chapter consisted in giving an interpreted transition system semantics to a number of key institutional notions.

In the following two sections we briefly summarize the basic ideas of three formal approaches to the representation of institutions which bear some essential similarities with ours. In particular, they all model institutions by means of transition systems and all account for deontic notions via an Anderson-like reduction.

5.8.1 Institutions in *Normative Temporal Logic*

Normative temporal logic (NTL, [Ågotnes et al., 2007]) is a generalization of Computational Tree Logic, better known as CTL ([Emerson, 1990]), enabling, instead of the usual CTL path quantifiers A and E, two families of path quantifiers $\{\mathrm{O}_\eta\}_{\eta \in NS}$ and $\{\mathrm{P}_\eta\}_{\eta \in NS}$. The set NS denotes the set of subrelations η of the temporal tree R such that $R \backslash \eta$ is a total relation. The intuitive idea behind this formalism is that every sub-relation η of R represents all the transitions (w, w') which are forbidden according to a normative system. The complement $R \backslash \eta$ denotes instead the legal transitions according to η and it is, for obvious reasons, required to be total.

It is thus possible in NTL to express that a certain temporal formula, e.g., $\bigcirc \phi$ (ϕ is true in the next step) is obligatory according to normative system η if an only if for all the η-compliant paths in the model (i.e., all paths that do not contain any element of η) ϕ is true at the next state. A number of deontic notions are similarly representable.

The difference between our DL based approach and NTL is, first of all, exactly the same that can be mentioned in comparing temporal logics with dynamic logics: dynamic logic supports an explicit representation of actions while temporal logic does not. In addition, our approach allows for a rich representation of the 'imposition' of institutional properties on states and transitions. However, the definition of normative systems as relations η could be thought of as a form of labeling of systems' transitions. Instead of labeling states as not norm-compliant, in NTL transitions are labeled as not norm-compliant. In a way each η can be viewed as a transition type FORBIDDEN(η) labeling the transitions which are not η-compliant.

5.8.2 Language *nC*+

Language *nC*+ ([Sergot, 2004; Sergot and Craven, 2006]) is a formalism for representing normative aspects of institutions. It extends language *C*+ ([Giunchiglia et al., 2004]) which is a formalism for specifying and reasoning about the effects of actions in general, and the persistence of facts over time, i.e., the so-called 'inertia' of facts. Essentially, *nC*+ adds to *C*+ the possibility of specifying the permitted or legal states of a transition system and its permitted or legal transitions. It does this by specifying sets of state and transition labeling rules, some of which handle deontic labels (green states and green transitions).

Technically speaking, $nC+$ is not a logic but a logic-based formalism for writing down finite labeled transition systems, which can model desired institutions including their deontic component. In this consists the main difference with our DL based approach: a TBox has many possible models, i.e., many transition systems, while an $nC+$ specification describes precisely one transition system. Our representation choice is of course dictated by our analysis. If an institution is viewed as a set of norms, then an institution specification should have a number of possible concrete systems all modeling those norms and not just one.

However, if we make the step from abstract institution specifications to concrete system specifications, then $nC+$ can indeed be thought of as a formalism for specifying institutions at what we called the infrastructural level, i.e, concrete institutions (Sections 5.5.1 and 5.5.2). While TBox specifications can just set further constraints on the possible sets of models, an $nC+$ would be able to fully describe a transition system modeling the abstract institution specification. In fact, at that level $nC+$ has a definite advantage over DL, namely that of being able to represent the persistence of state types over time.

5.8.3 Institutions and agents' protocols

The formal representation of institutions allows for the design and verification of protocols that can be used by agents in order to pursue their goals by complying, at the same time, with the norms of the institution. This is a delicate issue especially when agents have to operate in heavily regulated domains, where the number of norms to be complied with is high. In [Aldewereld et al., 2006a,b; Aldewereld, 2007], such issue is addressed by assuming a perspective on institutions which is very much related to the one defended in this chapter: institutions are thought of as sets of norms, and norms are formalized in a formalism based on temporal logic and an Anderson-like reduction of deontic notions.

The work presented in [Aldewereld et al., 2006b; Aldewereld, 2007] adapts then verification techniques from concurrent programming which allow for testing safety and liveness properties of protocols specified in linear-time temporal logic. By means of such techniques it is possible to test, against the formal specification of a set of norms, whether a given protocol is violation-free (safety) and if it successfully reaches the states it is suppose to reach (liveness). We think, given the similar transition systems semantics, that such techniques could be easily adapted, if not directly applied, to support the verification of protocols against norms specified in the framework presented in this chapter.

5.9 Conclusions

The chapter has shown how TBoxes can be used for representing the specification of institutions (Definition 5.4) and their infrastructures (Definition 5.5). We have provided several examples showing how key institutional notions such as the relation between abstract and concrete norms, roles, and norm enforcement can be represented and given a labeled transition systems semantics.

Our key representational choice, i.e., “norm = terminological axiom”, has been thoroughly discussed and the distinction between regulative and constitutive norms have been clarified in that light (Section 5.7): while constitutive and regulative norms obey exactly the same logic, they bear effects of different kinds for the mental states of the agents. While constitutive rules, once internalized by an agent, modify its beliefs, regulative rules modify its goals.

The boundaries of tractable specifications of institutions in DL have been addressed in Section 5.4. Finally, some related approaches sharing the understanding of institutions in terms of transition systems and Anderson-like reductions of deontic notions have been touched upon in Section 5.8.

Chapter 6

Organizations as 'Meaningful' Graphs

"Als je voor elke positie de beste speler kiest, heb je nog geen sterk elftal maar een team dat als los zand uiteen valt."

"If you choose the best player for each position, you still don't get a strong team but a group which falls apart as loose sand."

J. Cruijff, Dutch football folklore

Several methodologies for MAS are based on organizational structures as their cornerstones such as, for instance, OPERA [Dignum, 2003], TROPOS [Bresciani et al., 2004], and GAIA [Zambonelli et al., 2003]. However, it has been stressed in Chapter 1 that formal tools for the rigorous representation of organizational structure in MAS are not yet available and the study of structure is still mainly informal.

The chapter advances some proposals on how to rectify this situation by answering the fourth research question which, we recollect, concerns the following two issues:

1. What are the graph-theoretical properties of the links connecting the roles in the structure?
2. What is the 'meaning' of those links, that is to say, the effects they have on the activities of the agents operating the organization?

In order to find answers to these questions theses and established techniques are imported from sociological literature (in particular organization theory and social network analysis), adapted to MAS and partially extended with logic. By applying some formal methods which are quite commonplace in social theory (e.g., graph theory), this chapter aims at pointing to a whole body of literature and results which, we are convinced, could be profitably used with minor adaptations also for the analysis an design of MAS.

The results exposed build on [Grossi et al., 2005a, 2006d,c, 2007]. The exposition follows this outline. In Section 6.1 the basic intuitions and concepts grounding our analysis are introduced, which are imported from work on organization theory. Section 6.2 addresses the issue of the meaning of organizational structures in terms of the activities of the agents playing the roles in the structure. Labeled transition systems are used again as the underlying mathematical tool. Roughly, the central thesis will consist in viewing the links in a structure as statements of global properties of transition systems. Such global properties concern the activities that can be executed by agents playing given roles and the effects of those activities. It becomes then possible to link this semantic analysis of links with desirable graph-theoretical properties of organizational structures. If each link is related with specific activities, then certain configurations of links would make certain complex activities or interactions possible, whereas different configurations would make different activities and interaction patterns possible within the organization. Section 6.3 addresses this issue. We turn then to the issue of the adherence of organizational structures to desirable criteria such as, for instance, robustness and flexibility. A great deal of ongoing research in the field of organization-based MAS is devoted to compare and evaluate different types of organizations and their performances. Work on these issues varies from surveys comparing organizational paradigms [Horling and Lesser, 2004], to frameworks for representing and verifying organizational designs [Horling and Lesser, 2005; van der Broek et al., 2005], to studies concerning properties and performance of specific types of organizations [Scerri et al., 2004; So and Chon, 2005]. Sections 6.4 and 6.5 are devoted to the exposition of graph-theoretical metrics for the quantitative analysis of organizational structures. Section 6.6 discusses some related work and draws the attention to future research lines. Conclusions follow in Section 6.7.

6.1 Preliminaries

This section introduces the theoretical backbone of the chapter, which is borrowed from work on organization theory, and the notion of organizational structure as multigraph.

6.1.1 Elements of a theory of organizational structure

Organizations "represent rationally ordered instruments for the achievement of stated goals" ([Selznick, 1948]), that is, organizations arise in order to achieve specific objectives, and these objectives are pursued defining a number of subgoals contributing to the overall purpose of the organization. These subgoals identify the roles that are played in the organization. The relation between subgoals and overall objectives of the organization, i.e., the primitive decomposition of tasks within the organization, defines the essential form of organizational structure: "viewed in this light, formal organization is the structural expression of rational action" [Selznick, 1948]. Roles are the basic units over which this structure ranges determining the source of the "rational order" holding in the organization. The above quotes consider the decomposition of tasks as the source of structure within organizations: structures are the organizational tools for pursuing organizational goals. The chapter addresses these tools as formal objects in the context of MAS.

Work on organization in MAS[1] presents organizational structure as something essentially mono-dimensional, though it often, but only implicitly, considers a multiplicity of structured aspects: "authority", "communication", "delegation", "responsibility", "control", "decision-making", "power", etc. The thesis we hold here, which is inspired by foundational work on social and organization theory like [Selznick, 1948; Morgenstern, 1951; Giddens, 1984], is that organizations do not exhibit one single structural dimension, but that they are instead multi-structured objects. In particular, we view organizational structure as hiding at least three relevant dimensions which we call: power, coordination and control.

What characterizes organizations is the possibility for the agents enacting the roles to delegate to other agents some of the tasks they are supposed to execute. Delegation consists in the possibility for an agent enacting a role to transfer a given task to a somehow subordinated one. This transfer takes place in the form of a directed obligation ([Dignum, 1999]) from the agent enacting the first role to the agent enacting the subordinated one, the content of the obligation being the to-be-executed task. As a consequence of delegation, the addressee of the delegation becomes obliged to execute the task which belonged to the first agent. The possibility of delegating goals constitutes one of the essential aspects displaying what is usually called "delegation" or "power" structure of an organization ([Ioerger and Johnson, 2001]): who delegates to whom?

Since tasks can be properly executed or not, organizations need to engage in activities of performance control. In its simplest form, control consists in a monitoring activity triggering appropriate reactions to determine failures or violations. If an agent fails in executing one of the stated or delegated tasks, some sort of supervising agent should take over the execution of that task: organization calls for a form of supervision activity ([Giddens, 1984]). Potentially, the execution of any task can be object of control. Because of this, control can result in a complex activity, and it is not for nothing often viewed as "an organization within an organization" ([Morgenstern, 1951]). With respect to control the relevant structural question is: who controls (supervises) whom?

The activity of an organization also relies on the *coordination* structure, which is a broadly investigated topic in MAS studies. Here, following [Decker and Lesser, 1995; Grossi et al., 2004] we adopt a simple view on coordination, reducing it to the issue of the information with which agents enacting specific roles should be endowed in order to properly execute their tasks. Roles should have at disposal the information necessary for agents to appropriately enact them. This turns into

[1] See [Horling and Lesser, 2004] for a survey.

a knowledge problem of the state of the organization (or of part of it[2]) at a given moment. Agents should know when to act, that is, they should be informed about the status of the activities of the organization upon which their activities depend, and what they are supposed to do. As we observed above, delegation introduces a dynamics in the task distribution of an organization. The point is that once a task is delegated and a correspondent obligation arises for a specific agent, a certain amount of information might be required for that agent to execute that task. Such information should therefore 'flow' within the organization to that agent. Because of this, an information mechanism which can keep track of this dynamics is crucial for the performance of an organization.

> "The description of a delegation system [delegation structure] is incomplete unless the simultaneous signaling system [information structure] applied to it is also explicitly described" ([Morgenstern, 1951], p.17).

The coordination structure should guarantee that each agent is endowed with a representation of the actual state of the organization which is sufficient for it to properly enact its role. The question is then how the access and sharing of knowledge is structured within the organization: who informs whom?

To recapitulate, the *power structure* defines the task delegation patterns possible within the organization. The *coordination structure* concerns the flow of knowledge within the organization, and the *control structure* has finally to do with the task recovery functions of the organization. In other words, the existence of a power link between role r and role s implies that every delegation of tasks from agent r (agent enacting role r) to agent s (agent enacting role s) ends up in the creation of an obligation directed to agent s to execute the delegated task. If r and s are connected via a coordination link, then every information act from r to s ends up in creating the corresponding knowledge in agent s. Finally, a control link between r and s implies that agent r has to take over the tasks of agent s in case it fails to perform the requested tasks. As a result of this analysis, organizations will be represented as explicitly displaying a triple structure constrained on the basis of the interplay between the three notions of power, coordination, and control.

6.1.2 Two notes

It is worth stressing two points before we move on to the next sections. First, even though our analysis will focus on the three aspects of delegation, supervision and information introduced in the previous section, it is clear that these aspects are not the only ones playing a role in organizations. In particular, depending on the MAS application in mind, other dimensions of organizational activity can be

[2]Notice, in passing, that the amount of knowledge to be propagated through the organization also constitutes an important issue:

> If every competence [role] had full information about every other it might help but not necessarily; it would clearly be wasteful, if not physically impossible, for most organizations ([Morgenstern, 1951], p.28).

We do not address this issue here.

isolated (e.g., rights to access resources and their flow). However, it is our thesis that the method we follow in analyzing the three aspects introduced can be used for the analysis of any further structural dimension which might appear useful for a specific application domain. In a way, the work presented here is best read as an illustrative exposition of how certain formal tools can be applied to the analysis of the various aspects involved in organizational structures. The chosen aspects, i.e., delegation, supervision and information, have been chosen only because they appear to be quite recurrent in the work on organization theory upon which we have decided to ground our analysis ([Selznick, 1948; Morgenstern, 1951; Giddens, 1984]).

Second, the terms we use to refer to the three structural dimensions related to delegation, supervision and information, that is, power, control, and coordination are used in a technical and very specific way. The work presented has no ambition to provide full analysis of such multi-faceted phenomena such as power or coordination within organizations. We needed terms to denote the different dimensions of structure related to delegation, supervision and information and we settled for those ones. Different terminological choices are, needless to say, possible.

6.1.3 Organizational structure

A natural way of modeling the notion of organizational structure emerged in Section 6.1.1 is via directed multigraphs (or, multidigraphs), which we represent here as systems of relations.

Definition 6.1. (Organizational structure)
An organizational structure $\mathfrak{Os}$ *is a tuple:*

$$\langle Roles, R_{Pow}, R_{Coord}, R_{Contr} \rangle$$

where Roles is a non-empty finite set of roles, and $R_{Pow}, R_{Coord}, R_{Contr}$ *are three irreflexive binary relations on Roles characterizing the Power, respectively, the Coordination and the Control structures.*

For every R_k s.t. $k \in \{Pow, Coord, Contr\}$, we denote with $Roles_k$ the smallest subset of *Roles* such that, if $(x, y) \in R_k$ then $x, y \in Roles_k$. In other words, sets $Roles_k$ denote the set of roles involved in the structural dimension k. Each digraph $\langle Roles_k, R_k \rangle$ in $\mathfrak{Os}$ will be also referred to as the *structural dimension k* of $\mathfrak{Os}$.

We consider the roles on which the organizational structure ranges (i.e., the elements of set *Roles*) to be enacted by one and only one agent. The reason for this choice is illustrated by the following example taken, like the opening quote of the chapter, from the soccer world.

Suppose we need to model an organization for a soccer team implementing a 4-3-3 strategy. in such a way that the underlying organizational structure of the 4-3-3 strategy is made explicit. Three roles can be defined in every team: 'attacker', 'defender' and 'midfielder', which are connected by appropriate power, control and coordination relations. An option would be to model the organization via imposing complex enactment constraints such as: "the role 'attacker' should be enacted by three agents such that the first agent should communicate with the third one, the second agent should monitor the first and third ones, etc.". However, this would make implicit in the enactment constraints the power, coordination and control links that are present between all the various attackers in the 4-3-3 strategy. A better option would be to explicitly define three new roles, which can be seen as specializations of the 'attacker' role and which can be enacted by only one agent. The organizational links existing between these three new roles could thus be made explicit, and the resulting organizational structure satisfactorily modeled. This is the perspective we assume in the present work.

In practice, this boils down to a modeling issue: if two agents enacting a same role have to be connected by power coordination or control links, then two different roles have to be specified which substitute the first one and which are played by only one agent. This finer level of granularity allows for a more refined snapshot of the structure of an organization. An analysis at a level where roles do not specify the relative positions of all agents with respect to all the structural dimensions would just fall short, missing many possibly relevant structural links. It follows from this distinction that a study of the organizational structure ranging on *role types* would abstract from those power, coordination, and control links that might be present between the *role tokens* specializing the same role type (e.g., the three attackers in a 4-3-3 strategy). Here we are interested in the analysis of structure at the level of the actual agents' positions within the organization, and thus at a finer level of granularity. The elements of the set *Roles* in a $\mathfrak{Os}$ are then to be considered role tokens. In the rest of the paper, if not stated otherwise, we use the word role intending role token.

6.1.4 Enactment configuration

Roles are positions in a structure, but they are positions which are occupied by agents. An enactment configuration for an organizational structure $\mathfrak{Os}$ is a functional relation $EC \subseteq Roles \times Agents$ making explicit which agent out of a set *Agents* enacts which role: agent $i \in Agents$ enacts role $r \in Roles$ iff $EC(r, i)$. Relation EC is functional since, as stressed in the previous section, each role can be played by at most one agent.

A concrete organizational structure $\mathfrak{COs}$ is a structure $\langle \mathfrak{Os}, Agents, EC \rangle$, i.e, an organizational structure plus a set of agents and an enactment configuration.

6.2 Structure with Formal Meaning

Definition 6.1 makes the structural aspect of an organization explicit, at least as far as the three dimensions of power, coordination and control are considered. Before studying organizational structures in graph-theoretical terms it is important to first face a second representational issue. Even after deciding to treat organizational structures as as multidigraphs the problem remains of what each link in a structure actually means. While this problem is substantially disregarded in the network analysis of human organizations, probably because of their considerable complexity, it is hard to claim that such issue is of secondary importance for the design of multi-agent systems by means of organizational concepts.

Roughly, the existence of a link of a certain type between two roles r and s indicates that a certain action can be performed by the agent playing r with respect to the agent playing s (notice that this is what was called an executability law in the previous chapter), and that this action has some relevant consequences (notice that this is what was called an effect law in the previous chapter). The section is devoted to develop this simple intuition showing how organizational structures can obtain a semantics in terms of labeled transition systems.

To pursue this aim, we will conceive of institutional links as concise ways to refer to sets of terminological axioms. To draw a link in an organizational structure amounts to state a number of universal properties for the to-be-developed system. Such properties, which, in line with what was discussed in the previous chapters, are properties of transition systems, represent in some way the "meaning" of the organizational links. From the technical point of view the problem amounts to translate a $\mathfrak{COS}$ in a corresponding TBox. How this translation, which we denote with $\mathcal{T}$, is defined is exposed in the coming sections. The idea is to express the intuitive understanding of the relations of power, coordination and control which has been exposed in Section 6.1.1 via terminological axioms.

The formal framework we will work in is the one exposed in Section 5.1.2 of the previous chapter.

6.2.1 Institutional links

We say that role r has power over role s ($R_{Pow}(r,s)$ or $(r,s) \in R_{Pow}$) if and only if agents playing role r can always execute a delegation action towards agents playing role s and this action always result in a corresponding obligation for the recipient to execute the delegated action. The delegation action type is represented by the transition type forms DEL($\underline{a}$), where $\underline{a}$ is the delegated transition type form.

The translation we are looking for should therefore translate power links as follows.

Definition 6.2. *(Meaning of power links)*
Consider a $\mathfrak{COS} = \langle Roles, R_{Pow}, R_{Coord}, R_{Contr}, Agents, EC\rangle$. For all (r,s), if $R_{Pow}(r,s)$ then $\mathcal{T}(\mathfrak{COS})$ contains the following axioms for any transition type form $\underline{a}$ and agents $i, j \in Agents$:

$$rea(i,r) \sqcap rea(j,s) \sqcap \forall non\_\underline{a}(i).viol(i) \sqsubseteq \exists DEL(\underline{a},i,j).\top \tag{6.1}$$

$$rea(i,r) \sqcap rea(j,s) \sqcap \forall non\_\underline{a}(i).viol(i) \sqsubseteq \forall DEL(\underline{a},i,j).(\forall non\_\underline{a}(j)).viol(j) \tag{6.2}$$

$$id(rea(i,r) \sqcap rea(j,s)) \circ DEL(\underline{a},i,j) \sqsubseteq \underline{a}(i) \tag{6.3}$$

Intuitively, the fact that r is linked to s by a power link means that agents enacting role r, if put under the obligation to perform $\underline{a}$, can delegate action $\underline{a}$ to agents enacting role s (Formula 6.1), and they can do this successfully in the sense that they always create the corresponding obligation directed to the recipient of the delegation act (Formula 6.2). The last important aspect of the semantics of power links we have in mind is that agents playing role r actually execute $\underline{a}$ by delegating its execution to agents playing role s. In other words, if an agent enacting role r delegates $\underline{a}$ it does not end up in a violation state.

It goes without saying that this interpretation of power is extremely restrictive[3]. It is not not our aim to provide a semantic analysis of the notion of power as it has been done in Chapter 4 for the notion of counts-as. The point we are making here is rather of a methodological kind: structural links have a meaning, and this can be given in terms of transition systems as shown above for the notion of power adopted here.

Along the same lines we can provide a translation of the links in the control structure.

Definition 6.3. *(Meaning of control links)*
Consider a $\mathfrak{CDS} = \langle Roles, Agents, R_{Pow}, R_{Coord}, R_{Contr}, EC\rangle$. *For all* (r, s), *if* $R_{Contr}(r, s)$ *then* $\mathcal{T}(\mathfrak{CDS})$ *contains the following axioms for any transition type form* $\underline{a}$ *and agents* $i, j \in Agents$*:*

$$rea(i, r) \sqcap rea(j, s) \sqcap \forall non\_\underline{a}(j).viol(j) \sqsubseteq \forall non\_\underline{a}(j).(\forall non\_\underline{a}(i).viol(i)) \quad (6.4)$$

The fact that role r is linked to role s by a control link means that if agents playing role s ought to perform action a then agents playing role r ought to perform a in case the first agents fail to fulfill that obligation. Intuitively, a control link generates a sort of backup obligation addressed to the controllers. As such, control has to do with the deontic dimension of the organization. The above considerations concerning the restrictedness of the interpretation of power apply of course also to this notion of control.

To control, is here intended in the sense of adjusting deviant behaviour or recovering from failure. It concerns the generation of obligations triggered by the occurrence of violations. A different notion of control interpreted as some sort of sanctioning activity by an enforcer (see Chapter 5) would exhibit in fact the same logical pattern:

$$rea(i, r) \sqcap rea(j, s) \sqcap \forall \text{non\_}\underline{a}(j).viol(j) \sqsubseteq \forall \text{non\_}\underline{a}(j).(\forall \text{non\_SANCTION}(i, j).viol(i))$$

for any transition type form $\underline{a}$ and agents $i, j \in Agents$. If r controls s then agents playing role r ought to sanction agents playing role s in case they violate some norm.

Notice that Formulae 6.2 and 6.4 are effect laws, Formula 6.1 is an executability law and Formula 6.3 is a transition type inclusion statement.

However, structural links do not only have institutional meanings. The following section considers the semantic of structural links in terms of mentalistic notions, that is, in terms of the mental states of the agents involved in playing the roles of the link.

[3]See [Castelfranchi, 2003] for a comprehensive overview of this issue.

6.2.2 Mentalistic links

We have seen in Section 6.1.1 that organizational structures have to do –in a way eminently– with the flow of information within an organization. In this view, an essential ingredient of coordination between two roles is that agents enacting the first role can communicate with agents playing the second role and can do that effectively, that is, actually modifying the epistemic state of the recipient. This perspective takes into the picture well-investigated mentalistic notions such as knowledge and belief.

In what follows, we will interpret the coordination links of an organizational structure by means of the notion of knowledge. Coordination links have to do with the flow of knowledge between the agents enacting the roles of the organization. There is a number of formalisms which investigate in depth the relation between knowledge and dynamics (for example [Meyer and van der Hoek, 1995; Baltag and Moss, 2004]). For our purposes, it suffices here to show how we can represent knowledge in the framework of DL, thus avoiding to introduce further machinery. Again, the point we are stressing is that organizational structure can be given a semantics, in this case, of a mentalistic nature.

From a technical point of view, to represent knowledge in DL nothing else is needed than the introduction of special transition types IND(i), for any agent i, representing a relation of indistinguishability, for agent i, between system states. It is then needed to state that these transition types denote reflexive, symmetric and transitive relations. Such expressivity is not common in DL. It is however enabled for instance in DL $\mathcal{SROIQ}$ ([Horrocks et al., 2006]) which allows for TBoxes containing role assertions of the type: $\mathtt{Sym}(a)$, $\mathtt{Trans}(a)$, $\mathtt{Refl}(a)$, where a is an atomic transition type. The semantics of such assertions is the obvious one.

Assuming such new expressivity for the underling DL language, it becomes then possible to represent the notion of knowledge, and to give a mentalistic semantics to coordination links. The information action type is represented by the transition type forms INF(γ), where γ is a state type description belonging to a finite set of "communicable" state types[4].

Definition 6.4. *(Meaning of coordination links)*
Consider a $\mathfrak{CDs} = \langle Roles, R_{Pow}, R_{Coord}, R_{Contr}, Agents, EC\rangle$. *For all* (r, s), *if* $R_{Coord}(r, s)$ *then* $\mathcal{T}(\mathfrak{CDs})$ *contains the following axioms for any transition type form* $INF(\gamma)$ *and agents* $i, j \in Agents$:

$$rea(i, r) \sqcap rea(j, s) \sqcap \forall IND(i).\gamma \sqsubseteq \exists INF(\gamma, i, j).\top \qquad (6.5)$$

$$rea(i, r) \sqcap rea(j, s) \sqcap \forall IND(i).\gamma \sqsubseteq \forall INF(\gamma, i, j).(\forall IND(j).\gamma) \qquad (6.6)$$

Intuitively, the fact that r is linked to s by a coordination link means that if agents enacting role r know that γ they can always inform agents enacting role s that γ (Formula 6.5), and they can do this successfully in the sense that they always create the corresponding knowledge in the recipient of the information act (Formula 6.6).

[4]The reason for this restriction is that we want to keep the alphabet of the language finite.

6.2.3 Enactment links

An enactment configuration EC is translated by means of the state type forms $rea(r)$, for any $r \in Roles$.

Definition 6.5. *(Meaning of enactment links)*
Consider a $\mathfrak{COS} = \langle Roles, R_{Pow}, R_{Coord}, R_{Contr}, Agents, EC\rangle$*. For all* (r, i)*, if* $EC(r, i)$ *then* $\mathcal{T}(\mathfrak{COS})$ *contains the following axiom:*

$$rea(i, r) \equiv \top \tag{6.7}$$

Intuitively, an enactment configuration fixes which agents are playing which roles as a global property of the system.

6.2.4 Institutional links based on mentalistic notions

Especially when considering concrete organizational structures $\mathfrak{COS}$, i.e., $\mathfrak{OS}$ which are instantiated by a given group of agents $\mathfrak{OS}$, the mentalistic dimension becomes essential in the analysis of the organization. Organizational structure relies in that case also on how much of the organization itself is known to the agents. In this respect, Definitions 6.2 and 6.3 should be restated by making the knowledge ingredient explicit.

Definition 6.6. *(Power and control based on knowledge)*
Consider a $\mathfrak{COS} = \langle Roles, R_{Pow}, R_{Coord}, R_{Contr}, Agents, EC\rangle$*. For all* (r, s)*, if* $R_{Pow}(r, s)$ *then* $\mathcal{T}(\mathfrak{COS})$ *contains the following axioms for any transition type form* $\underline{a}$ *and agents* $i, j \in Agents$*:*

$$rea(i, r) \sqcap rea(j, s) \sqcap \forall IND(i).(\forall non\_\underline{a}(i).viol(i)) \sqsubseteq \exists DEL(\underline{a}, i, j).\top \tag{6.8}$$

$$\begin{aligned} rea(i, r) \sqcap rea(j, s) \sqcap \forall IND(i).(\forall non\_\underline{a}(i).viol(i)) \sqsubseteq \forall DEL(\underline{a}, i, j).\\ \forall IND(i).(\forall non\_\underline{a}(j)).viol(j)) \end{aligned} \tag{6.9}$$

and if $R_{Contr}(r, s)$ *then* $\mathcal{T}(\mathfrak{COS})$ *contains the following axioms for any transition type form* $\underline{a}$ *and agents* $i, j \in Agents$*:*

$$\begin{aligned} rea(i, r) \sqcap rea(j, s) \sqcap \forall IND(i).(\forall non\_\underline{a}(j).viol(j)) \sqsubseteq \forall non\_\underline{a}(j).\\ (\forall IND(i)(\forall non\_\underline{a}(i).viol(i))) \end{aligned} \tag{6.10}$$

This translation uses as pre and post conditions of the agents' organizational activities the agents' knowledge about the relevant institutional states (e.g., directed obligations), rather than those states themselves. So, Formulae 6.8 and 6.9 concern the knowledge by the delegating agent that: 1) before delegating the task, it is under an obligation to perform it; 2) after delegating the task the recipient is under an obligation to perform it. Similarly, Formula 6.10 states in the antecedent that the controller knows the controlee is under an obligation, and in the consequent that the controller knows it is put under that obligation if the controlee fails.

	Roles in institutions	Roles in organizations
Status	YES	YES
Enactment/Deactment	YES	NO
Abstractness/Concreteness	YES	NO
Agents' Mental States	NO	YES
Executability of Actions	NO	YES

Table 6.1: Features of role specification in institutions and organizations

6.2.5 Organizations and institutions

The analysis exposed in this section opens up a direct connection between the institutional paradigm in MAS design and the organizational one. A direct consequence of interpreting links as sets of terminological axioms is that institution and organization specification can, in effect, partially overlap. Consider the specification of a concrete an institution $\mathfrak{CIns}$ (Definition 5.6) and denote with $\mathbb{C}(\mathfrak{CIns})$ the set of all subsumptions logically following from $\mathfrak{CIns}$. Similarly, denote with $\mathbb{C}(\mathcal{T}(\mathfrak{COs}))$ the set of all subsumptions logically following from the translation of $\mathfrak{COs}$. It might well be the case that $\mathbb{C}(\mathfrak{CIns}) \cap \mathbb{C}(\mathcal{T}(\mathfrak{COs})) \neq \emptyset$, that is to say, they partially specify the same system.

This point of contact is of particular interest with respect to the notion of role. From the structural point of view a role is just a position in a structure, that is to say, a set of links. From the institutional perspective instead, it is a set of norms. These sets of two different 'things' partially overlap if studied in terms of the properties they express for transition systems. Roles as sets of norms state terminological axioms specifying how the role can be enacted, deacted, and what kind of status the agents acquire by enacting the role(see Section 5.2.3). Roles as sets of links state terminological axioms also specifying the status acquired by agents playing certain roles (Formulae 6.2 and 6.4) but, on the one hand, disregarding how that role can be enacted or deacted and, on the other hand, specifying the activities (e.g., delegation or information) that can be executed while enacting the role and, possibly, also their mental effects on the interacting agents.

There is therefore a central overlap in institutional and organizational role specifications, which consists in the specification of the status acquired by agents playing certain roles. Both institutions and organizations specify what an agent ought to, is permitted to, or has the right to do. On the other hand, there is an essential difference also in relation with the status specification. We have seen, in the previous chapter, that an essential feature of institutions is to connect abstract activities and state of affairs (i.e., transition and state types) to concrete ones. This is not the case in organizations where this kind of information can be difficultly represented in the form of links between roles. Organizational specifications are rather concerned with the activities that can be executed by playing a certain role (Formula 6.1), which is, as studied in Section 5.5, an infrastructural issue from which an institutional specification abstracts. A further important difference is that while institutional specifications state rules prescribing how a given role can be played by agents, structural specifications of organizations abstract from this aspect, which can also be difficultly represented in terms of links between roles. In a way, organization specifications really look at the system from the point of view of what can be done while playing a role, whereas institution specifications also consider how that point of view can be reached. To recap, organizations can be thought of as systems specifications lying in between institution specifications and infrastructure specifications.

There is also, however, one essential characteristic feature of organizational specifications, namely, the bringing agents' mental states (Formulae 6.5 and 6.6) into the picture. Institution specifications are forced to abstract therefrom since, from an institutional perspective, agents should be considered as black boxes. On the contrary, the types of agents interacting within an organization can be known in advanced by the designer. This is aligned with the common view within MAS research which sees institutions as particularly suitable for the design of open MAS, i.e., MAS where the agents interacting in the system are heterogeneous and created by different users ([Vázquez-Salceda, 2004]). The concern with abstract descriptions, typical of institution specifications, can instead be dropped in closed MAS, where a set of agents of a known type is taken as given, and it only awaits a suitable organization.

Table 6.1 recapitulates what just discussed. These considerations answer our fifth research question. We will briefly come back again to this issue in Section 6.5.4.

6.3 'Meaningful' links and structural properties

The formal specification of the transition systems semantics of organizational structures can guide MAS designers to recognize some desirable properties of the to-be-designed system.

6.3.1 Graph theory and organizations

Some graph-theoretical properties have a clear organizational sense, which manifests itself also without making the meaning of links formal. For example, in a power structure the existence of a unique source for the R_{Power} relation would express the so-called *unity of command principle* or, to use the terminology of [Morgenstern, 1951], the existence of a *highest competence*, that is, a sort of source of all tasks of the organization. Similarly, if the indegree of all the elements that are not the source is 1, then commands flow through the power structure according to what is usually called, in organization theory, an *unambiguous chains of commands*. If the indegree is higher than 1, then the delegation can give rise to redundancies and ambiguities, generating the possibility of authority conflicts, as formally described in [Friedell, 1967].

However, the formal semantics of structure is of definite help in the study of how the three dimensions of power, coordination and control interact, and how they should interact in view of some desirable design properties. The present section addresses this issue showing how a formally specified semantics of structure can aid the understanding of structural properties concerning the interaction of the power, coordination and control dimensions.

6.3.2 Some notation

Before getting started it is worth recollecting some standard graph theoretical notions which will be used in the rest of the chapter. An R_k-path (of length n) is a sequence $\langle x_1, ..., x_{n+1}\rangle$ of distinct elements of *Roles* s.t. $\forall i$ $1 \leq i \leq n$, $(x_i, x_{i+1}) \in R_k$. A R_k-semipath (of length n) is a sequence $\langle x_1, ..., x_{n+1}\rangle$ of distinct elements of *Roles* s.t. $\forall i$ $1 \leq i \leq n$, $(x_i, x_{i+1}) \in R_k$ or $(x_{i+1}, x_i) \in R_k$. A *source* in *Roles* is an element s s.t. $\forall d \in Roles$ with $d \neq s$ there exists a R_k-path from s to d. The *indegree* $id_k(d)$ of a point d in structure k is the number of elements d_1 s.t. $(d_1, d) \in R_k$. The *outdegree* $od_k(d)$ of a point d in structure k is the number of elements d_1 s.t. $(d, d_1) \in R_k$. We say a point d to be incident w.r.t. a k link if $1 \leq id_k(d)$, and it is said to have emanating k links if $1 \leq od_k(d)$.

6.3.3 Structural interplay

One essential structural design question is how the three structural dimensions of power, coordination and control should be arranged in order to guarantee the correct functioning of the organization.

Structure is necessary in order to reassign goals via new obligations (delegation). Because of this dynamics organizations need to distribute relevant knowledge (information), and implement forms of performance assessment and recovery (monitoring). Somehow, the interplay between these structural dimensions lies in the delegation activity and is therefore based on the power relation. This is in perfect accordance with many foundational investigations in the theory of organizations ([Selznick, 1948; Morgenstern, 1951]). In particular:

> "[. . .] delegation is the primordial organizational act, a precarious venture which requires the continuous elaboration of formal mechanisms of *coordination* and *control*" ([Selznick, 1948], p. 25).

This observation can be distilled in the two following principles: the structure of the organization should see to it that each agent is always aware of its obligations (a sort of "ought implies know" principle); the structure of the organization should see to it that the tasks allocated to the agents are always executed (a "successful performance" principle).

Intuitively, the implementation of the "ought implies know" principle can be met by aligning the coordination structure with the power structure, so that every delegation action can be followed by a corresponding information action. The successful performance can never be guaranteed as the agents are autonomous and subject to failure. The control structure cannot guarantee a full implementation of the "successful performance" principle. The control structure generates a new obligation for the controller each time the obligation has not been met by the controlled agent. In principle, also the controller can then violate this obligation leaving the goal not achieved. The more levels of controls are enacted, the stronger the principle can be thought of being implemented. This is again an instance of the "control of the controllers" problem, which we have already addressed by dealing with the issue of norm-enforcement within institutions (see Section 5.6). It is interesting to notice that the institutional side of the problem, and the structural one are just two faces of the same coin. In discussing Definition 6.3 we have in fact observed that the sanctioning issue and the control one exhibit identical formal patterns when expressed as state type subsumptions.

In short, if we want to guarantee to some extent that after a delegation through a power relation the delegated action is actually performed, we should take care that the agent that is obliged to perform the action also knows about this. This can only be ensured through a successful inform action to that agent through a coordination link that is either direct or indirect. If we also want to have a back-up in case of failure we should make sure that there is an agent monitoring the delegated action through a control link.

This leads us to formulate this soundness criterion for organizational structures.

Definition 6.7. (Sound $\mathfrak{Os}$)
A sound organizational structure is a tuple: $\langle Roles, R_{Pow}, R_{Coord}, R_{Contr} \rangle$ where Roles is the finite set of roles, and $R_{Pow}, R_{Coord}, R_{Contr}$ are three irreflexive binary relations on Roles such that $\forall r, s \in Roles$:

$$(r,s) \in R_{Pow} \Rightarrow \textit{there exists an } R_{Coord}\textit{–path from } r \textit{ to } s;$$
$$(r,s) \in R_{Pow} \Rightarrow \textit{there exists a role } t \in Roles \textit{ s.t. } R_{Contr}(t,s).$$

The occurrence of a power relation between role r and role s requires: the existence of a (finite) coordination path from r to s so that effective informative actions can transmit the relevant knowledge of agents enacting role r to agents enacting role s; and the existence of at least one element t (which, notice, might be r itself) which is in a control relation with s.

6.3.4 Why structural soundness?

By considering what the links are intended to mean, the question about how the different strctural dimension should interact can be given a more precise answer. This section shows what are the 'semantic' grounds for requiring soundness as formulated in Definition 6.7. We illustrate this point by means of a simple example which restates, in formal fashion, the considerations advanced in the previous section.

Example 6.1. *(Structural properties and the semantics of links) Consider the following simple $\mathfrak{COs} = \langle \{r,s,t\}, \{(r,s)\}, \{(r,s)\}, \{(t,s)\}, \{i,j,k\}, \{(r,i),(s,j),(t,k)\} \rangle$ (see Figure 6.1 for* *a graphical representation). And consider the translation $\mathcal{T}$ constrained according to Definitions 6.6, 6.4 and 6.5. The following subsumptions are logically implied by $\mathcal{T}(\mathfrak{CO}\mathfrak{s})$ for any $\underline{a}$:*

$$rea(i,r) \sqcap rea(j,s) \sqcap rea(k,t) \equiv \top \tag{6.11}$$

$$\begin{aligned}
DEL(\underline{a}, i, j) &\sqsubseteq \underline{a}(i) && (6.12)\\
\forall IND(i).(\forall non_\underline{a}(i).viol(i)) &\sqsubseteq \exists DEL(\underline{a}, i, j).\top && (6.13)\\
\forall IND(i).(\forall non_\underline{a}(j).viol(j)) &\sqsubseteq \exists INF(\forall non_\underline{a}(j).viol(j), i, j).\top && (6.14)\\
\forall IND(i).(\forall non_\underline{a}(i).viol(i)) &\sqsubseteq \exists(DEL(\underline{a}, i, j)\\
&\quad \circ INF(\forall non_\underline{a}(j).viol(j)), i, j)).\top && (6.15)
\end{aligned}$$

$$\begin{aligned}
\forall non_\underline{a}(j).viol(j) &\sqsubseteq \forall non_\underline{a}(j).(IND(k).\\
&\quad (\forall non_\underline{a}(k).viol(k))) && (6.16)\\
\forall IND(i).(\forall non_\underline{a}(j).viol(j)) &\sqsubseteq \forall INF(\forall non_\underline{a}(j).viol(j), i, j).\\
&\quad \forall IND(j).\underline{a}(j).viol(j)) && (6.17)\\
\forall IND(i).(\forall non_\underline{a}(i).viol(i)) &\sqsubseteq \forall DEL(\underline{a}, i, j).(\forall IND(i).\\
&\quad (\forall non_\underline{a}(j).viol(j))) && (6.18)\\
\forall IND(i).(\forall non_\underline{a}(i).viol(i)) &\sqsubseteq \forall(DEL(\underline{a}, i, j) \circ INF(\forall non_\underline{a}(j).\\
&\quad viol(j)), i, j)).(\forall IND(j).\\
&\quad (\forall non_\underline{a}(j).viol(j))) && (6.19)\\
\forall IND(i).(\forall non_\underline{a}(i).viol(i)) &\sqsubseteq \forall(DEL(\underline{a}, i, j) \circ INF(\forall non_\underline{a}(j).\\
&\quad viol(j)), i, j)).(\forall IND(j).\\
&\quad (\forall non_\underline{a}(j).viol(j)) \sqcap \forall non_\underline{a}(j).\\
&\quad (\forall IND(k).(\forall non_\underline{a}(k).viol(k))) && (6.20)
\end{aligned}$$

These subsumptions follow quite directly from Definitions 6.6, 6.4 and 6.5. Let us have a closer look at them. Formula 6.11 states the enactment configuration holding in the system. Formula 6.12 states that, for agent i, delegating $\underline{a}$ to j is a way of doing $\underline{a}$. Formulae 6.13-6.15 state executability laws holding given $\mathcal{T}(\mathfrak{CO}\mathfrak{s})$. The most interesting is Formula 6.15 which expresses that it is possible for agent i to first delegate $\underline{a}$ to agent j and then informing j that it has to perform $\underline{a}$. This results from the organizational structure at issue, where the pair (r,s) belongs to both the power and the coordination structure.

Formula 6.16 expresses that all states in which agent j ought to perform $\underline{a}$ are states where if j does not perform $\underline{a}$ then it is the case that agent k knows it ought to perform $\underline{a}$. It concerns the exercise of control by k on j. Formulae 6.17 and 6.18 deal with the effect of the information and, respectively, delegation actions under this structural configuration. If i knows that j is obliged to perform $\underline{a}$ then by informing j about this, j also knows that it is obliged to perform $\underline{a}$ (Formula 6.17), and i always knows that j is obliged to perform $\underline{a}$ after it delegates $\underline{a}$ to j (Formula 6.18).

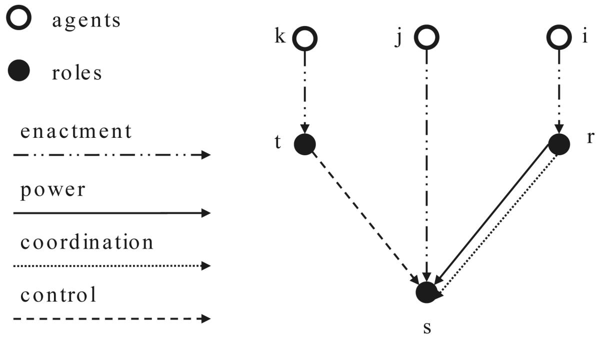

Figure 6.1: Pictorial representation of Example 6.1.

Finally, Formula 6.20 is what we were precisely looking for. It brings it all together stating that: if agent i ought to do $\underline{a}$, then by i delegating $\underline{a}$ to j and then informing j that it ought to perform $\underline{a}$, the system always reaches a state where j knows it ought to perform $\underline{a}$ and where if j does not comply with this obligation then agent k would know it ought to take over the execution of $\underline{a}$.

The $\mathfrak{COs}$ considered in Example 6.1 is a very simple instance of a sound organization. It is easy to see that soundness is what in fact guarantees that Formula 6.20 follows by applying the translation. In a way, Formula 6.20 shows in a formal way what the impact is of a certain structural property, in this case soundness, on the agents taking part in the organization.

The example considered is obviously a toy one, based on the simple semantics of the structural links presented above. However, it does convey the idea about the added value of addressing the transition system semantics of organizational structure by showing how it can provide a rigorous ground for the design of organizational structures themselves: if I want a property such as the one expressed in Formula 6.20 to hold in my system, what structural properties do I need?

Once the properties needed are identified, the question arises also about how much does a given organizational structure adhere to the chosen property. This type of questions is addressed in the next sections which make direct use of Definition 6.1.

6.4 Measuring Organizational Structure

This section presents some equations measuring specific graph-theoretical aspects of organizational structures[5].

6.4.1 Completeness, Connectedness, Economy

Completeness and connectedness of an $\mathfrak{Os}$ have to do with how strongly roles are linked with one another within one of the structural dimensions k. How much does the given structure approximate the structure where all directed links are present (*completeness*)? And how much is the given structure split in fragments (*connectedness*)?

$$\text{Completeness}_k(\mathfrak{Os}) = \frac{|R_k|}{|Roles_k| * (|Roles_k| - 1)} \tag{6.21}$$

$$\text{Connectedness}_k(\mathfrak{Os}) = 1 - \frac{|\text{DISCON}_k|}{|Roles_k| * (|Roles_k| - 1)} \tag{6.22}$$

with $|R_k| > 0$ and DISCON_k is the set of ordered pairs (x, y) of $Roles_k$ s.t. there is neither a R_k-semipath from x to y nor from y to x, i.e., the set of disconnected ordered pairs of the structural dimension $\langle Roles_k, R_k \rangle$. The condition $|R_k| > 0$ states that the structural dimension k does indeed exist. If the structure does not exist it cannot be measured. As a consequence, $\text{Completeness}_k > 0$. Stating that $\text{Completeness}_k(\mathfrak{Os}) = 0$ means thus that $R_k = \emptyset$ and hence that no structure at all is given. In practice, formula 6.21 measures the fraction of the actual links of the dimension $\langle Roles_k, R_k \rangle$ on all the available ones and formula 6.22 measures how 'not disconnected' that dimension is. With respect to connectedness, an important notion is that of cutpoint or, in an organizational reading, *liason role* [Harary, 1959a], i.e., a role whose removal decreases the connectedness of the structure.

The *economy* of a given $\mathfrak{Os}$ expresses a kind of balance between the two concerns of keeping the structure connected and of minimizing the number of links, i.e., minimizing completeness:

$$\text{Economy}_k(\mathfrak{Os}) = 1 - \frac{|R_k| - (|Roles_k| - 1)}{|Roles_k| * (|Roles_k| - 1) - (|Roles_k| - 1)} \tag{6.23}$$

with $|R_k| > 0$. The equation is based on the intuition according to which the most 'economical' digraph of n points consists of $n-1$ links, i.e., the minimum number of links which is still sufficient to keep the digraph connected. Indeed, the nominator of the fraction, consists of the number of links in the structural dimension k which are in excess or in defect w.r.t. the optimum of $n-1$ links. The denominator denotes instead the absolute number of links in excess in k. If $|R_k| = n - 1$ then the value of

[5] Equations 6.22, 6.23 and 6.24 below are an adaptation of equations presented in [Krackhardt, 1994]. The other equations are our contribution.

$\mathrm{Economy}_k(\mathfrak{Os})$ is optimal, i.e., equal to 1. The equation measures, therefore, how much k is 'not expensive' in terms of links. Notice that $\mathrm{Economy}_k(\mathfrak{Os}) = 1$ does not imply $\mathrm{Connectedness}_k(\mathfrak{Os}) = 1$, it does only imply that there are enough links in R_k for it to be possibly connected. If the existence of symmetric links in R_k is assumed, then $n - 1$ links are clearly not enough any more for guaranteeing connectedness. On the other hand, notice also that $\mathrm{Economy}_k(\mathfrak{Os})$ can assume a value greater than 1. That indicates a sort of '*over-efficiency*' of k. In this case, it is easy to see that, if $\mathrm{Economy}_k(\mathfrak{Os}) > 1$ then $\mathrm{Connectedness}_k(\mathfrak{Os}) < 1$. In other words, if the economy measures of $\mathfrak{Os}$ is lower than the optimal value 1, then $\mathfrak{Os}$ has more links than the ones necessary for $\mathfrak{Os}$ to be connected. If economy is instead higher than the optimal value 1, than there are in $\mathfrak{Os}$ too few links for it to be connected.

6.4.2 Unilaterality, Univocity, Flatness

The properties of unilaterality and univocity express the tendency of an $\mathfrak{Os}$ to display, respectively, an orientation in its links (*unilaterality*), and the absence of redundant links ending up in the same role (*univocity*). Do the links of an $\mathfrak{Os}$ always have a 'direction' or does the $\mathfrak{Os}$ allow, so to say, 'peer-to-peer' connections? And how many of those connections are such that no role has more than one incident link of the same structural dimension?

$$\mathrm{Unilaterality}_k(\mathfrak{Os}) = 1 - \frac{|\mathrm{SIM}_k|}{|R_k|} \tag{6.24}$$

$$\mathrm{Univocity}_k(\mathfrak{Os}) = \frac{|\mathrm{IN}_k|}{|Roles_k|} \tag{6.25}$$

$$\mathrm{Flatness}_k(\mathfrak{Os}) = 1 - \frac{|\mathrm{CUT}_k|}{|Roles_k|} \tag{6.26}$$

with $|R_k| > 0$ and SIM_k denotes the set of links (x, y) in R_k s.t. (y, x) is also in R_k, i.e., $|\mathrm{SIM}_k|$ is twice the number of symmetric links in k; IN_k denotes the set of roles x in $Roles_k$ s.t. $id_k(x) = 1$ or $id_k(x) = 0$, i.e., the set of roles which either have indegree equal to 1 in k or they are a source of k or of some subgraphs of k; and CUT_k denotes the set of roles x s.t. $od_k(x) \leq 1$ and $id_k(x) \leq 1$, that is to say, the set of roles which are at the same time addresser and addressee of k links. Intuitively, equation 6.24 measures how much asymmetry is present in k, while equation 6.25 measures how much a dimension k is univocal or "non ambiguous". The most univocal structures are assumed to be either the ones in which every point, except the source, has one and only one incident link (like in trees), or the ones in which exactly all points have only one incident link (like in cycles). Finally, equation 6.26 measures the relative amount of points in dimension k which are not intermediate point in a k-path, in other words the amount of points the removal of which would not determine a cut in any k-path. Obviously, the lowest value of flatness is provided by cycles.

Intuitively, unilaterality has to do with the level of subordination present in a structure. Consider the R_{Coord} dimension. The higher the number of unilaterality, the lower the amount of 'peer-to-peer' information exchange within $\mathfrak{Os}$. Univocity has to do with the level of conflict and redundancies of a given structure. Consider the R_{Pow} dimension. The higher the level of univocity, the more unambiguous is the chain of commands, as well as the more fragile once a link happens to be removed. See also [Friedell, 1967] for similar investigations on this issue. Flatness instead, has to do with the length of paths available within a given structure. We will see in Section 6.5 that long paths of the control dimension can be useful in order to implement levels of control on the controller roles themselves.

6.4.3 Detour, Overlap, Cover and Chain

The properties we address in this section do not concern structural dimensions taken in isolation, like the one just investigated, but instead how the different dimensions of an $\mathfrak{Os}$ interact with one another [6].

The properties we call *detour* and *overlap* regard the degree to which a structural dimension j 'follows' a structural dimension k, meaning by this the degree to which j establishes corresponding paths for each link of k, so that the roles that are related by R_k links are the same as those that are related by R_j-paths.

$$\mathrm{Detour}_{jk}(\mathfrak{Os}) = \frac{|\mathrm{PATH}_{jk}|}{|R_k|} \tag{6.27}$$

with $|R_k| > 0$ and the set PATH_{jk} is defined as the set of ordered pairs (x, y) s.t. $(x, y) \in R_k$ and there exists a R_j-path from x to y. Equation 6.27 measures the relative amount of R_j-paths between the elements of $Roles_k$ which have the same direction of the links in R_k. A special case of detour is the overlap. In fact, to measure how much does a dimension j overlap with a dimension k, it suffices to define a set LINK_{jk} corresponding to a PATH_{jk} where the R_j-paths are of length 1, i.e., simple links, and hence: $\mathrm{LINK}_{jk} \equiv R_k \cap R_j$. A set LINK_{jk} consists then of all the pairs (x, y) which are in R_k and in R_j, that is to say, of all x, y which are linked in R_k and in R_j.

$$\mathrm{Overlap}_{jk}(\mathfrak{Os}) = \frac{|\mathrm{LINK}_{jk}|}{|R_k|} \tag{6.28}$$

with $|R_k| > 0$. Intuitively, the more j-pairs correspond to k-pairs, the more j overlaps k in $\mathfrak{Os}$.

The property we call *in-cover* concerns the extent to which all the incident roles of k are also incident roles of a dimension j. In other words, we say that a dimension j *in-covers* a dimension k if all the roles which are addressees of a k link, are also addressees of a j link.

$$\mathrm{InCover}_{jk}(\mathfrak{Os}) = \frac{|\mathrm{IN}_j^+ \cap \mathrm{IN}_k^+|}{|\mathrm{IN}_k^+|} \tag{6.29}$$

[6] In mathematical sociology the study of the interaction of a number of different social relations within a social structure has often be indicated as crucial [Harary, 1959a]:

> "An actual group of people generally has more than one relation simultaneously operating. [. . .] The study of the influence of various relations on each other is in its infancy. However, this appears to be an extremely important field of endeavor" ([Harary, 1959a], p. 402).

Such issue has remained —to our knowledge— hardly investigated.

with $|R_k| > 0$ and the set IN_i^+ is defined as the set of all elements x in $Roles_i$ such that $1 \leq id_i(x)$. The equation describes then how many of the incident roles of k are also incident roles in j.

The usefulness of these measures for capturing aspects of the structural interplay can already be shown in relation with Definition 6.7. Readers might have noticed that, via the equations just exposed, it is possible to provide a quantification of the degree to which a given $\mathfrak{Os}$ adheres to the soundness principle concerning the interplay of the three dimensions of power, coordination and control. In fact, if we have $\mathrm{Detour}_{Coord-Power}(\mathfrak{Os}) = 1$ and $\mathrm{InCover}_{Contr-Pow}(\mathfrak{Os}) = 1$ then, following Definition 6.7, $\mathfrak{Os}$ is sound. Lower degrees of these measures would thus determine lower adherence to the soundness principle. Notice also that maximum soundness is trivially obtained via an overlap of both coordination and control structures on the power structure: that is to say, if $\mathrm{Overlap}_{Coord-Power}(\mathfrak{Os}) = 1$ and $\mathrm{Overlap}_{Contr-Power}(\mathfrak{Os}) = 1$, then $\mathfrak{Os}$ is (maximally) sound.

Equation 6.29 can be easily modified in order to capture analogous properties which we call *out-cover* and *chain*. The first one concerns the extent to which all the roles with emanating links in a dimension k are also roles with emanating links in a dimension j. The second one concerns the extent to which a dimension j is 'incident' to the emanating links in a dimension k, in the sense that the roles with incident links in j contain the roles with emanating links in k.

$$\mathrm{OutCover}_{jk}(\mathfrak{Os}) = \frac{|\mathrm{OUT}_j^+ \cap \mathrm{OUT}_k^+|}{|\mathrm{OUT}_k^+|}, \tag{6.30}$$

$$\mathrm{Chain}_{jk}(\mathfrak{Os}) = \frac{|\mathrm{IN}_j^+ \cap \mathrm{OUT}_k^+|}{|\mathrm{OUT}_k^+|}, \tag{6.31}$$

with $|R_k| > 0$, IN_i^+ is as defined above and OUT_i^+ is the set of all elements x in $Roles_i$ such that $1 \leq od_i(x)$. Notice that the chain measure can be viewed as an inter-structural version of the flatness measure.

Before ending the section, it is worth noticing that all structural measures defined above range between 0 and 1 except economy which can get values higher than 1. Despite this, we saw that the optimal value of $\mathrm{Economy}_k(\mathfrak{Os})$ is still 1 (higher values determine over-efficiency). Whether a given $\mathfrak{Os}$ enjoys a property at its optimal level, can therefore be handled as a matter of approximation of the corresponding measure to 1. For example, the more $\mathrm{Economy}_k(\mathfrak{Os})$ approximates value 1 the more $\mathfrak{Os}$ enjoys economy.

6.4.4 An example

In order to illustrate the above measures, an example is here provided and discussed. Consider the $\mathfrak{Os}$ depicted in Figure 6.2. It is specified as follows:
$Roles = \{a, b, c, d, e, f, g, h\}$,

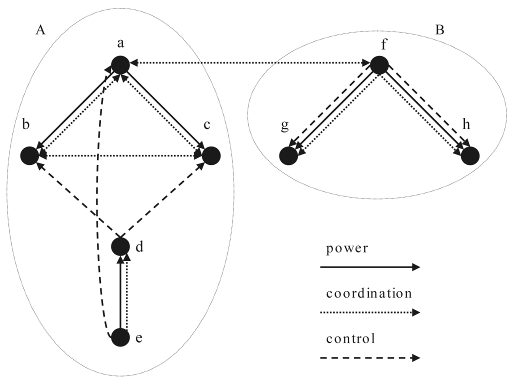

Figure 6.2: Example of organizational structure.

$R_{Pow} = \{(a,b),(a,c),(e,d),(f,g),(f,h)\}$,
$R_{Coord} = \{(a,b),(a,c),(b,a),(c,a),(b,c),(c,b),(e,d),(f,g),(f,h)\}$,
$R_{Contr} = \{(d,b),\ (e,a),(d,c),(f,g),(f,h)\}$.
We then have that: $Roles_{Pow} = Roles_{Coord} = Roles_{Contr} = \{a,b,c,d,e,f,g,h\}$.

Such a $\mathfrak{Os}$ specifies an organization where two substructures A and B are connected via a symmetric coordination link. It is what we may call, following [Horling and Lesser, 2004], a form of *federation*.

Substructure B is a typical form of highly centralized hierarchy: all connections move from the source f to the subordinated roles g and h. Indeed, it exhibits the optimal level of *efficiency*, *unilaterality*, *univocity* and *flatness* (equal to 1) for all three structural dimensions. Completeness and connectedness are also the same for all three dimensions, respectively equal to $\frac{2}{6}$ and to 1. Besides, there is a full reciprocal *overlap* (equal to 1) of all the three dimensions which, as shown above in Section 6.4.3, implies the soundness of the structure.

Substructure A, instead, displays a slightly more complex pattern. It hides two disconnected power hierarchies composed by roles a, b and c and, respectively, roles d and e. In fact, we have that $\text{Completeness}_{Pow}(A) = \frac{3}{20}$ and $\text{Connectedness}_{Pow}(A) = \frac{7}{10}$. Besides, the coordination structure is much more complete than the power one ($\text{Completeness}_{Coord}(A) = \frac{7}{20}$). This is due to the full connection holding between roles a, b and c. As to the interplay of the different dimensions in A, it is easily seen that $\mathfrak{Os}$ is not maximally sound since $\text{InCover}_{Contr-Pow}(A) = \frac{2}{3}$. This is due to the fact that role d is not object of control although it is subordinated, in the power structure, to role e. In case e would delegate to d a task, a failure in accomplishing this task would not be recovered. This would definitely constitute a weak spot in an organization designed according to this structure. Interestingly, there is minimum overlap between R_{Contr} and R_{Pow}: $\text{Overlap}_{Contr-Pow}(A) = 0$. This embodies a sort of

	Pow	Coord	Contr
Compl.$_k(\mathfrak{Os})$	$\frac{5}{56}$	$\frac{11}{56}$	$\frac{5}{56}$
Conn.$_k(\mathfrak{Os})$	$\frac{1}{4}$	$\frac{31}{56}$	$\frac{26}{56}$
Econ.$_k(\mathfrak{Os})$	$\frac{51}{49}$	$\frac{45}{49}$	$\frac{51}{49}$
Unil.$_k(\mathfrak{Os})$	1	$\frac{3}{11}$	1
Univ.$_k\mathfrak{Os})$	1	$\frac{5}{8}$	1
Flat.$_k\mathfrak{Os})$	1	$\frac{1}{2}$	1

	Coord-Pow	Contr-Pow	Pow-Contr
Detour$_{jk}(\mathfrak{Os})$	1	$\frac{2}{5}$	$\frac{2}{5}$
Overlap$_{jk}(\mathfrak{Os})$	1	$\frac{2}{5}$	$\frac{2}{5}$
InCover$_{jk}(\mathfrak{Os})$	1	$\frac{4}{5}$	$\frac{4}{5}$
OutCover$_{jk}(\mathfrak{Os})$	1	$\frac{2}{3}$	$\frac{2}{3}$
Chain$_{jk}(\mathfrak{Os})$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

	Coord-Contr	Contr-Coord	Pow-Coord
Detour$_{jk}(\mathfrak{Os})$	$\frac{2}{5}$	$\frac{2}{9}$	$\frac{4}{9}$
Overlap$_{jk}(\mathfrak{Os})$	$\frac{2}{5}$	$\frac{2}{9}$	$\frac{5}{9}$
InCover$_{jk}(\mathfrak{Os})$	1	$\frac{5}{6}$	$\frac{5}{6}$
OutCover$_{jk}(\mathfrak{Os})$	$\frac{2}{3}$	$\frac{2}{5}$	$\frac{3}{5}$
Chain$_{jk}(\mathfrak{Os})$	$\frac{1}{3}$	$\frac{3}{5}$	$\frac{2}{5}$

Figure 6.3: Example of structural measures.

complete "*separation of concerns*" between the power and the control dimensions, in the sense that controller roles are never in a power position with respect to the controlled roles. This is obviously a sensible design requirement for preventing connivances between controllers and roles in power positions. On the other hand, OutCover$_{Pow-Contr}(A) = \frac{1}{2}$ and OutCover$_{Coord-Contr}(A) = \frac{1}{2}$ show that, although no role is at the same time in a power and in a control position w.r.t. the same roles, there are controllers in A (one out of two) which have the possibility to delegate tasks and communicate with other roles (role e). Worth noticing is also the following: Chain$_{Contr-Pow}(A) = \frac{1}{2}$, that is, one out of two roles in a power position are subjected to control. Interestingly, the only uncontrolled role in a power position is the controller role e itself, and in fact no control of the controller is implemented: Flatness$_{Contr}(A) = 1$.

After discussing the two substructures in isolation we focus now on the federation $\mathfrak{Os}$ emerging by the joining of the two substructures via a symmetric coordination link between roles a and f. The resulting structural measures of $\mathfrak{Os}$ are the one listed in figure 6.3. Let us comment upon them. First of all, none of the three dimensions is connected (with coordination being the most connected among the three). This means that within each dimension, there exist unrelated clusters of roles. In particular, the roles in a controlling position within substructure A cannot communicate with the rest of the federation. It follows that all dimensions happen to display high values of economy and even over-efficiency, like in the case of power and control. As to the degree of unilaterality and univocity, power and control enjoy a degree equal to 1, and they thus display typically hierarchical features. On the other hand, coordination is highly reciprocal except, as we have already noticed, within substructure B, and it maintains a high degree of univocity keeping therefore a low level of redundancies in coordination as well. As to the interplay between the different dimensions, $\mathfrak{Os}$ inherits the flaw of substructure A which prevents it from enjoying the maximum degree of InCover$_{Contr-Pow}$, jeopardizing soundness. Coordination, instead, fully overlaps power guaranteeing the necessary flow of communication along the paths of delegation.

6.5 Structural Evaluation of Agent Organizations

As the example showed, the structural measures captured in equations 6.21-6.25 and 6.27-6.31 would be already enough for a quantified comparison of organizational structures. What is still lacking, is to give those measures an interpretation, in terms of commonly used criteria such as robustness, flexibility and efficiency. In this section the metrics developed in the previous section are used to provide hints about the adherence of a given organizational structure to those criteria. Questions we aim at shedding light on are of the type: Is the coordination structure flexible (enough)? Is the power structure efficient (enough)? Is the interplay between power and control structure robust (enough)?

It should be stressed that we do not claim those notions to be understandable only on the basis of structural considerations. We rather address what, just by looking at the structure of an organization, can be said about its robustness, flexibility and efficiency. As a matter of fact, considerations about structure have always been relevant both in organizational sciences and multi-agent systems for explaining why, for instance, a network is more flexible than a hierarchy. Here, we try to provide just a more steady handle for considerations of this kind. The semantics of power coordination and control exposed in Section 6.1.1 and made formal in Section 6.2 is used as a background for guiding the ink between the graph-theoretical metrics and the criteria considered.

6.5.1 Robustness

To say it with [Stimson, 1996], "robustness is simply a measure of how stable the yield is in the face of anticipated risks. That is, the maintenance of some desired system characteristics despite fluctuations in the behavior of its component parts or its environment [. . .]. Adding robustness thus adds complexity" .

Robustness asks for redundancies in the structural dimensions used for dividing tasks within an organizations, i.e., the power and the coordination structures. Redundancy for a power structure means low values of the Univocity$_{Pow}$ measure, and for a coordination structure also a low degree of the Unilaterality$_{Coord}$ in order to allow for symmetric coordination links. In particular, symmetric coordination links can substitute broken power links allowing for bilateral negotiations of tasks to replace direct delegation. Therefore, a high Overlap$_{Coord-Pow}$ would be a sign of robustness. In addition, the coordination structure determines how well information can disseminate over the organization. So, one can easily claim that the more complete and more connected (Completeness$_{Coord}$ and Connectedness$_{Coord}$) is the coordination structure is the more robust is the organization.

For the same reasons the control structure plays an important role for the robustness of an organization allowing for failure detection and reaction. It can be required that each role in the power and coordination structures is controlled, suggesting a high degree of the following measures: Chain$_{Contr-Pow}$, i.e., the control of agents in power positions; Chain$_{Contr-Coord}$, i.e., the control of roles from which coordination links depart; InCover$_{Contr-Coord}$, i.e., the control of roles to which coordination links are directed. It is also sensible to require controlling roles to be in power positions in order to possibly delegate the tasks they might have to take over from the controlled agents. This aspect can be fostered by high values of OutCover$_{Pow-Contr}$, which also call for high values of OutCover$_{Pow-Coord}$ in order for the delegation to be followed by appropriate information.

Furthermore, every role in the control structure can be required to have a high in-degree (every role is monitored by many other roles), which corresponds to a low level of Univocity$_{Contr}$. The number of control levels can also be increased, so that as many controllers as possible are, in turn, controlled. This has to do with the well-known "control of the controllers" issue which we already touched upon in Section 6.3 and corresponds to a low degree of Flatness$_{Contr}$ (long control paths are enabled).

To recapitulate these considerations, the following measures can be sensibly considered to foster robustness:

Completeness$_{Coord}$	1	Overlap$_{Coord-Pow}$	1
Connectedness$_{Coord}$	1	Chain$_{Contr-Pow}$	1
Univocity$_{Pow}$	0	Chain$_{Contr-Coord}$	1
Unilaterality$_{Coord}$	0	InCover$_{Contr-Coord}$	1
Univocity$_{Contr}$	0	OutCover$_{Pow-Contr}$	1
Flatness$_{Contr}$	0	OutCover$_{Pow-Coord}$	1

For instance, the maximum enhancement of robustness obtainable via modification of the connectedness measure is yielded by value 1. In other words, the more Connectedness$_{Coord}$ approximates 1, the more the structure is robust. As to univocity, the optimal value for increasing robustness is instead 0.

Getting back to the organizational structure $\mathfrak{Os}$ discussed in the example in Section 6.4.4, we can easily note that robustness is not its forte. Nevertheless, it does score well in the robustness-related measures concerning the interaction between the three structures: $\text{OutCover}_{Pow-Contr}(\mathfrak{Os}) = \frac{2}{3}$, $\text{OutCover}_{Pow-Coord}(\mathfrak{Os}) = \frac{3}{5}$, $\text{Chain}_{Contr-Coord}(\mathfrak{Os}) = \frac{3}{5}$ and $\text{InCover}_{Contr-Coord}(\mathfrak{Os}) = \frac{5}{6}$.

6.5.2 Flexibility

We start again with a quote from organizational theory: "Flexible organizations are a looser co-operative association than classic hierarchical organizations. [. . .] Flexible organizations are continually in flux and are able to adapt in a flexible way to changing circumstances" [Schoemaker, 2003].

The redistribution of tasks within an organization can be achieved via delegation through the power structure. However, an articulated power structure hinders flexibility constraining the distribution of tasks to predisposed patterns. This suggests that, for enhancing flexibility at a structural level, low degrees of both $\text{Completeness}_{Pow}$ and $\text{Connectedness}_{Pow}$ are required.

The control structure plays also a role from the point of view of flexibility since it can function as a link between different parts of the power structure. Whenever an agent enacting a role in the power structure fails on a task, its controller should react and have the power to possibly redistribute the task. Structurally, this corresponds to high values of $\text{Chain}_{Contr-Pow}$ and $\text{OutCover}_{Pow-Contr}$.

Network organizations and teams, instead, where no power structure exists, are commonly indicated as the paradigmatic example of flexible organizations [Powell, 1990]. What becomes essential is therefore a coordination structure through which the knowledge, concerning which agent might be capable to handle the new task, flows within the whole organization. The more roles are connected through this structure the more likely the right agent can be found to perform a new task. Completeness and connectedness ($\text{Completeness}_{Coord}$ and $\text{Connectedness}_{Coord}$) are therefore also linked to the enhancement of flexibility.

To recapitulate our considerations, the following are the measures that result in fostering flexibility:

$\text{Completeness}_{Pow}$	0	$\text{Completeness}_{Coord}$	1
$\text{Connectedness}_{Pow}$	0	$\text{Connectedness}_{Coord}$	1
$\text{Chain}_{Contr-Pow}$	1	$\text{OutCover}_{Pow-Contr}$	1

With respect to the flexibility of the structure $\mathfrak{Os}$ in the example, we see that it has indeed a small power structure (connectedness and completeness are very low) and a reasonably connected coordination structure (= $\frac{31}{56}$). These two aspects both enhance flexibility. This is indeed what we would expect, being $\mathfrak{Os}$ a form of "federation", that is, a form of organization which retains some purely hierarchical aspects (in its substructures) but exhibiting better flexibility. It scores well also w.r.t. the OutCover measure between power and control: $\text{OutCover}_{Pow-Contr}(\mathfrak{Os}) = \frac{2}{3}$.

6.5.3 Efficiency

According to [Etzioni, 1964], efficiency mostly refers to the amount of resources used by the organization to perform its tasks. However, organizational structure plays a role in this sense, since "links are not without cost in a social system" [Krackhardt, 1994].

The existence of a power structure guarantees efficient distribution of tasks, and a tree is the most efficient structure to cover all roles. Such a structure is obtained imposing value 1 for all the following measures: $\text{Connectedness}_{Pow}$ (a disconnected power structure generates fragments with independent power), Economy_{Pow} (maximum economy without over-efficiency), $\text{Unilaterality}_{Pow}$ (no peer-to-peer connections) and Univocity_{Pow} (no conflicts in the chain of command).

As to coordination and control, economy (Economy) should also be required to be 1 for the simple reason that this would minimize the amount of links. In addition, the most efficient way in order to guarantee soundness (Definition 6.7) consists in mirroring the power dimension, therefore obtaining high levels for all measures of overlap, that is: Overlap w.r.t. the related dimensions of $Coord - Pow$, $Contr - Pow$, as well as $Pow - Coord$ and $Pow - Contr$ (overlap needs to hold in both directions in order to force coincidence). This keeps the number of links minimal and avoids the creation of further roles with mere coordination and control tasks. It follows that a fully hierarchical organization (such as substructure B described in the example of Section 6.4.4 where all structures follow the same pattern forms the most efficient organization possible, at least from a structural perspective.

These are the thus the measures we consider to maximise efficiency:

$\text{Connectedness}_{Pow}$	1	$\text{Unilaterality}_{Pow}$	1
Economy_{Pow}	1	Univocity_{Pow}	1
Economy_{Coord}	1	Economy_{Contr}	1
$\text{Overlap}_{Coord-Pow}$	1	$\text{Overlap}_{Contr-Pow}$	1
$\text{Overlap}_{Pow-Coord}$	1	$\text{Overlap}_{Pow-Contr}$	1

The structure $\mathfrak{Os}$ of the example in Section 6.4.4 incorporates a very efficient power structure: unilaterality and univocity are optimal (equal to 1) as well as the overlap between coordination and power. On the other hand, the power structure covers only a small fraction of the whole organization ($\text{Connectedness}_{Pow}(\mathfrak{Os}) = \frac{1}{4}$). As a consequence, distribution of tasks via delegation can only partially take place.

6.5.4 Metrics vs. meaning

Before closing the section, we deem worth making a final remark concerning the two faces, so to say, of the chapter. On the one hand we have addressed structural links from the point of view of their meaning in terms of transition systems. On the other hand, we have studied them from the point of view of their mere graph-theoretical features.

	Institutions	**Organizations**
Roles' Status	YES	YES
Enactment/Deactment of Roles	YES	NO
Abstractness/Concreteness	YES	NO
Agents' Mental States	NO	YES
Executability of Actions	NO	YES
Structural Criteria	NO	YES

Table 6.2: Features supported by institution and organization specifications

In Section 6.2.5, on the grounds of the transition systems semantics used for interpreting structural links, we have compared the organizational and the institutional views on role specification. Such comparison has highlighted similarities and differences between the organizational and institutional paradigms as they have been understood in this work. There is one final remark to be made on this issue, which completes our answer to the fifth research question. A key feature of organizational specifications resides in the possibility of studying them graph-theoretically. In fact, the design of a concrete MAS might aim at incorporating an organizational structure enjoying a desirable degree of some structural properties according to the metrics presented in this section. While institutional specifications are not transparent to considerations about the robustness, flexibility or efficiency of the system, organizational specifications are. At the same time, by relating structural links to state and transition type subsumptions as proposed in Section 6.2, it becomes possible to check how adding or removing a link, and thus changing the value of some structural property of an $\mathfrak{Os}$, modifies the system specification in terms of the corresponding terminological axioms. This can be easily done by comparing the TBox $\mathcal{T}(\mathfrak{Os})$ of the organization before the links are added or removed, with the TBox $\mathcal{T}(\mathfrak{Os}')$ after the modification of the structure. The possibility of formally accounting for this sort of design feedback is a direct result of the attribution of meaning to structural links.

Table 6.1 can now be extended to Table 6.2. Institution and organization specifications each display specific features addressing different aspects of social interaction. It is our claim that the design process of a MAS would greatly benefit from the realization, at a design phase, of those different features. In a sense, a designer should be aware of when it is most appropriate to think in institutional terms, and when in organizational ones. We hope that the results presented in this chapter can contribute to the development of such an awareness.

6.6 Related and Future Work

The present section discusses some related work and points to some relevant future research lines.

6.6.1 Organizational structure and responsibility

The concept of responsibility is central to a theory of collective agency and organizations. Responsibility issues arise any time a group of agents acts collectively in order to achieve certain objectives. Plans are made for the collective action of the group and specific agents are stated to be "responsible" for certain tasks. If something goes wrong certain agents might be found "responsible" for what happened, they might be held "accountable" and be "blamed". The notion of responsibility displays different nuances all related with particular aspects of collective agency and, predominantly, obligation and knowledge.

The way obligations and knowledge flow withing groups of agents is in turn related with the organizational structure those groups display. The possibility to delegate tasks to subordinated agents, or to successfully inform other agents about the actual state of the organization, or the possibility to put effective monitoring and recovery mechanisms in place are all aspects influencing the assessment of responsibilities within organizations. If an agent is appointed to perform a specific task, but it does not get the necessary knowledge for correctly performing it, can it be held responsible for a failure in the execution of the plan? And in what sense precisely is it responsible? Again, if an agent is appointed to a task but it delegates it to a subordinated agent, does the failure of the subordinated agent determines a form of responsibility for the first agent? And in what sense?

We have addressed these questions in [Grossi et al., 2004, 2006h] claiming that responsibility issues within groups of agents are essentially related with the way groups are organized in order to pursue their objectives. In a nutshell, the less a group of agents is organized, the more blurred becomes the assessment of responsibilities within the group. In that work, the notions of power coordination and control structure, with their formal meaning, are used in order to ground four different notions of responsibility within organized groups of agents.

6.6.2 Organizational structure in management sciences

The chapter has imported a number of techniques from mathematical sociology (in particular from [Harary et al., 1965; Friedell, 1967; Krackhardt, 1994]), as it is evident especially from Section 6.4.3. We want now to point at the work on organizational structure presented in [Malone, 1987; Malone and Smith, 1988] which has been developed in the area of management sciences and which shares a number of essential similarities with what was presented especially in Section 6.4.3.

In that work structures are also represented as graphs, although not as multigraphs: $\langle Roles, R\rangle$ where R is an areflexive relation on $Roles$. The interesting representational aspect of that work is that the elements of $Roles$ are labeled, e.g., some roles are called "product managers" (choosing which tasks should be performed by the organization), "task processors" (executing tasks), "functional managers" (choosing which processor should perform the tasks chosen by the manager). In other words, properties are assigned to roles which are independent from the structure R. This allows for more fine-grained distinctions between structures. Two organizational structures might in fact have exactly the same form, i.e., be isomorphic, but they might differ with respect to the functions assigned to each role in the structure. In that view, an organizational structure becomes exactly something like a Kripke model: $\langle Roles, R\rangle, \mathcal{I}$ where $\mathcal{I}$ is the evaluation assigning function labels to each role.

It becomes then natural, along this line, to view the notion of $\mathfrak{Os}$ presented here as a multi-modal Kripke frame, and then add an evaluation assigning functions to roles in the form of propositions. Organizational structures should then be more properly represented as Kripke models. Modal operators would, in this case, represent direct reachability of roles with some properties along one of the structural dimensions k. As an example, formula $\langle Power\rangle\, \phi$ would express that a role of type ϕ is reachable in one step along the power dimension. This would add some interesting further aspects to our analysis, and in particular, the exact characterization of the notion of position within a structure in terms of the modal logic notion of bisimulation (cf. [Kamps and Marx, 2002]): two roles occupy the same positions in two organizational structures iff they are bisimilar with respect to those organizational structures.

6.6.3 Fine-tuning structure and its meaning

The chapter has been built on the distinction between semantic aspects vs. graph-theoretical aspects of organizational structures. There is still much to pursue in both direction.

The obvious question to be addressed from the graph-theoretical point of view is, at this stage, whether organizations can be designed which maximize the adherence to all the three criteria touched upon in Section 6.5. From a structural point of view and as intuition suggests, it is easy to show that this is not possible. Consider, for instance, the coordination structure. In fact, efficiency increases when Economy_{Coord} approximates 1. Maximum robustness and flexibility both require Economy_{Coord} equal to 0, while maximum efficiency requires Economy_{Coord} equal to 1:

	Robust	Flexible	Efficient
Economy_{Coord}	0	0	1

Intuitively, both robustness and flexibility increase the number of structural links and thus the costs of the organizational overhead, while efficiency reduces these overhead costs. Similar problems exist, for instance, for the power structure. The robustness criterion requires as many redundancies as possible, and therefore low levels of univocity, while flexibility demands the structure to be as small as possible and therefore with very low degrees of completeness. A number of similar incompatibilities can be detected and mathematically investigated.

Since it is not possible to maximize the adherence to all properties at the same time, the point consists then in finding suitable compromise solutions. This issue corresponds to what in organization theory falls under the label "*synthesis problems*", that is, the questions concerning "which structures are best suited to solve optimally certain types of problems" [Harary, 1959b]. Should for instance flexibility be privileged over efficiency? In other words, choices should be made between the concurrent criteria. An extensive analysis of the interdependencies between equations 6.21-6.31 could provide useful insights on this type of issues.

From the point of view of the study of the meaning of structural links far deeper analysis of the one proposed in Section 6.2 can be provided. In particular, an interesting way to go would be to analyze links in terms of different possible semantics for the speech acts ([Searle, 1969]) occurring in the literature on agent communication languages (ACLs) along the lines proposed, for instance, in [Dignum, 2006]. Structural links would then correspond to the possibility of performing certain speech acts or engage in some more complex communication protocol with precise consequences. Interestingly, also the literature on ACL recognizes the importance of distinguishing between mentalistic and institutional, or broadly speaking normative, semantics for speech acts (see, for instance, [Agerri, 2007]).

6.7 Conclusions

The chapter has put forth some proposals in order to lay the ground for a formal theory of organizations based on the notion of organizational structure. Following foundational work on organization theory and in particular [Morgenstern, 1951], the research presented started by stressing that organizational structures consist of a number of different relations among roles and, consequently, that they are better represented as multidigraphs. Two lines have then been pursued.

The first line has shown how to interpret the links holding between different roles in a structure in terms of labeled transition systems (Section 6.2): the existence of a link always expresses something about the actions that the agent playing that role can perform, or ought to perform, and what are the consequences of those actions. This insight has also motivated a clear thesis about the relation between the institutional and the organizational paradigms in MAS design (Section 6.2.5), which is summarized in Table 6.2. To put it in a nutshell, organizations concern the executability of actions by agents playing certain roles and their effects on the system including their effects on the mental states of other interacting agents, while institutions concern the specification of the (institutional) effects of the (concrete) actions of the agents taking part in the system, and they abstract from both agents' mental states and the executability of actions. Finally, it has been shown that by committing to a formal interpretation of the links of an organizational structure it becomes possible to analytically motivate the reasonableness of certain graph-theoretical properties of structure, e.g. soundness (Definition 6.7).

The second line has addressed the issue of the influence of organizational structures on the performance of organizations, aiming at providing a quantitative method for analyzing, comparing and evaluating different types of structures. We proceeded as follows. First we provided a number of measures for quantifying the adherence of organizational structures to specific graph-theoretical features (Section 6.3). Second, these measures have been put in relation with the organizational criteria of robustness, efficiency and flexibility of an organization (Section 6.5).

Chapter 7

Conclusions

"For it is most true that Cicero saith of them [the philosophers] somewhere; that there can be nothing so absurd but may be found in the books of philosophers. And the reason is manifest. For there is not one of them that begins his ratiocination from the definitions or explications of the names they are to use; which is a method that hath been used only in geometry, whose conclusions have thereby been made indisputable".

T. Hobbes, "Leviathan", Ch. 5

Roughly speaking, the results presented in this thesis are of two types. They are summarized in Table 7.1.

Results of the first type concern the development of precise views of institutions and organizations as they are advanced in some informal philosophical and social literature. In particular, we have committed to and investigated the following central thesis. Institutions are normative systems and normative systems impose terminologies ([Pufendorf, 1688]), which are stated by means of constitutive rules ([Searle, 1969, 1995]). They define contexts, i.e., the sets of situations ([Stalnaker, 1998]) which make the rules of the normative system true. Contexts are defined on different languages ([Shoham, 1991]) and they can be ordered from more concrete to more abstract ones. By means of such ordering the abstractness of different normative systems can also be expressed. On this ground we provided an analysis of the notions of open-texture and counts-as, and we analyzed the difference between constitutive and regulative rules.

- The open-texture of normative (legal) terms is related with the abstractness of normative systems. A same term can get different and conflicting interpretations in different concrete variants of an abstract normative system.
- Counts-as statements are statements talking about institutions viewed as terminologies defining contexts. They can either express what logically follows

from a given terminology (classificatory counts-as); or what follows from a given terminology which does not hold in general (proper classificatory counts-as); or the axioms of the terminology (constitutive counts-as).

- Regulative rules can be thought of as a special type of constitutive ones, that is, they are those constitutive rules defining the concept of violation for a given institution. Constitutive and regulative rules obey the very same logic. However, the way they are internalized by agents is different: regulative norms motivate, while constitutive norms determine belief updates.

As to organizations, we have upheld the following view. Organizations consist of structures laid upon the set of roles of the organization ([Morgenstern, 1951]). These structures have a precise impact on the activities of the agents enacting the roles of the organizations, and they exhibit degrees of adherence to structural properties which are related to their performance.

- Links between roles can establish what activities can actually be performed by the agents. In this sense they concern infrastructural aspects of the MAS, that is, what kind of system transitions are possible and under what conditions.
- Links between roles can also specify what kind of effects can be determined by the performance of certain activities by the agents enacting the roles at issue. Some of such effects can be of an institutional nature, and they therefore depend on the being in force of sets of constitutive rules. Others can be of a mentalistic nature, e.g., stating that between two roles certain speech acts can successfully be performed which change the mental state of the recipient.
- The performance criteria of robustness, flexibility, and efficiency can be addressed from a structural point of view, highlighting how they are linked to formal properties of the organizational structure.

Results of the second type concern the study of formal tools to support with exact methods the analysis of the aforementioned positions. We have been working with Description and Modal logics, and Graph Theory.

- Description logics have been used to represent and study terminologies, and therefore institutions, as TBoxes. A contextual version of DL have been proposed in order to capture, within the same formalism, different TBoxes and reason about them. In that framework, contexts have been modeled as sets of DL models. This has led to a theory of contexts as elements of Boolean Algebras with Operators which naturally accounts for context operations and for the relations between abstract and concrete contexts. By contextualizing DL terminologies it was also possible to formally characterize open-texture concepts and distinguish them from arbitrary ones. Finally, DL has also been used to represent the effects of organizational structures on agents' activities laying a common formal ground, i.e., labeled transition systems, for the analysis of institutions as well as of organizations.

	First group	**Second group**
NOTIONS	*INFORMAL ANALYSIS*	*FORMAL ANALYSIS*
Institution	Terminolgy	TBox
Context	Defined by terminologies	Set of DL models
Abstractness	Relation between contexts	(fcs_i) $\preccurlyeq$-statement
Open-texture	Concepts with non-empty core	$\mathfrak{Core}_{\mathbb{M}}$-statement
Norm	(Contextual) Subsumption	$(\xi :)$ $\sqsubseteq$-statement
Constitutive	Terminological axiom	$\sqsubseteq$-statement in a TBox
Regulative	Terminological axiom about violation	$\sqsubseteq$-statement in a TBox
Counts-as	(Proper) Classificatory/constitutive	$\Rightarrow^{cl}_{c}$ / $\Rightarrow^{cl+}_{c}$ / $\Rightarrow^{co}_{c,\Gamma}$
Organization	Structure	Multigraph
Effects	Effect and executability laws	$\sqsubseteq$-statements
Properties	Structural properties	Graph-theoretical metrics

Table 7.1: Summary of main theses and results

- Modal logics have first been used in relation with the theory of contexts proposed. It has been shown that reasoning with context descriptions in contextual terminologies is an instance of a specific modal logic, namely propositional release logic. The analysis of counts-as conditionals have been conducted in normal modal logics, accounting for the non-standard features of those conditionals by means of the interaction of three normal operators.

- Graph Theory have been used in order to provide formal representations of organizational structures and, in addition, to provide a quantitative handle on the discussion of the properties that such structures can exhibit.

The results just summarized have provided a way to make the institutional and organizational "handcuffs" formal, and thus "visible" in some way. Throughout the whole work we have pointed to a number of research lines which, in our view, would take the results presented here one step further towards the development of formal techniques for the "design of invisible handcuffs" for MAS. It is worth recollecting them at this stage. They are, essentially, of three kinds.

First, the formal frameworks developed in Chapters 2, 3 and 4 should incorporate a dynamic element taking care of how contexts —and therefore normative systems— can be defined and evolve in time as a result of specific actions (e.g., legislative rulings as suggested in Section 4.8.6).

Secondly, the analysis of organizational structure proposed in Chapter 6 in terms of graph-theory could be pushed further in its logical foundations. As sketched in Section 6.6.2, modal logic techniques could be used for capturing key logical aspects of reasoning with social networks, thus developing a fully-fledged logic of organizational structures.

Thirdly, and probably more importantly, the issue of norm implementation addressed in Chapter 5 should now be tackled from a game-theoretic perspective in the light of implementation theory and mechanism design, as we have already stressed in Section 1.2.3. The point is to understand how a given set of rules can be implemented in a society of agents via appropriate mechanisms and, possibly, to evaluate the impact of different set of rules, implemented by different mechanisms, on one same society. This would finally endow a designer with tools to better understand the impact of different "invisible handcuffs" on the to-be-regulated society, thereby improving the quality of his/her legislative action.

All in all, aim of this thesis has been to deploy formal methods for analyzing the notions of institution and organization. The quote from [Hobbes, 1651], by which this concluding chapter was opened, displays quite an optimistic faith in the use of exact methods as a way of acquiring *indisputable* knowledge about some subject matter. It is rather our hope that the investigations presented have shown the possibility of *disputing* about theories of institutions and organizations in exact and rigorous ways.

Appendix A

Completeness of Logics $\mathbf{K45^{ij}_n}$, $\mathbf{Cxt^u}$ and $\mathbf{Cxt^{u,-}}$

"Was vernünftig ist, das ist wirklich; und was wirklich ist, das ist vernünftig."

"What is real is rational; what is rational is real."

G. W. F. Hegel, "Grundlinien der Philosophie des Rechts. Naturrecht und Staatswissenschaft", Vorrede

This appendix proves soundness —"what is rational is real" ([Hegel, 1821])— and completeness —"what is real is rational" ([Hegel, 1821])— of the logics introduced in Chapter 4 for the analysis of counts-as: $\mathbf{K45^{ij}_n}$, $\mathbf{Cxt^u}$ and $\mathbf{Cxt^{u,-}}$. We will make use of the canonical model technique.

A.1 Preliminaries

Some facts are provided which will be used in the rest of the appendix.

A.1.1 Logics $\mathbf{K45^{ij}_n}$ and $\mathbf{Cxt^u}$

Logics $\mathbf{K45^{ij}_n}$ and $\mathbf{Cxt^u}$ are normal modal logics, i.e., the axiomatization of every modality $[i]$ contains all tautologies of propositional calculus, axiom `K` and is closed under rules `MP` and `N`. A normal modal logic Λ is strongly complete w.r.t. a class $\mathfrak{F}$ of frames if for any set of formulae Φ and formula ϕ, if Φ semantically entails ϕ then ϕ is derivable from Φ in Λ: if $\Phi \models_{\mathfrak{F}} \phi$ then $\Phi \vdash_{\Lambda} \phi$[1]. The following result about

[1]It might be worth recalling that strong completeness generalizes weak completeness, where the set of formulae Φ is empty.

strong completeness is key for the proofs of the next sections. The reader is referred to [Blackburn et al., 2001] for the proof.

Proposition A.1. *(Redefining strong completeness)*
A normal modal logic Λ is strongly complete w.r.t. a class of frames $\mathfrak{F}$ iff every Λ-consistent set of formulae is satisfiable on some $\mathcal{F} \in \mathfrak{F}$, i.e., it has a model $\mathcal{M}$ built on a frame $\mathcal{F}$ in class $\mathfrak{F}$.

Some well-known definitions and general results of modal completeness theory for normal modal logics are now listed, which concern the canonical model construction. We refer the reader to [Blackburn et al., 2001] for further details.

Let us, first of all, recall some facts about maximal consistent sets. Let Λ be a multi-modal normal logic. A maximal Λ-consistent set of formulae in a multi-modal language $\mathcal{L}_n$ is a set Φ s.t.: (a) $\perp$ is not derivable in Λ from Φ (i.e., Λ-consistency of Φ); (b) every set properly including Φ is Λ-inconsistent. Every maximal Λ-consistent set Φ is such that: $\Lambda \subseteq \Phi$; Φ is closed under rule MP; for all formulae ϕ either $\phi \in \Phi$ or $\neg\phi \in \Phi$; for all formulae $\phi, \psi : \phi \vee \psi \in \Phi$ iff $\phi \in \Phi$ or $\psi \in \Phi$.

We can now report the notion of canonical model for a normal modal logic Λ.

Definition A.1. *(Canonical model for logic Λ)*
The canonical model $\mathcal{M}^\Lambda$ for a normal modal logic Λ in the multi-modal language $\mathcal{L}_n$ is the structure $\left\langle W^\Lambda, \{R_i^\Lambda\}_{1\leq i\leq n}, \mathcal{I}^\Lambda \right\rangle$ where:

1. *The set W^Λ is the set of all maximal Λ-consistent sets.*

2. *The canonical relations $\{R_i^\Lambda\}_{1\leq i\leq n}$ are defined as follows: for all $w, w' \in W^\Lambda$, if for all formulae ϕ, $\phi \in w'$ implies $\langle i \rangle \phi \in w$, then $wR_i^\Lambda w'$.*

3. *The canonical interpretation $\mathcal{I}^\Lambda$ is defined by $\mathcal{I}^\Lambda(p) = \{w \in W^\Lambda \mid p \in w\}$.*

We briefly recall three key lemmata of completeness theory for normal modal logics. For the proofs we refer the reader again to [Blackburn et al., 2001].

Lemma A.1. *(Existence Lemma)*
For any normal modal logic Λ and any state $w \in W^\Lambda$, it holds that: if $\langle i \rangle \phi \in w$ then there exists a state $w' \in W^\Lambda$ such that $wR_i^\Lambda w'$ and $\phi \in w'$.

Lemma A.2. *(Truth Lemma)*
For any normal modal logic Λ and any formula ϕ, it holds that: $\mathcal{M}^\Lambda, w \models \phi$ iff $\phi \in w$.

Finally, we will also make use of the notion of point-generated subframe. Given a frame $\mathcal{F} = \langle W, \{R_i\}_{1\leq i\leq n}\rangle$, a point-generated subframe $\mathcal{F}^w$ of $\mathcal{F}$ is a structure $\left\langle W^w, \{R_i^w\}_{1\leq i\leq n}\right\rangle$ such that: (a) W^w is the set of states $w' \in W$ such that there exists, for any R_i, a finite R_i-path from w to w'; (b) $R_i^w = R_i \cap (W^w \times W^w)$, i.e., each R_i^w is the restriction of R_i on W^w. We will refer to the set of states which are accessible from a state w via a relation R_i^w as the set $r_i^w(w)$. The following result is of interest.

Lemma A.3. *(Generated subframes preserve validity)*
Let $\mathfrak{F}$ be a class of frames and $g(\mathfrak{F})$ be the class of point-generated subframes of the frames in $\mathfrak{F}$. It holds that, for all formulae ϕ in language $\mathcal{L}_n$: $\mathfrak{F} \models \phi$ iff $g(\mathfrak{F}) \models \phi$.

Completeness of $\mathbf{K45_n^{ij}}$ and $\mathbf{Cxt^u}$ is proven in Section A.2 and, respectively, Section A.3.

A.1.2 Logic $\mathbf{Cxt^{u,-}}$

In contrast to $\mathbf{K45_n^{ij}}$ and $\mathbf{Cxt^u}$, logic $\mathbf{Cxt^{u,-}}$ is quite more than a normal modal logic. It is built on a language containing a set $\mathbb{N}$ of nominals ($\mathcal{L}_n^{u,-}$, see Section 4.5.1), its axiomatics contains rule `Name` (see Section 4.5.3), and its models state conditions on the possible valuations of one type of propositional variables in the language, i.e., the nominals (see Section 4.5.2).

Let us call modal logics with names the normal modal logics on a language $\mathcal{L}_n$ with nominals extended with rule `Name`, axioms `Most` and `Least` and the axioms of the universal modality $[u]$. In the case of modal logics with names, strong completeness is defined as follows. Let Λ be a modal logic with names. Logic Λ is strongly complete w.r.t. the class $\mathfrak{F}$ of frames if for any set of formulae Φ and formula ϕ, if Φ semantically entails ϕ *in all surjective models* (see Section 4.5.2 for the definition of surjective model) built on a frame in $\mathfrak{F}$ then ϕ is derivable from Φ in Λ: if $\Phi \models_{\mathfrak{F}} \phi$ then $\Phi \vdash_{\Lambda} \phi$. Proposition A.1 should now be restated for modal logics with names.

Proposition A.2. *(Redifining strong completeness for modal logics with names)*
A modal logic Λ with names is strongly complete w.r.t. the class of frames $\mathfrak{F}$ iff every Λ-consistent set Φ of formulae is satisfiable on some surjective model built on a frame in class $\mathfrak{F}$.

Proof. [$\Leftarrow$] From right to left we argue by contraposition. If Λ is not strongly complete w.r.t. the class $\mathfrak{F}$ then there exists a set of formulae $\Phi \cup \{\phi\}$ s.t. $\Phi \models_{\mathfrak{F}} \phi$ and $\Phi \nvdash_{\Lambda} \phi$. It follows that $\Phi \cup \{\neg\phi\}$ is Λ-consistent but not satisfiable on any surjective model built on a frame in class $\mathfrak{F}$. [$\Rightarrow$] From left to right we argue per absurdum. Let us assume that $\Phi \cup \{\neg\phi\}$ is Λ-consistent but not satisfiable in any surjective model built on a frame in class $\mathfrak{F}$. It follows that $\Phi \models_{\mathfrak{F}} \phi$ and hence $\Phi \cup \{\neg\phi\}$ is not Λ-consistent, which is impossible. □

Strong completeness of logic $\mathbf{Cxt^{u,-}}$ is dealt with in Section A.4. Observe already that we will need to introduce a new kind of canonical model and to prove new Truth and Existence Lemmata. What proven in Section A.4 relies on general results exposed in [Gargov and Goranko, 1993] and [Blackburn et al., 2001].

A.2 Soundness and completeness of $\mathbf{K45^{ij}_n}$

To facilitate readability we recollect the axiom schemata of $\mathbf{K45^{ij}_n}$:

(P)	all tautologies of propositional calculus
(K)	$[i](\phi_1 \rightarrow \phi_2) \rightarrow ([i]\phi_1 \rightarrow [i]\phi_2)$
(4^{ij})	$[i]\phi \rightarrow [j][i]\phi$
(5^{ij})	$\neg[i]\phi \rightarrow [j]\neg[i]\phi$
(Dual)	$\langle i \rangle \phi \leftrightarrow \neg[i]\neg\phi$
(MP)	IF $\vdash \phi_1$ AND $\vdash \phi_1 \rightarrow \phi_2$ THEN $\vdash \phi_2$
(N^i)	IF $\vdash \phi$ THEN $\vdash [i]\phi$

where i, j denote elements of the finite set of indexes C.

The proof of soundness is routinary. It is well-known that inference rules MP and N preserve validity on any class of frames[2]. Providing the soundness of $\mathbf{K45^{ij}_n}$ w.r.t. CXT frames boils than down to checking the validity of axioms 4^{ij} and 5^{ij}.

Theorem A.1. *(Soundness of* $\mathbf{K45^{ij}_n}$ *w.r.t.* CXT *frames)*
Logic $\mathbf{K45^{ij}_n}$ *is sound w.r.t. the class of* CXT *frames.*

Proof. The validity of 4^{ij} is proven showing that its contrapositive has no countermodel. Such countermodel $\mathcal{M}$ would contain a state w such that for a given formula ϕ, $\mathcal{M}, w \models \langle j \rangle \langle i \rangle \phi$ and $\mathcal{M}, w \models \neg \langle i \rangle \phi$. Hence, by the semantics, $\exists w' \in W_i$ s.t. $\mathcal{M}, w \models \phi$ and $\nexists w' \in W_i$ s.t. $\mathcal{M}, w \models \phi$, which is impossible. The validity of 5^{ij} is proven in the same way. Suppose there is a model $\mathcal{M}$ and a state w such that $\mathcal{M}, w \models \langle i \rangle \phi$ and $\mathcal{M}, w \models \neg[j] \langle i \rangle \phi$. Hence, by the semantics, $\exists w' \in W_i$ s.t. $\mathcal{M}, w \models \phi$ and $\nexists w' \in W_i$ s.t. $\mathcal{M}, w \models \phi$. □

As to completeness, the desired result is obtained in two steps.

1. First, via the canonical model, it is proven that logic $\mathbf{K45^{ij}_n}$ is complete with respect to the class of i-j transitive (if wR_iw' and $w'R_jw''$ then wR_jw''), and i-j euclidean (if wR_iw' and wR_jw'' then $w'R_jw''$) frames[3].
2. Second, it is proven that if $\mathfrak{TE}$ is the class of of i-j transitive and i-j euclidean frames, then for every $\phi \in \mathcal{L}_n$: $\mathfrak{TE} \models \phi$ iff CXT $\models \phi$.

Theorem A.2. *(Completeness of* $\mathbf{K45^{ij}_n}$*)*
Logic $\mathbf{K45^{ij}_n}$ *is strongly complete w.r.t. the class of i-j transitive and i-j euclidean frames.*

[2]See [Blackburn et al., 2001].

[3]In [Nayak, 1994], frames with this property are called, respectively, hyper-transitive and hyper-euclidean.

Proof. By Proposition A.1, given a $\mathbf{K45_n^{ij}}$-consistent set Φ of formulae, it suffices to find a model state pair $(\mathcal{M}, w)$ such that: (a) $\mathcal{M}, w \models \Phi$, and (b) the frame $\mathcal{F}$ on which $\mathcal{M}$ is based is i-j transitive and i-j euclidean. Let $\mathcal{M}^{\mathbf{K45_n^{ij}}} = \left\langle W^{\mathbf{K45_n^{ij}}}, \{R_i^{\mathbf{K45_n^{ij}}}\}_{i \in C}, \mathcal{I}^{\mathbf{K45_n^{ij}}} \right\rangle$ be the canonical model of logic $\mathbf{K45_n^{ij}}$, and let Φ^+ be any maximal consistent set in $W^{\mathbf{K45_n^{ij}}}$ extending Φ. By Lemma A.2 it follows that $\mathcal{M}^{\mathbf{K45_n^{ij}}}, \Phi^+ \models \Phi$, which proves (a). It remains to be proven that $\left\langle W^{\mathbf{K45_n^{ij}}}, \{R_i^{\mathbf{K45_n^{ij}}}\}_{i \in C} \right\rangle$ enjoys i-j transitivity (b.1) and i-j euclidicity (b.2). To prove (b.1) consider three states $w, w', w'' \in W^{\mathbf{K45_n^{ij}}}$ such that $wR_j^{\mathbf{K45_n^{ij}}}w'$ and $w'R_i^{\mathbf{K45_n^{ij}}}w''$. Suppose then that $\phi \in w''$. As $w'R_i^{\mathbf{K45_n^{ij}}}w''$ and $wR_j^{\mathbf{K45_n^{ij}}}w'$, it follows that $\langle i \rangle \phi \in w'$ and then that $\langle j \rangle \langle i \rangle \phi \in w$. Now, w is a maximal consistent set of logic $\mathbf{K45_n^{ij}}$, it therefore contains formula $\langle j \rangle \langle i \rangle \phi \rightarrow \langle i \rangle \phi$ (i.e., the contrapositive of axiom 4^{ij}), hence $\langle i \rangle \phi \in w$ and thus $wR_i^{\mathbf{K45_n^{ij}}}w''$ which completes the proof of (b.1). Analogously, to prove (b.2) consider three states $w, w', w'' \in W^{\mathbf{K45_n^{ij}}}$ such that $wR_j^{\mathbf{K45_n^{ij}}}w'$ and $wR_i^{\mathbf{K45_n^{ij}}}w''$. Suppose then that $\phi \in w''$. It follows that $\langle i \rangle \phi \in w$ and since w is a maximal consistent set of logic $\mathbf{K45_n^{ij}}$, it contains formula $\langle i \rangle \phi \rightarrow [j] \langle i \rangle \phi$ (i.e., axiom 5^{ij}) and hence $[j] \langle i \rangle \phi \in w$. From this and from $wR_i^{\mathbf{K45_n^{ij}}}w''$ it follows that $\langle i \rangle \phi \in w''$, that is to say, for any formula ϕ it is the case that: if $\phi \in w'$, then $\langle i \rangle \phi \in w''$. Now, by Definition A.1, this implies that $w'R_i^{\mathbf{K45_n^{ij}}}w''$, which proves (b.2). □

Lemma A.4. *(Semantic equivalence for* Cxt *frames)*
Consider the class $\mathfrak{TE}$ *of i-j transitive and i-j euclidean frames. For every* $\phi \in \mathcal{L}_n$, $\mathfrak{TE} \models \phi$ *iff* Cxt $\models \phi$. *That is,* Cxt *frames and* $\mathfrak{TE}$ *frames define the same logic.*

Proof. [$\Leftarrow$] From right to left: for every ϕ, Cxt $\models \phi$ implies $\mathfrak{TE} \models \phi$. The proof is obtained by showing that a Cxt frame is always i-j transitive and i-j euclidean. By proposition 4.1, for all $w, w' \in W$, $w' \in W_i$ iff wR_iw'. To prove i-j transitivity, suppose that wR_iw' ($w' \in W_i$) and $w'R_jw''$ ($w'' \in W_j$). It follows therefore that wR_jw''. The proof of i-j euclidicity is perfectly analogous. Suppose that wR_iw' ($w' \in W_i$) and wR_jw'' ($w'' \in W_j$), hence $w'R_jw''$. [$\Rightarrow$] From left to right: for every ϕ, $\mathfrak{TE} \models \phi$ implies Cxt $\models \phi$. In this case, the proof is obtained by showing that every i-j transitive and i-j euclidean frame, which is also point-generated, is a context frame. By Lemma A.3, it holds that for every ϕ, $\mathfrak{TE} \models \phi$ iff $g(\mathfrak{TE}) \models \phi$. Now, let $\mathcal{F}^w$ be any frame in $g(\mathfrak{TE})$ generated by some state w. In order to prove the desired result, it suffices to show that every i-j transitive and i-j euclidean frame $\mathcal{F}^w$ generated by state w is a Cxt frame. By Proposition 4.1, this is proven by showing that for every $R_i^w \in \{R_i^w\}_{i \in C}$, $w'R_i^w w''$ iff $w'' \in r_i^w(w)$. This amounts to prove that for every w', w'' if there exists an R_i-path from w to w' and from w to w'', then $w'R_iw''$ iff $w'' \in r_i(w)$. From left to right, if there exists an R_i-path from w to w' and $w'R_iw''$, then by transitivity (which is a special case of i-j transitivity) wR_iw'', that is, $w'' \in r_i(w)$. From right to left, if there exists an R_i-path from w to w' and $w'' \in r_i(w)$, then wR_iw'' and hence, by euclidicity, $w'R_iw''$. □

Corollary A.1. *(Completeness of* $\mathbf{K45^{ij}_n}$ *w.r.t.* Cxt *frames)*
Logic $\mathbf{K45^{ij}_n}$ *is strongly complete w.r.t. the class of* Cxt *frames.*

Proof. Follows directly from Theorem A.2 and Lemma A.4. □

A.3 Soundness and completeness of $\mathbf{Cxt^u}$

We recollect the axiomatization of logic $\mathbf{Cxt^u}$:

(P)	all tautologies of propositional calculus
($\mathtt{K}^i$)	$[i](\phi_1 \to \phi_2) \to ([i]\phi_1 \to [i]\phi_2)$
(4^{ij})	$[i]\phi \to [j][i]\phi$
(5^{ij})	$\neg[i]\phi \to [j]\neg[i]\phi$
($\mathtt{T}^u$)	$[u]\phi \to \phi$
($\subseteq .ui$)	$[u]\phi \to [i]\phi$
(Dual)	$\langle i \rangle \phi \leftrightarrow \neg[i]\neg\phi$
(MP)	If $\vdash \phi_1$ and $\vdash \phi_1 \to \phi_2$ then $\vdash \phi_2$
($\mathtt{N}^i$)	If $\vdash \phi$ then $\vdash [i]\phi$

where i, j denote elements of the set of indexes C and u denotes the universal context index in C.

On the grounds of the results of the previous section, the proof of soundness and completeness of $\mathbf{Cxt^u}$ w.r.t. $\text{Cxt}^\top$ can be easily obtained. Soundness boils down to prove that axioms $\mathtt{T}^u$ and $\subseteq .ui$ are valid in $\mathbf{Cxt^u}$ frames.

Theorem A.3. *(Soundness of* $\mathbf{Cxt^u}$ *w.r.t.*$\text{Cxt}^\top$ *frames)*
Logic $\mathbf{Cxt^u}$ *is sound w.r.t. the class of* $\text{Cxt}^\top$ *frames.*

Proof. Trivial, given the interpretation of the $[u]$-operator as universal quantification on all the states in the domain W of the frame. □

Let $\mathfrak{TE}^\sim$ be the class of frames satisfying the following properties: they are i-j transitive, i-j euclidean; they contain an equivalence relation R_u such that for all $i \in C$, $R_i \subseteq R_u$. Again, completeness w.r.t. the relevant class of frames is proven in two steps.

1. Logic $\mathbf{Cxt^u}$ is first proven to be complete w.r.t. the class of $\mathfrak{TE}^\sim$ frames.

2. It is then proven that for any formula ϕ on $\mathcal{L}_n$: $\mathfrak{TE}^\sim \models \phi$ iff $\text{Cxt}^\top \models \phi$.

Theorem A.4. *(Completeness of* $\mathbf{Cxt^u}$*)*
Logic $\mathbf{Cxt^u}$ *is strongly complete w.r.t. the class* $\mathfrak{TE}^\sim$ *frames.*

Proof. By Proposition A.1, given a $\mathbf{Cxt^u}$-consistent set Φ of formulae, it suffices to find a model state pair $(\mathcal{M}, w)$ such that: (a) $\mathcal{M}, w \models \Phi$, and (b) the frame on which $\mathcal{M}$ is based is i-j transitive and i-j euclidean and contains a universal relation. Claim (a) is proven by making use of Lemma A.2. It remains to be proven that the frame $\left\langle W^{\mathbf{Cxt^u}}, \{R_i^{\mathbf{Cxt^u}}\}_{i\in C}\right\rangle$ of the canonical model enjoys i-j transitivity and i-j euclidicity (b.1) and that there exists a relation $R_u^{\mathbf{Cxt^u}}$ for $u \in C$ such that $R_u^{\mathbf{Cxt^u}}$ is an equivalence relation (b.2) and for every $i \in C$, $R_i^{\mathbf{Cxt^u}} \subseteq R_u^{\mathbf{Cxt^u}}$ (b.3). Claim (b.1) follows from Theorem A.2 since $\mathbf{Cxt^u}$ extends $\mathbf{K45_n^{ij}}$. As to (b.2), it follows from (b.1) that each $R_i^{\mathbf{Cxt^u}}$ is transitive and euclidean and, therefore, so is $R_u^{\mathbf{Cxt^u}}$. The proof of the reflexivity of $R_i^{\mathbf{Cxt^u}}$ is then routinary. Finally, claim (b.3) needs to be proven. Consider two states $w, w' \in W^{\mathbf{Cxt^u}}$ such that $wR_i^{\mathbf{Cxt^u}}w'$. Suppose then that $\phi \in w'$. It follows that $\langle i \rangle \phi \in w$. Since w is a maximal $\mathbf{Cxt^u}$-consistent set, it contains formula $\langle i \rangle \phi \rightarrow \langle u \rangle \phi$ (i.e., the contrapositive of axiom $\subseteq .ui$) and therefore $\langle u \rangle \phi \in w$. Hence, by Definition A.1, $wR_u^{\mathbf{Cxt^u}}w'$. □

Lemma A.5. *(Semantic equivalence for* $\text{Cxt}^{\top}$ *frames)*
For any formula ϕ on $\mathcal{L}_n$: $\mathfrak{TE}^{\sim} \models \phi$ iff $\text{Cxt}^{\top} \models \phi$. That is, $\text{Cxt}^{\top}$ frames and $\mathfrak{TE}^{\sim}$ frames define the same logic.

Proof. The proof is analogous to the proof of Lemma A.4. [$\Leftarrow$] The direction from right to left (for every ϕ, $\text{Cxt}^{\top} \models \phi$ implies $\mathfrak{TE}^{\sim} \models \phi$) is straightforwardly proven by observing that every $\text{Cxt}^{\top}$ frame represents a frame containing a universal relation R_u. In fact, a relation R_u is universal iff it holds that: for any $w, w' \in W$, wR_uw' iff $w' \in W$ (notice that this is a special case of Proposition 4.1). But every universal relation is an equivalence relation, which also includes all R_i's for any $i \in C$. That all $\text{Cxt}^{\top}$ frames are i-j transitive and i-j euclidean follows from Lemma A.4. This completes the proof of the right-to-left direction. [$\Rightarrow$] From left to right: for every ϕ, $\mathfrak{TE}^{\sim} \models \phi$ implies $\text{Cxt}^{\top} \models \phi$. Lemma A.4 has proven that every i-j transitive and i-j euclidean frame generated by state w is a Cxt frame. Consider now the relation R_u^w of the point-generated subframe $\mathcal{F}^w$ of a frame $\mathcal{F} \in \mathfrak{TE}^{\sim}$ containing an equivalence relation R_u such that for all $i \in C$, $R_i \subseteq R_u$. To obtain the desired result —via Lemma A.3— it suffices to show that the relation R_u^w is universal on W^w, which is trivial. □

Corollary A.2. *(Completeness of* **Cxt^u** *w.r.t.* $\text{Cxt}^{\top}$ *frames)*
Logic $\mathbf{Cxt^u}$ *is strongly complete w.r.t. the class of* $\text{Cxt}^{\top}$ *frames.*

Proof. Follows directly from Theorem A.4 and Lemma A.5. □

A.4 Soundness and completeness of $\mathbf{Cxt^{u,-}}$

We recollect the axiomatization of logic $\mathbf{Cxt^{u,-}}$:

(P)	all tautologies of propositional calculus
($\mathtt{K}^i$)	$[i](\phi_1 \to \phi_2) \to ([i]\phi_1 \to [i]\phi_2)$
(4^{ij})	$[i]\phi \to [j][i]\phi$
(5^{ij})	$\neg[i]\phi \to [j]\neg[i]\phi$
($\mathtt{T}^u$)	$[u]\phi \to \phi$
($\subseteq .ui$)	$[u]\phi \to [i]\phi$
(Least)	$\langle u \rangle\, \nu$
(Most)	$\langle u \rangle\, (\nu \wedge \phi) \to [u](\nu \to \phi)$
(Covering)	$[c]\phi \wedge [-c]\phi \to [u]\phi$
(Packing)	$\langle -c \rangle\, \nu \to \neg\, \langle c \rangle\, \nu$
(Dual)	$\langle i \rangle\, \phi \leftrightarrow \neg[i]\neg\phi$
(Name)	IF $\vdash \nu \to \theta$ THEN $\vdash \theta$, for ν not occurring in θ
(MP)	IF $\vdash \phi_1$ AND $\vdash \phi_1 \to \phi_2$ THEN $\vdash \phi_2$
($\mathtt{N}^i$)	IF $\vdash \phi$ THEN $\vdash [i]\phi$

where i, j are metavariables for the elements of C, c denotes elements of the set of atomic context indexes K; ν ranges over the set $\mathbb{N}$ of nominals, and θ in rule Name denotes a formula in which the nominal denoted by ν does not occur.

Theorem A.5. *(Soundness of* $\mathbf{Cxt^{u,-}}$ *w.r.t.* $\textsc{Cxt}^{\top,\backslash}$ *frames)*
Logic $\mathbf{Cxt^u}$ *is sound w.r.t. the class of* $\textsc{Cxt}^{\top,\backslash}$ *frames.*

Proof. It suffices to show that axioms Covering and Packing are valid in $\textsc{Cxt}^{\top,\backslash}$ frames by just noticing that in $\textsc{Cxt}^{\top,\backslash}$ frames, for any atomic context index c, family $\{W_c, W_{-c}\}$ is a bipartition of the domain W: $W \subseteq W_c \cup W_{-c}$, i.e., family $\{W_c, W_{-c}\}$ is a covering of W; and $W_c \cap W_{-c} = \emptyset$, i.e., $\{W_c, W_{-c}\}$ is a packing of W. □

Let $\mathfrak{IE}^{\top,\backslash}$ be the class of frames satisfying the following properties: they are i-j transitive, i-j euclidean; they contain a universal relation R_u; the set of relations $\{R_i\}_{i\in C}$ is such that, for any atomic context index c and states $w, w' \in W$: wR_uw' implies wR_cw' or $wR_{-c}w'$; and wR_cw' implies not $wR_{-c}w'$. Again, completeness w.r.t. the $\textsc{Cxt}^{\top,\backslash}$ frames is proven in two steps.

1. Logic $\mathbf{Cxt^{u,-}}$ is first proven to be complete w.r.t. the class of $\mathfrak{IE}^{\top,\backslash}$ frames.

2. It is then proven that for any formula ϕ on $\mathcal{L}^{u,-}_n$: $\mathfrak{IE}^{\top,\backslash} \models \phi$ iff $\textsc{Cxt}^{\top,\backslash} \models \phi$.

Some facts need to be proven about the canonical model of logic $\mathbf{Cxt^{u,-}}$. Given the presence of nominals in the language, and of rule Name in the axiomatics, the standard techniques for normal modal logics need to be extended. In particular, its canonical model should be built on maximal consistent *named* sets. The general definition is the following one. Let Λ be a given logic on a multi-modal language $\mathcal{L}_n$ with nominals. A maximal Λ-consistent named set of formulae of the multi-modal language $\mathcal{L}_n$ with nominals is a set Φ s.t.: (a) $\perp$ is not derivable in Λ from Φ (i.e., Λ-consistency of Φ); (b) every set properly including Φ is Λ-inconsistent. Every maximal Λ-consistent set Φ is such that: $\Lambda \subseteq \Phi$; Φ is closed under rule `MP` and `Name`; for all formulae ϕ either $\phi \in \Phi$ or $\neg\phi \in \Phi$; for all formulae $\phi, \psi : \phi \vee \psi \in \Phi$ iff $\phi \in \Phi$ or $\psi \in \Phi$.

Lemma A.6. *(Maximal Λ-consistent named sets)*
Maximal Λ-consistent named sets always contain at least one nominal.

Proof. Let Φ be a maximal Λ-consistent set of formulae on $\mathcal{L}_n$ with nominals. Suppose per absurdum that $\forall \nu \in \mathbb{N}$, $\neg\nu \in \Phi$. It follows that for every ν there exists a finite conjunction θ of formulae from Φ such that: $\vdash \nu \rightarrow \neg\theta$. Now, either ν occurs in θ and thus $\nu \in \Phi$, or ν does not occur in θ and therefore, by rule `Name`, $\neg\theta \in \Phi$ which is impossible. $\square$

Obviously, the standard properties of maximal Λ-consistent sets still obtain. The canonical model of $\mathbf{Cxt^{u,-}}$ should be built with maximal $\mathbf{Cxt^{u,-}}$-consistent named sets. In addition, the canonical model should be surjective. The following results show how this can be done.

Consider, first of all, that since logic $\mathbf{Cxt^{u,-}}$ extends logic $\mathbf{Cxt^{u}}$, we know by Theorem A.4 that the canonical model of logic $\mathbf{Cxt^{u,-}}$ will contain an equivalence relation $R_u^{\mathbf{Cxt^{u,-}}}$ such that for every $i \in C$, $R_i^{\mathbf{Cxt^{u,-}}} \subseteq R_u^{\mathbf{Cxt^{u,-}}}$. Recall also that every equivalence relation yields a partition on its domain. The clusters of the partition yielded by $R_u^{\mathbf{Cxt^{u,-}}}$ on $W^{\mathbf{Cxt^{u,-}}}$ containing state w is denoted as the set $r_u^{\mathbf{Cxt^{u,-}}}(w)$.

Lemma A.7. *(Maximal* $\mathbf{Cxt^{u,-}}$*-consistent named sets)*
The following facts hold for maximal $\mathbf{Cxt^{u,-}}$*-consistent named sets:*

1. *Each nominal in* $\mathbb{N}$ *is contained in at least one maximal* $\mathbf{Cxt^{u,-}}$*-consistent set.*
2. *If a nominal is contained in a maximal* $\mathbf{Cxt^{u,-}}$*-consistent set* $w \in W^{\mathbf{Cxt^{u,-}}}$ *then it is not contained in any other maximal* $\mathbf{Cxt^{u,-}}$*-consistent set* $w' \in W^{\mathbf{Cxt^{u,-}}}$ *which is accessible from* w *via* $R_u^{\mathbf{Cxt^{u,-}}}$*. In other words, if two maximal* $\mathbf{Cxt^{u,-}}$*-consistent sets contain the same nominal, and belong to the same cluster of the partition of* $W^{\mathbf{Cxt^{u,-}}}$ *yielded by* $R_u^{\mathbf{Cxt^{u,-}}}$*, then they are the same set.*

Proof. Clause 1 follows easily from Lemma A.1 and the fact that every state $w \in W^{\mathbf{Cxt^{u,-}}}$ contains formula $\langle u \rangle \nu$ (axiom `Least`). Clause 2 is proven in two steps. (a) Given a nominal $\nu \in \Phi$, for any maximal $\mathbf{Cxt^{u,-}}$-consistent set Φ it is proven that for all ϕ: $\phi \in \Phi$ iff $[u](\nu \rightarrow \phi) \in \Phi$. (b) Given two maximal $\mathbf{Cxt^{u,-}}$-consistent sets Φ and Φ', if $\nu \in \Phi, \Phi'$ and $\Phi R_u^{\mathbf{Cxt^{u,-}}} \Phi'$ then $\Phi = \Phi'$. Let us prove (a). From left to right. We assumed a nominal $\nu \in \Phi$, hence if $\phi \in \Phi$ then $\nu \wedge \phi \in \Phi$, being Φ a maximal $\mathbf{Cxt^{u,-}}$-consistent set. The set Φ also contains formula $\phi \rightarrow \langle u \rangle \phi$ (i.e., the contrapositive of axiom $\mathtt{T}^u$) and $\langle u \rangle (\nu \wedge \phi) \rightarrow [u](\nu \rightarrow \phi)$ (i.e., axiom `Most`) from which it follows that $\langle u\rangle(\nu \wedge \phi) \in \Phi$ and hence that $[u](\nu \rightarrow \phi) \in \Phi$. From right to left: for any $\phi \in \Phi$, if $[u](\nu \rightarrow \phi) \in \Phi$ then by axiom $\mathtt{T}^u$ we obtain $\nu \rightarrow \phi \in \Phi$ and then by $\mathtt{MP}$ $\phi \in \Phi$. Let us prove (b) per absurdum. Suppose $\Phi \neq \Phi'$. Then there should exist a formula ϕ such that $\phi \in \Phi$ and $\phi \notin \Phi'$ and hence $\neg\phi \in \Phi'$. From (a) it follows that $[u](\nu \rightarrow \phi) \in \Phi$ and since $\Phi R_u^{\mathbf{Cxt^{u,-}}} \Phi'$ we obtain that $\nu \rightarrow \phi \in \Phi'$ and via $\mathtt{MP}$ $\phi \in \Phi'$, which is impossible. □

Clause 1 just states that all nominals get a denotation, that is to say, the interpretation function from nominals to singletons is defined on every nominal. Clause 2 is particularly interesting. It states that the same nominal can in fact belong to different maximal $\mathbf{Cxt^{u,-}}$-consistent sets if these sets are not related via $R_u^{\mathbf{Cxt^{u,-}}}$. To put it otherwise, nominals behave as real names if they refer to sets in a same cluster in the partition yielded by $R_u^{\mathbf{Cxt^{u,-}}}$. It follows that interpreting nominals on a generated frame corresponding to some cluster $r_u^{\mathbf{Cxt^{u,-}}}(w)$ ensures that they will behave like names.

Definition A.2. *(Canonical model for logic* $\mathbf{Cxt^{u,-}}$*)*
The canonical model $\mathcal{M}^{\mathbf{Cxt^{u,-}}}$ *for logic* $\mathbf{Cxt^{u,-}}$ *in language* $\mathcal{L}_n^{u,-}$ *is the structure:*

$$\left\langle W^{\mathbf{Cxt^{u,-}}}, \{R_i^{\mathbf{Cxt^{u,-}}}\}_{i\in C}, \mathcal{I}^{\mathbf{Cxt^{u,-}}} \right\rangle$$

where:

- *Set* $W^{\mathbf{Cxt^{u,-}}}$ *is the set of maximal* $\mathbf{Cxt^{u,-}}$*-consistent named sets which are* $[u]$*-connected to a given maximal* $\mathbf{Cxt^{u,-}}$*-consistent named set* w*, that is:*

$$W^{\mathbf{Cxt^{u,-}}} = \{w' \mid \{\phi \mid [u]\phi \in w\} \subseteq w'\}.$$

- *The canonical relations* $\{R_i^{\mathbf{Cxt^{u,-}}}\}_{i\in C}$ *and interpretation* $\mathcal{I}^{\mathbf{Cxt^{u,-}}}$ *are defined as in Definition A.1.*

It can now be shown that nominals behave like proper names since they all denote one and only one element in $W^{\mathbf{Cxt^{u,-}}}$. The canonical model of $\mathbf{Cxt^{u,-}}$ is therefore surjective.

Corollary A.3. *(Nominals are names in* $\mathcal{M}^{\mathbf{Cxt^{u,-}}}$*)*
Let $\mathcal{M}^{\mathbf{Cxt^{u,-}}} = \left\langle W^{\mathbf{Cxt^{u,-}}}, \{R_i^{\mathbf{Cxt^{u,-}}}\}_{i\in C}, \mathcal{I}^{\mathbf{Cxt^{u,-}}} \right\rangle$ *be the canonical model of logic* $\mathbf{Cxt^{u,-}}$*. It is the case that: for every nominal* ν*,* $\mathcal{I}^{\mathbf{Cxt^{u,-}}}(\nu)$ *is the only element of* $W^{\mathbf{Cxt^{u,-}}}$ *containing* ν*.*

Proof. It follows directly from Definition A.2 Lemmata A.6, A.7. □

Now, what we still miss is a new version of the truth lemma (Lemma A.2). In effect, this boils down to prove that there are enough maximal $\mathbf{Cxt^{u,-}}$-consistent named sets to support an existence lemma (Lemma A.1).

Lemma A.8. *(Truth Lemma for logic* $\mathbf{Cxt^{u,-}}$*)*
Let $\mathcal{M}^{\mathbf{Cxt^{u,-}}} = \left\langle W^{\mathbf{Cxt^{u,-}}}, \{R_i^{\mathbf{Cxt^{u,-}}}\}_{i\in C}, \mathcal{I}^{\mathbf{Cxt^{u,-}}} \right\rangle$ *be the canonical model of logic* $\mathbf{Cxt^{u,-}}$*. It holds that:* $\mathcal{M}^{Cxt^{u,-}}, w \models \phi$ *iff* $\phi \in w$*.*

Proof. The proof is, as usual, on the complexity of ϕ. The interesting case concerns modalities. It needs to be proven that if $\langle i\rangle\phi \in w$ then there exists a state $w' \in W^{Cxt^{u,-}}$ such that $wR_i^{Cxt^{u,-}}w'$ and $\phi \in w'$. This can be shown as usual by building w' on the set $\{\psi \mid [i]\psi \in w\} \cup \{\phi\}$. Such set can be proven consistent in the usual way. What matters here, is to prove that $\{\psi \mid [i]\psi \in w\}$ contains at least one nominal since, as a result, w' will be named. The desired fact is proven per absurdum like in the proof of Lemma A.6 using rule `Name`. Hence, set $\{\psi \mid [i]\psi \in w\} \cup \{\phi\}$ is consistent and named, therefore, it can be extended to the desired w'. □

We can now prove strong completeness with respect to $\mathfrak{IE}^{\top,\setminus}$ frames.

Theorem A.6. *(Completeness of* $\mathbf{Cxt^{u,-}}$*)*
Logic $\mathbf{Cxt^{u,-}}$ *is strongly complete w.r.t. the class of* $\mathfrak{IE}^{\top,\setminus}$ *frames, that is, frames satisfying the following clauses:*

1. *They are i-j transitive, i-j euclidean.*
2. *They contain a universal relation* R_u*.*
3. *The set of relations* $\{R_i\}_{i\in C}$ *is such that, for any atomic context index* c *and states* $w, w' \in W$*: (3.a)* wR_uw' *implies* wR_cw' *or* $wR_{-c}w'$*; and (3.b)* $wR_{-c}w'$ *implies not* wR_cw'*.*

Proof. By Proposition A.2, given a $\mathbf{Cxt^{u,-}}$-consistent set Φ of formulae, it suffices to find a model state pair $(\mathcal{M}, w)$ such that: (a) $\mathcal{M}, w \models \Phi$, and (b) the frame $\mathcal{F}$ on which $\mathcal{M}$ is based satisfies clauses 1-3 and $\mathcal{M}$ is surjective. Claim (a) is proven by making use of Lemma A.8. As to claim (b), it follows from Corollary A.3 that $\left\langle W^{\mathbf{Cxt^{u,-}}}, \{R_i^{\mathbf{Cxt^{u,-}}}\}_{i\in C}, \mathcal{I}^{\mathbf{Cxt^{u,-}}}\right\rangle$ is surjective. It remains to be proven that the frame $\left\langle W^{\mathbf{Cxt^{u,-}}}, \{R_i^{\mathbf{Cxt^{u,-}}}\}_{i\in C}\right\rangle$ of the canonical model satisfies Clauses 1-3. Clause 1 and Clause 2 are proven to be satisfied by Theorem A.4 since $\mathbf{Cxt^{u,-}}$ extends $\mathbf{K45_n^{ij}}$ and $\mathbf{Cxt^{u}}$ and by considering that the frame of the canonical model is generated. Claims (3.a) and (3.b) of clause 3 remain to be proven. To prove claim (3.a) it has to be shown that: for any atomic context index c and states $w, w' \in W^{\mathbf{Cxt^{u,-}}}$, $wR_u^{\mathbf{Cxt^{u,-}}}w'$ implies $wR_c^{\mathbf{Cxt^{u,-}}}w'$ or $wR_{-c}^{\mathbf{Cxt^{u,-}}}w'$. Consider two states $w, w' \in W^{\mathbf{Cxt^{u,-}}}$ such that $wR_u^{\mathbf{Cxt^{u,-}}}w'$ and suppose that $\phi \in w'$. Since w is a maximal $\mathbf{Cxt^{u,-}}$-consistent named set, it contains formula $\langle u\rangle\phi \rightarrow (\langle c\rangle\phi \vee \langle -c\rangle\phi)$ (i.e., the contrapositive of axiom `Covering`) and therefore $\langle c\rangle\phi \vee \langle -c\rangle\phi \in w$. For the properties of maximal consistent sets it follows that either $\langle c\rangle\phi \in w$ or $\langle -c\rangle\phi \in w$, and hence by Definition A.2, either $wR_c^{\mathbf{Cxt^{u,-}}}w'$ or $wR_{-c}^{\mathbf{Cxt^{u,-}}}w'$, which proves (3.a). As to (3.b), it should be proven that for any atomic context index c and states $w, w' \in W^{\mathbf{Cxt^{u,-}}}$, $wR_{-c}^{\mathbf{Cxt^{u,-}}}w'$ implies not $wR_c^{\mathbf{Cxt^{u,-}}}w'$. Suppose that $wR_{-c}^{\mathbf{Cxt^{u,-}}}w'$. By Clause 1 in Lemma A.7 we know that w' should contain at least one nominal. Suppose it to be v. By Clause 2 of this theorem, from $wR_{-c}^{\mathbf{Cxt^{u,-}}}w'$ it follows that $wR_u^{\mathbf{Cxt^{u,-}}}w'$ and from this, by Clause 2 in Lemma A.7, we know that there is no $w'' \in r_u^{\mathbf{Cxt^{u,-}}}(w)$ such that $v \in w''$. By Definition A.2 it follows that $\langle -c\rangle v \in w$. Now, w is a maximal $\mathbf{Cxt^{u,-}}$-consistent named set and it contains thus formula $\langle -c\rangle v \rightarrow \neg\langle c\rangle v$ (i.e., axiom `Packing`). It follows that $\neg\langle c\rangle v \in w$ and it is therefore not the case that $wR_c^{\mathbf{Cxt^{u,-}}}w'$, which proves claim (3.b). □

Lemma A.9. *(Semantic equivalence for* $\textsc{Cxt}^{\top,\backslash}$ *frames)*
For any formula ϕ *on* $\mathcal{L}_n^{u,-}$: $\mathfrak{TE}^{\top,\backslash} \models \phi$ *iff* $\textsc{Cxt}^{\top,\backslash} \models \phi$. *That is,* $\textsc{Cxt}^{\top,\backslash}$ *frames and* $\mathfrak{TE}^{\top,\backslash}$ *frames define the same logic.*

Proof. The proof is similar to the proofs of Lemmata A.4 and A.5. [$\Leftarrow$] From right to left: for every ϕ, $\textsc{Cxt}^{\top,\backslash} \models \phi$ implies $\mathfrak{TE}^{\top,\backslash} \models \phi$. The results follow by the application of Proposition 4.1. From $W = W_c \cup W_{-c}$ for any atomic context identifier c, it follows that for every $w, w' \in W$, $wR_u w'$ implies $wR_c w'$ or $wR_{-c} w'$. And from $W_c \cap W_{-c} = \emptyset$ for any atomic context identifier c, it follows that for every $w, w' \in W$, $wR_{-c} w'$ implies not $wR_c w'$. [$\Rightarrow$] From left to right: for every ϕ, $\mathfrak{TE}^{\top,\backslash} \models \phi$ implies $\textsc{Cxt}^{\top,\backslash} \models \phi$. Frames in $\mathfrak{TE}^{\top,\backslash}$ already contain a universal relation. It just needs to be shown that for any atomic index c: (a) $W^w \subseteq r_c(w) \cup r_{-c}(w)$ and (b) $r_c(w) \cap r_{-c}(w) \subseteq \emptyset$. Both claims are straightforwardly proven by observing that for any atomic context index c and states $w', w'' \in W^w$: $w' R_u^w w''$ (i.e., $w'' \in W^w$) implies $w' R_c^w w''$ (i.e., $w'' \in r_c(w)$) or $w' R_{-c}^w w''$ (i.e., $w'' \in r_{-c}(w)$); and $w' R_c^w w''$ (i.e., $w'' \in r_c(w)$) implies not $w' R_{-c}^w w''$ (i.e., $w'' \notin r_{-c}(w)$). □

Corollary A.4. *(Completeness of* $\mathbf{Cxt^{u,-}}$ *w.r.t.* $\textsc{Cxt}^{\top,\backslash}$ *frames)*
Logic $\mathbf{Cxt^{u,-}}$ *is strongly complete w.r.t. the class of* $\textsc{Cxt}^{\top,\backslash}$ *frames.*

Proof. Follows directly from Theorem A.6 and Lemma A.9. □

Bibliography

Ågotnes, T., Van der Hoek, W., Rodríguez-Aguilar, J. A., Sierra, C. and Wooldridge, M. (2007). On the logic of normative systems. In M. Veloso, editor, *Twentieth International Joint Conference on Artificial Intelligence, IJCAI07*, pages 1175–1180. AAAI Press, California.

Agerri, R. (2007). Norms of conversation in a framework for agent communication languages. In G. Boella, L. van der Torre and H. Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany.

Akman, V. and Surav., M. (1996). Steps toward formalizing context. *AI Magazine*, volume 17(3):pages 55–72.

Alchourrón, C. E. and Bulygin, E. (1971). *Normative Systems*. Springer Verlag.

Aldewereld, H. (2007). *Autonomy vs. Conformity. An Institutional Perspective on Norms and Protocols*. Siks dissertation series no. 2007-10, Universiteit Utrecht.

Aldewereld, H., Grossi, D., , Dignum, F. and Vázquez-Salceda, J. (2006a). Designing normative behavior via landmarks. In O. Boissier, J. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowski, J. S. ao Sichman and J. Vázquez-Salceda, editors, *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems. Proceedings of ANIREM 2005 and OOOP 2005*, volume 3913 of *LNCS*, pages 157–169. Springer-Verlag.

Aldewereld, H., Vázquez-Salceda, J., Dignum, F. and Meyer, J.-J.Ch. (2006b). Verifying norm compliance of protocols. In O. Boissier, J. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowsk, J. S. ao Sichman and J. Vázquez-Salceda, editors, *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems. Proceedings of ANIREM 2005 and OOOP 2005*, volume 3913 of *LNCS*, pages 222–236. Springer-Verlag.

Anderson, A. (1957). The formal analysis of normative concepts. *American Sociological Review*, volume 22:pages 9–17.

Anderson, A. (1958). A reduction of deontic logic to alethic modal logic. *Mind*, volume 22:pages 100–103.

Areces, C., Blackburn, P. and Marx, M. (2000). The computational complexity of hybrid temporal logics. *Logic Journal of the IGPL*, volume 8(5):pages 653–679. Oxford University Press.

Atkinson, K. and Bench-Capon, T. (2005). Levels of reasoning with legal cases. In P. E. Dunne and T. Bench-Capon, editors, *Proceedings of the ICAIL 2005 Workshop on Argumentation in AI and Law*, IAAIL Workshop Series, pages 1–11. Wolf Legal Publishers, Bologna.

Azzoni, G. (2003). Il cavallo di Caligola. In *Ontologia sociale potere deontico e regole costitutive*, pages 45–54. Quodlibet, Macerata, Italy.

Baader, F., Brandt, S. and Lutz, C. (2005). Pushing the $\mathcal{EL}$ envelope. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence IJCAI-05*. Morgan-Kaufmann Publishers, Edinburgh, UK.

Baader, F., Calvanese, D., McGuinness, D., Nardi, D. and Patel-Schneider, P. (2002). *The Description Logic Handbook*. Cambridge University Press, Cambridge, UK.

Baltag, A. and Moss, L. (2004). Logic for epistemic programs. *Synthese*, volume 139(2):pages 165–224. Special issue on Logic, Rationality and Action.

Benerecetti, M., Bouquet, P. and Ghidini, C. (2000). Contextual reasoning distilled. *Journal of Experimental and Theoretical Artificial Intelligence (JETAI)*, volume 12(3):pages 279–305.

van Benthem, J. (2002). Extensive games as process models. *Journal of Logic, Language and Information*, volume 11:pages 289–313.

van Benthem, J. (2005). Logic in games. Lecture Notes of the ILLC graduate course on Logic, Language and Information, Universiteit van Amsterdam, Amsterdam, The Netherlands.

van Benthem, J. (2006). Logic in philosophy. ILLC Preprint. PP-2005-21. To appear in Logic and Philosophy. Jacquette, J. and Gabbay, D. and Thagard, P. and Woods, J. (eds.).

van Benthem, J., van Eijck, J. and Kooi, B. (2006). Logics of communication and change. *Information and Computation*, volume 204(11):pages 1620–1662.

van Benthem, J., van Eijck, J. and Stebletsova, V. (1994). Modal logic, transition systems and processes. *Journal of Logic and Computation*, volume 4(5):pages 811–855.

Blackburn, P., de Rijke, M. and Venema, Y. (2001). *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, UK.

Boella, G. and van der Torre, L. (2003). Attributing mental attitudes to normative systems. In *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 942–943. ACM Press, New York, NY, USA.

Boella, G. and Van der Torre, L. (2004). Regulative and constitutive norms in normative multiagent systems. In D. Dubois, C. A. Christopher A. Welty and M. Williams, editors, *Proceedings of KR2004, Whistler, Canada*, pages 255–266.

Boella, G. and van der Torre, L. (2005). Constitutive norms in the design of normative multiagent systems. In F. Toni and P. Torroni, editors, *CLIMA VI*, LNAI 3900, pages 303–319. Springer-Verlag.

Boella, G., van der Torre, L. and Verhagen, H., editors (2007). *Normative Multiagent Systems*, number 07122 in Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany.

Boissier, O., Padget, J., Dignum, V., Lindemann, G., Matson, E., Ossowski, S., Sichman, J. S. and Vázquez-Salceda, J., editors (2006). *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems. Proceedings of ANIREM 2005 and OOOP 2005*, number 3913 in LNAI. Springer, Utrecht, The Netherlands. Revised and selected papers.

Bouquet, P., Giunchiglia, F., van Harmelen, F., Serafini, L. and Stuckenschmidt, H. (2004). Contextualizing ontologies. *Journal of Web Semantics*, volume 1(4):pages 325–343.

Bradshaw, J. M., Dutfield, S., Carpenter, B., Jeffers, R. and Robinson, T. (1995). KAoS: A Generic Agent Architecture for Aerospace Applications. In T. Finin and J. Mayfield, editors, *Proceedings of the CIKM '95 Workshop on Intelligent Information Agents*. Baltimore, Maryland, US.

Bresciani, P., Giorgini, P., Giunchiglia, F., Mylopoulos, J. and Perini, A. (2004). Tropos: An agent-oriented software development methodology. *Journal of Autonomous Agents and Multi-Agent Systems*, volume 8(3):pages 203–236.

van der Broek, E. L., Jonker, C. M., Sharpanskykh, A., Treur, J. and Yolum, P. (2005). Formal modeling and analysis of organizations. In O. Boissier, V. Dignum, E. Matson and J. Sichman, editors, *Proceedings ofthe International Workshop on Organizations in Multi-Agent Systems (OOOP)*. AAMAS'05, Utrecht, The Netherlands.

Broersen, J. (2003). *Modal action Logics for Reasoning about Reactive Systems*. Siks dissertation series no. 2003-2, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.

Bulygin, E. (1992). On norms of competence. *Law and Philosophy 11*, pages 201–216.

Burgess, A. (1962). *A Clockwork Orange*. Penguin Books, London, UK, 2000 edition.

Buvač, S., Buvač, S. V. and Mason, I. A. (1994). The semantics of propositional contexts. *Proceedings of the 8th ISMIS. LNAI-869*, pages 468–477.

Buvač, S. V. and Mason, I. A. (1993). Propositional logic of context. *Proceedings AAAI'93*, pages 412–419.

Castelfranchi, C. (2000). Engineering social order. In *Engineering Societies in the Agents World*, volume 1972 of *LNAI*, pages 1–18. Springer-Verlag.

Castelfranchi, C. (2003). The micro-macro constitution of power. *ProtoSociology*, volume 18:pages 208–268.

Castelfranchi, C. (2004). Formalizing the informal?: Dynamic social order, botton-up social control, and spontaneous normative relations. *Journal of Applied Logic*, volume 1(1-2):pages 47–92.

Conte, R. and Castelfranchi, C. (1995). *Cognitive and Social Action*. UCL Press, London, UK.

d'Altan, P., Meyer, J.-J.Ch. and Wieringa, R. (1993). An integrated framework for ought-to-be and ought-to-do constraints. In Y. H. Tan, editor, *Working Papers of the Workshop om Deontic and Non-Monotonic Logics*, pages 14–45. Rotterdam.

Debray, R. (1997). *L' État Séducteur*. Gallimard, Paris.

Decker, K. and Lesser, V. (1995). Designing a family of coordination algorithms. In V. Lesser, editor, *Proceedings of ICMAS'95*, pages 73–80. The MIT Press: Cambridge, MA, USA, San Francisco, CA, USA.

Dignum, F. (1999). Autonomous agents with norms. *Artificial Intelligenc and law*, volume 7:pages 69–79.

Dignum, F. (2002). Abstract norms and electronic institutions. In *Proceedings of the International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA '02), Bologna*, pages 93–104.

Dignum, F., Dignum, V., Koenig, S., Kraus, S., Singh, M. P. and Wooldridge, M., editors (2005). *The 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005)*. ACM Press, Utrecht, The Netherlands.

Dignum, V. (2003). *A Model for Organizational Interaction*. SIKS Dissertation Series.

Dignum, V. (2006). Coordinating tasks in agent organizations. or: Can we ask you to read this paper? In *Preproceedings of COIN@ECAI'06*.

Dignum, V., Fornara, N., Noriega, P., Boella, G., Boissier, O., Matson, E. and Vázquez-Salceda, J., editors (2007). *Coordination, Organizations, Institutions, and Norms in Multi-agents Systems. Proceedings of COIN 2006*, LNAI. Springer, Hakodate, Japan and Riva del Garda, Italy. Forthcoming.

Emerson, E. (1990). Temporal and modal logic. In J. v. Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B: Formal Models and Semantics, chapter 14, pages 996–1072. MIT Press, Boston, US.

Esteva, M., de la Cruz, D. and Sierra, C. (2002). ISLANDER: an electronic institutions editor. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 1045–1052. ACM Press, New York, NY, USA.

Esteva, M., Rodríguez-Aguilar, J., Rosell, B. and Arcos, J. (2004). Ameli: An agent-based middleware for electronic institutions. In *Third International Joint Conference on Autonomous Agents and Multi-agent Systems*. New York, US.

Etzioni, A. (1964). *Modern Organizations*. Prentice Hall.

Fagin, R., Halpern, J. and Vardi, M. (1988). Belief, awareness and limited reasoning. *Artificial Intelligence*, volume 34:pages 39–76.

Fararo, T. J. (1997). Reflections on mathematical sociology. *Sociological Forum*, volume 12:pages 73–101.

Fine, K. (1975). Vagueness, truth and logic. *Synthese*, volume 30:pages 265–300.

Fox, M. S. (1988). An organizational view of distributed systems. *Distributed Artificial Intelligence*, pages 140–150.

van Fraassen, B. C. (1966). Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, volume 63:pages 481–495.

Friedell, M. F. (1967). Organizations as semilattices. *American Sociological Review*, volume 32:pages 46–54.

Gargov, G. and Goranko, V. (1993). Modal logic with names. *Journal of Philosophical Logics*, volume 22:pages 607–636.

Gelati, J., Rotolo, A., Sartor, G. and Governatori, G. (2004). Normative autonomy and normative co-ordination: Declarative power, representation, and mandate. *Artif. Intell. Law*, volume 12(1-2):pages 53–81.

Ghidini, C. and Giunchiglia, F. (2001). Local models semantics, or contextual reasoning = locality + compatibility. *Artificial Intelligence*, volume 127(2):pages 221–259.

Giddens, A. (1984). *Social Theory and Modern Sociology*. Polity Press, Cambridge, UK.

Gini, M., Ishida, T., Castelfranchi, C. and Johnson, W. L., editors (2002). *The 1st International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS 2002)*. ACM Press, Bologna, Italy.

Giunchiglia, F., Lee, E., Lifschitz, J., McCain, N. and Turner, H. (2004). Nonmonotonic causal theories. *Artificial Intelligence*, volume 153:pages 49–104.

Goble, L. and Meyer, J.-J.Ch., editors (2006). *Proceedings of DEON'06*, number 4048 in LNAI. Springer, Utrecht, The Netherlands.

Goldman, A. I. (1976). *A Theory of Human Action*. Princeton University Press, Princeton, US.

Goranko, V. and Passy, S. (1992). Using the universal modality: Gains and questions. *Journal of Logic and Computation*, volume 2(1):pages 5–30. Oxford University Press.

Governatori, G., Gelati, J., Rotolo, A. and Sartor, G. (2002). Actions, institutions, powers. preliminary notes. In *International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'02)*, pages 131–147.

Grossi, D., Aldewereld, H. and Dignum, F. (2006a). Ubi lex ibi poena. designing norm enforcement in electronic institutions. In *Preproceedings of COIN@AAMAS'06*.

Grossi, D., Aldewereld, H., Vázquez-Salceda, J. and Dignum, F. (2006b). Ontological aspects of the implementation of norms in agent-based electronic institutions. *Computational & Mathematical Organization Theory (CMOT)*, volume 12(2/3):pages 251–275.

Grossi, D. and Dignum, F. (2004). From abstract to concrete norms in agent institutions. To be published in Proceedings of FAABS III Workshop, Washington.

Grossi, D., Dignum, F., Dastani, M. and Royakkers, L. (2005a). Foundations of organizational structure in multi-agent systems. In *Proceedings of AAMAS'05*, pages 690–697.

Grossi, D., Dignum, F., Dignum, V., Dastani, M. and Royakkers, L. (2006c). Structural aspects of the evaluation of agent organizations. In *Preproceedings of COIN@ECAI'06*.

Grossi, D., Dignum, F., Dignum, V., Dastani, M. and Royakkers, L. (2006d). Structural evaluation of agent organizations. In *Proceedings of AAMAS'06*, pages 1110 – 1112. ACM Press, Hakodate, Japan.

Grossi, D., Dignum, F. and Meyer, J.-J.Ch. (2005b). Context in categorization. In L. Serafini and P. Bouquet, editors, *Proceedings of CRR'05, Workshop on Context Representation and Reasoning*, volume 136 of *CEUR Workshp Proceedings*. Paris.

Grossi, D., Dignum, F. and Meyer, J.-J.Ch. (2005c). Contextual taxonomies. In J. Leite and P. Toroni, editors, *Post-proceedings of CLIMA V, 5th International Workshop on Computational Logic in Multi-Agent Systems*, LNAI 3487, pages 33–51. Springer-Verlag.

Grossi, D., Dignum, F. and Meyer, J.-J.Ch. (2006e). Contextual terminologies. In F. Toni and P. Torroni, editors, *Post-proceedings of CLIMA VI, 6th International Workshop on Computational Logic in Multi-Agent Systems*, LNAI 3900, pages 284–302. Springer-Verlag.

Grossi, D., Dignum, F. and Meyer, J.-J.Ch. (2007). A formal road from institutional norms to organizational structures. In *Proceedings of 6th International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS 2007*. Honolulu, Hawai'i, US.

Grossi, D., Dignum, F., Royakkers, L. and Meyer, J.-J.Ch. (2004). Collective obligations and agents: Who gets the blame. In A. Lomuscio and D. Nute, editors, *Proceedings of DEON'04*, pages 129–145.

Grossi, D., Meyer, J.-J.Ch. and Dignum, F. (2005d). Modal logic investigations in the semantics of counts-as. In *Proceedings of the Tenth International Conference on Artificial Intelligence and Law (ICAIL'05)*, pages 1–9. ACM.

Grossi, D., Meyer, J.-J.Ch. and Dignum, F. (2006f). Classificatory aspects of counts-as: An analysis in modal logic. *Journal of Logic and Computation*, volume 16(5):pages 613–643. Oxford University Press.

Grossi, D., Meyer, J.-J.Ch. and Dignum, F. (2006g). Counts-as: Classification or constitution? An answer using modal logic. In L. Goble and J.-J.Ch. Meyer, editors, *Proceedings of DEON 2006*, LNAI 4048, pages 115–130.

Grossi., D., Meyer, J.-J.Ch. and Dignum, F. (2007). The many faces of counts-as: A formal analysis of constitutive-rules. To appear in the Journal of Applied Logic.

Grossi, D., Royakkers, L. and Dignum, F. (2006h). Organizational structure and responsibility: An analysis in a dynamic logic of organized collective agency. To appear in the Journal of Artificial Intelligence and Law.

Hage, J. and Verheij, B. (1999). The law as a dynamic interconnected system of states of affairs. *IJHCS: International Journal of Human-Computer Studies*, volume 51:pages 1043–1077.

Hannoun, M., Boissier, O., Sichman, J. S. and Sayettat, C. (2000). MOISE: An organizational model for multi-agent systems. In *IBERAMIA-SBIA*, pages 156–165.

Harary, F. (1959a). Graph theoretic methods in the management sciences. *Managemenet Science*, volume 5(4):pages 387–403.

Harary, F. (1959b). Status and contrastatus. *Sociometry*, volume 22(1):pages 23–43.

Harary, F., Norman, R. and Cartwright, D. (1965). *Structural Models: An Introduction to the Theory of Directed Graphs*. John Wiley & Sons, New York.

Hare, R. M. (1952). *The Language of Morals*. Oxford University Press, Oxford, UK.

Hart, H. L. A. (1958). Positivism and the separation of law and morality. *Harvard Law Review*, volume 71:pages 593–629.

Hart, H. L. A. (1961). *The Concept of Law*. Clarendon Press, Oxford, UK.

Hegel, G. W. F. (1821). *Grundlinien der Philosophie des Rechts. Naturrecht und Staatswissenschaft*. von Helmut Reichelt, Frankfurt am Main. 1972.

Hindriks, F. (2006). But where is the university? Manuscript.

Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca.

Hobbes, T. (1651). *Leviathan*. Penguing Classics, c.b. macpherson, 1968 edition.

Hohfeld, W. N. (1911). Some fundamental legal conceptions as applied in judicial reasoning. *Yale Law Journal*, volume 23(16):pages 16–59.

Horling, B. and Lesser, V. (2004). A survey of multi-agent organizational paradigms. Computer Science Technical Report 04-45, University of Massachusetts.

Horling, B. and Lesser, V. (2005). Using ODML to model and design organizations for multi-agent systems. In O. Boissier, V. Dignum, E. Matson and J. Sichman, editors, *Proceedings ofthe International Workshop on Organizations in Multi-Agent Systems (OOOP)*. AAMAS'05, Utrecht, The Netherlands.

Horrocks, I., Kutz, O. and Sattler, U. (2006). The even more irresistible $\mathcal{SROIQ}$. In P. Doherty, J. Mylopoulos and C. Welty, editors, *Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KR 2006)*, pages 57–67. AAAI Press, Menlo Park, California.

Hübner, J. F., Sichman, J. S. and Boissier, O. (2002). Moise+: Towards a structural functional and deontic model for mas organization. In *Proceedings of the First Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'02)*. ACM Press, Bologna, Italy.

Husserl, E. (1988). *Vorlesungen über Ethik und Wertlehre (1908–1914)*, volume 28 of *Husserliana*. Kluwer Academic Publisher, Dordrecht.

Ioerger, T. and Johnson, C. (2001). A formal model of responsibilities in agent-based teamwork. In *Proceedings of the fifth international conference on Autonomous Agents, Montreal, Canada*. ACM Press.

Jackson, M. O. (2001). A crash course in implementation theory. *Social Choice and Welfare*, volume 18:pages 655–708.

Jennings, N. R., Sierra, C., Sonnenberg, L. and Tambe, M., editors (2004). *The 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*. IEEE Computer Society, New York, NY, USA.

Jones, A. J. I. and Sergot, M. (1992). Deontic logic in the representation of law: towards a methodology. *Artificial Intelligence and Law 1*.

Jones, A. J. I. and Sergot, M. (1993). On the characterization of law and computer systems. *Deontic Logic in Computer Science*, pages 275–307.

Jones, A. J. I. and Sergot, M. (1996). A formal characterization of institutionalised power. *Journal of the IGPL*, volume 3:pages 427–443.

Jónsson, B. and Tarski, A. (1951). Boolean algebras with operators: Part I. *American Journal of Mathematics*, volume 73:pages 891–939.

Jónsson, B. and Tarski, A. (1952). Boolean algebras with operators: Part II. *American Journal of Mathematics*, volume 75:pages 127–162.

Kamps, J. and Marx, M. (2002). Notions of indistinguishability for semantic web languages. In I. Horrocks and J. A. Hendler, editors, *Proceedings of The Semantic Web - ISWC 2002, First International Semantic Web Conference*, number 2342 in Lecture Notes in Computer Science, pages 30–38. Springer, Sardinia, Italy.

Kanger, S. (1985). On realization of human rights. *Acta Philosophica Fennica*, volume 38:pages 71–79. Holmström, G. and Jones, A. (eds.).

D. Harel amd Kozen, D. and Tiuryn, J. (1984). Dynamic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic: Volume II: Extensions of Classical Logic*, pages 497–604. Reidel, Dordrecht, The Netherlands.

Krabbendam, J. and Meyer, J.-J.Ch. (1999). Contextual deontic logics. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems*, pages 347–362. IOS Press, Amsterdam.

Krabbendam, J. and Meyer, J-J.Ch. (2000). Release logics for temporalizing dynamic logic, orthogonalising modal logics. In M. Barringer, M. Fisher, D. Gabbay and G. Gough, editors, *Advances in Temporal Logic*, pages 21–45. Kluwer Academic Publisher.

Krackhardt, D. (1994). Graph theoretical dimensions of informal organizations. In C. M. Carley and M. J. Prietula, editors, *Computational Organization Theory*, pages 89–110. Lawrence Erlbaum Associates.

Kraus, S., Lehmann, D. J. and Magidor, M. (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, volume 44(1-2):pages 167–207.

Levesque, H. and Brachman, R. J. (1987). Expressiveness and tractability in knowledge representation and reasoning. *Computational Intelligence*, volume 3:pages 78–93.

Levesque, H. J. (1990). All I Know: A study in autoepistemic logic. *Artificial Intelligence*, volume 42(2–3):pages 263–309.

Lewis, D. (1999). Many, but almost one. In *Papers in Metaphysics and Epistemology*, pages 164–182. Cambridge University Press.

Lindahl, L. (2004). Deduction and justification in the law. the role of legal terms and concepts. *Ratio Juris*, volume 17(2):pages 182–202.

Lindahl, L. and Odelstad, J. (2006). Open and closed intermediaries in normative systems. In T. van Engers, editor, *Proceedings of the Nineteenth JURIX Conference on Legal Knowledge and Information Systems (JURIX 2006)*, pages 91–100.

Lindahl, L. and Odelstad, J. (2007). Intermediaries and intervenients in normative systems. To appear in the Journal of Applied Logic.

Lomuscio, A. and Nute, D., editors (2004). *Proceedings of DEON'04*.

Lomuscio, A. and Sergot, M. (2000). Investigations in grounded semantics for multi-agent systems specification via deontic logic. Technical report, Imperial College.

Lomuscio, A. and Sergot, M. (2003). Deontic intepreted systems. *Studia Logica*, volume 75(1):pages 63–92.

Lopez, F., Luck, M. and d'Inverno, M. (2006). A normative framework for agent-based systems. *Computational and Mathematical Organization Theory*, volume 12:pages 227–250.

Łukasiewicz, J. (1951). *Aristotle's Syllogistic, From the Standpoint of Modern Formal Logic*. Clarendon Press, Oxford, UK.

Makinson, D. (1999). On a fundamental problem of deontic logic. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems. New Studies in Deontic Logic and Computer Science*, volume 49 of *Frontiers in Artificial Intelligence and Applications*, pages 29–53. IOS Press, Amsterdam.

Makinson, D. (2005). *Bridges from Classical to Nonmonotonic Logic*, volume 5 of *Texts in Computing*. King's College Publications.

Makinson, D. and van der Torre, L. (2000). Input-output logics. *Journal of Philosophical Logic*, volume 29(4):pages 383–408.

Malone, T. and Smith, S. A. (1988). Modeling the performance of organizational structures. *Operational Research*, volume 36(3):pages 421–436.

Malone, T. W. (1987). Modeling coordination in organizations and markets. *Manageement Sciences*, volume 33(10):pages 1317–1332.

McCarthy, J. (1980). Circumscription—a form of non-monotonic reasoning. *Artificial Intelligence*, volume 13:pages 27–39.

McCarthy, J. (1986). Notes on formalizing contexts. In T. Kehler and S. Rosenschein, editors, *Proceedings of the Fifth National Conference on Artificial Intelligence*, pages 555–560. Morgan Kaufmann, Los Altos, California.

Meyer, J.-J.Ch. (1988). A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, volume 29(1):pages 109–136.

Meyer, J.-J.Ch., de Boer, F., van Eijk, R. M., Hindriks, K. V. and van der Hoek, W. (2001). On programming KARO agents. *Logic Journal of the IGPL*, volume 9(2).

Meyer, J.-J.Ch. and van der Hoek, W. (1995). *Epistemic Logic for AI and Computer Science*, volume 41 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, UK.

Meyer, J.-J.Ch. and van der Hoek, W. (1991). Non-monotonic reasoning by monotonic means. In J. van Eijck, editor, *JELIA '90: Proceedings of the European Workshop on Logics in AI, London, UK*, 478, pages 399–411. Springer-Verlag, London, UK.

Meyer, J.-J.Ch. and Wieringa, R. J., editors (1993). *Deontic Logic in Computer Science. Normative System Specification. Proceedings of the 1st International Workshop on Deontic Logic in Computer Science (DEON'91)*. John Wiley & Sons, Amsterdam, The Netherlands.

Moore, G. E. (1903). *Principia Ethica*. Cambridge University Press.

Morgenstern, O. (1951). Prolegomena to a theory of organizations. Manuscript. URL `http://qss.stanford.edu/~godfrey/Morgenstern/prolegom.pdf`

Nayak, P. (1994). Representing multiple theories. In *Proceedings of the 12th National Conference on Artificial Intelligence, Volume 2. Seattle, WA, USA*, pages 1154–1160. AAAI Press.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, Cambridge, UK.

Omicini, A., Ricci, A., Viroli, A., Castelfranchi, C. and Tummolini, L. (2004). Coordination artifacts: Environment-based coordination for intelligent agents. In *AAMAS'04*.

Passy, S. and Tinchev, T. (1985). Pdl with data constants. *Information Processing Letters*, volume 20:pages 35–41. Elsevier Science Publishers.

Passy, S. and Tinchev, T. (1991). An essay in combinatorial dynamic logic. *Information and Computation*, volume 93:pages 263–332.

Pauly, M. (2005a). Changing the rules of play. *Topoi*, volume 24(2):pages 209–220.

Pauly, M. (2005b). Programming and verifying subgame perfect mechanisms. *Journal of Logic and Computation*, volume 15(3):pages 295–316.

Pauly, M. and Wooldridge, M. (2003). Logic for mechanism design–a manifesto. In *Proceedings of the 2003 Workshop on Game Theory and Decision Theory in Agent Systems (GTDT-2003)*. Melbourne, Australia.

Pavese, C. (1936). I mari del sud. In *Le Poesie*. Einaudi, Torino, Italy, 1998 edition.

Pawlak, Z. (1991). *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academic, Dordrecht, The Netherlands.

Peczenik, A. (1989). *On Law and Reason*. Kluwer, Dordrecht.

Pitt, J. (1999). ALFEBIITE project. European Project. IST 1999 10298.
URL `http://www.iis.ee.ic.ac.uk/~alfebiite/ab-overview-page.htm`

Pörn, I. (1977). *Action theory and social science. Some formal models*. Reidel Publishing Company, Dordrecht, The Netherlands.

Powell, W. W. (1990). Neither market nor hierarchy: Network forms of organizations. *Research in Organizational Behavior*, volume 12:pages 295–336.

Prakken, H. (1997). *Logical Tools for Modelling Legal Arguments*. Kluwer, Dordrecht, The Netherlands.

Prakken, H. and Vreeswijk, G. (2002). Logics for defeasible argumentation. *Handbook of Philosophical Logic*, volume IV:pages 218–319. Second Edition.

Pufendorf, S. (1688). *De Jure Naturae et Gentium*. Clarendon Press, 1934 edition.

Quine, W. (1960). Variables explained away. *Proceedings of the American Philosophical Society*, volume 104(3):pages pp. 343–347.

Quine, W. V. (1971). Algebraic logic and predicate functors. In R. Rudner and I. Scheffler, editors, *Logic and art: lessons in honor of Nelson Goodman*. Bobbs-Merrill, Indianapolis.

Rao, A. S. and Georgeff, M. P. (1991). Modeling rational agents within a BDI-architecture. In J. Allen, R. Fikes and E. Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, pages 473–484. Morgan Kaufmann publishers Inc.: San Mateo, CA, USA.

Rawls, J. (1955). Two concepts of rules. *The Philosophical Review*, volume 64(1):pages 3–32.

Ricciardi, M. (1997). Constitutive rules and institutions. Paper presented at the meeting of the Irish Philosophical Club, Ballymascanlon.

Rosenschein, G. S., Wooldridge, M., Sandholm, T. and Yokoo, M., editors (2003). *The 2nd International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS 2003)*. ACM Press, Melbourne, Victoria, Australia.

Royakkers, L. and Dignum, F. (1997). Defeasible reasoning with legal rules. In D. Nute, editor, *Defeasible Deontic Logic*, pages 263–286. Kluwer, Dordrecht, The Netherlands.

Ruiter, D. W. P. (1997). A basic classification of legal institutions. *Ratio Juris*, volume 10:pages 357–371.

Ryle, G. (1949). *The Concept of Mind*. Penguin Books, London.

Sartor, G. (2006). Fundamental legal concepts: A formal and teleological characterization. *Artificial Intelligence and Law*, volume 14:pages 101–142.

Scerri, P., Xu, Y., Liao, E., Lai, J. and Sycara, K. (2004). Scaling teamwork to very large teams. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*. ACM PRess, New York, NY.

Schaerf, M. and Cadoli, M. (1995). Tractable reasoning via approximation. *Artificial Intelligence*, volume 74(2):pages 249–310.

Schild, K. (1991). A correspondence theory for terminological logics: preliminary report. In *Proceedings of IJCAI-91, 12th International Joint Conference on Artificial Intelligence*, pages 466–471. Sidney, AU.

Schoemaker, M. (2003). Identity in flexible organizations: Experiences. In *Dutch Organizations Creativity and Innovation Management*, volume 12.

Searle, J. (1969). *Speech Acts. An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, UK.

Searle, J. (1995). *The Construction of Social Reality*. Free Press.

Selznick, P. (1948). Foundations of the theory of organization. *American Sociological Review*, volume 13:pages 25–35.

Sergot, M. (2004). $(C+)^{++}$: An action language for modeling norms and institutions. Technical Report 8, Department of Computing, Imperial College, London.

Sergot, M. and Craven, R. (2006). The deontic component of action language $nC+$. In L. Goble and J.-J.Ch. Meyer, editors, *Proceedings of the 8th International Workshop on Deontic Logic in Computer Science, DEON 2006*, number 4048 in LNAI. Springer, Utrecht, The Netherlands.

Shoham, Y. (1991). Varieties of context. In *Artificial intelligence and mathematical theory of computation: papers in honor of John McCarthy*, pages 393–407. Academic Press Professional, Inc.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. J.M. Dent & Sons, London, UK, 1910 edition.

Smith, B. (2001). Fiat objects. *Topoi*, volume 2(20):pages 131–148.

So, Y. and Chon, K. (2005). A performance model for tree-structuresd multiagent organizations in faulty environments. In O. Boissier, V. Dignum, E. Matson and J. Sichman, editors, *Proceedings ofthe International Workshop on Organizations in Multi-Agent Systems (OOOP)*. AAMAS'05, Utrecht, The Netherlands.

Sørensen, A. B. (1978). Mathematical models in sociology. *Annual Review of Sociology*, volume 4:pages 345–371.

Stalnaker, R. (1998). On the representation of context. In *Journal of Logic, Language, and Information*, volume 7, pages 3–19. Kluwer.

Stimson, W. (1996). *The Robust Organization: Transforming Your Company Using Adaptive Design*. Irwin Professional Publishing.

Svetonius, T. C. (110). *De Vita Caesarum*. Project Gutemberg, 2003–2006 edition.

Tarski, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, volume 4(3):pages 341–376.

Tarski, A. (1983). The establishment of scientific semantics. In *Logic, Semantics, Metamathematics*, pages 401–408. Hackett, Indianapolis.

Thorburn, W. M. (1918). The myth of Occam's razor. *Mind*, volume 27:pages 345–353.

Varzi, A. (2000). Vague names for sharp objects. In L. Obrst and I. Mani, editors, *Proceedings of the Workshop on Semantic Approximation, Granularity, and Vagueness*, pages 73–78. AAAI Press.

Vázquez-Salceda, J. (2004). *The role of Norms and Electronic Institutions in Multi-Agent Systems*. Birkhäuser Verlag AG.

Vázquez-Salceda, J., Aldewereld, H. and Dignum, F. (2004). Norms in multiagent systems: from theory to practice. *International Journal of Computer Systems Science & Engineering*, volume 20(4):pages 95–114.

Von Wright, G. H. (1963). *Norm and Action. A Logical Inquiry*. Routledge, London.

Weiss, G. and Stone, P., editors (2006). *The 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006)*. ACM Press, Hakodate, Japan.

Wittgenstein, L. (1921). *Logisch-philosophische Abhandlung*, volume 14. Annalen der Naturphilosophie.

Wittgenstein, L. (1958). *Philosophical Investigations*. Blackwell, Oxford, UK.

Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *Knowledge Engineering Review*, volume 10(2):pages 115–152.

von Wright, G. H. (1951). Deontic logic. *Mind 60*, pages 1–15.

von Wright, G. H. (1983). Norms, truth and logic. In *Practical Reason: Philosophical Papers*, volume 1. Cornell University Press, Ithaca, New York.

Wygralak, M. (1996). *Vaguely Defined Objects*. Kluwer Academic Publishers.

Zambonelli, F., Jennings, N. and Wooldridge, M. (2003). Developing multiagent systems: the GAIA methodology. *ACM Transactions on Software Engineering and Methodology*, volume 12(3):pages 317–370.

Designing Invisible Handcuffs—*Summary*

The work presented in this thesis suggestively moves from the consideration that what makes social notions difficult to grasp lies probably in their intrinsic "invisibility". Like Smith's notorious "invisible hand", institutions and organizations are something which just cannot be pointed at by simply looking at the outside world. Still, they are there and they are almost ubiquitous. The prime research question of the work is therefore foundational. If we aim at a science of such things as institutions and organizations, how should we think of them? What concepts and conceptual apparatuses should we use to represent and reason about institutions and organizations? Or, to carry on the metaphor, how can we make those things somehow "visible"?

Interestingly, recent developments in the field of multi-agent systems (MAS) have clearly pointed at the need for an answer to such question. In MAS a number of autonomous pieces of software —the agents— interact in order to execute complex tasks. In such systems the crux of the matter is to design agents' interaction in such a way that, on the one hand, agents remain autonomous and, on the other hand, that the system exhibits global desirable properties. In human societies, institutions and organizations have developed as means for pursuing this exact aim. They set invisible boundaries —"handcuffs"— for the activities of the individuals in the society. If such "handcuffs" need to be designed in order to coordinate software agents, then a formal theory of institutions and organizations needs to be found which can ground the design process. Essentially, aim of the present work is to advance precise proposals for the development of such a formal theory.

From a methodological point of view, the work proceeds by first committing to precise views of institutions and organizations which can be found in the literature on social and legal sciences. Such views are then formalized and the resulting formal theory is finally discussed in its formal aspects as well as in the theoretical implications it bears for the notions of institution and organization thus analyzed.

The work presented conceives of institutions as systems of constitutive rules. Following Searle, constitutive rules are statements of the type "X counts as Y in context C" —the so-called counts-as statements— and they underlie the whole construction of institutional reality. It is our thesis that by means of these state-

ments institutional qualifications are *imposed* on the to-be-regulated domain which provide norms for agents' conduct. A typical example taken from the game of football could be the rule "the off-side situation counts as a violation of the rules of the game". Institutions can therefore be viewed as the imposition of complex conceptualizations specified in terms of counts-as statements. From a formal point of view, this suggests to represent institutions as description logic terminologies or taxonomical boxes (TBoxes). Taxonomical boxes are sets of subsumption statements describing the relations considered by the institutions to hold between the concepts it uses to conceptualize the to-be-regulated domain. The subsumption statement corresponding to the aforementioned counts-as statement would be the following one: "$\texttt{offside} \sqsubseteq \texttt{violation}$".

Of course, many different institutions coexist which might disagree on the way they look at the same domain. This motivates the formal analysis of the notion of context and the related one of contextual terminology which, from the point of view of the formal machinery deployed in our analysis, underpins the whole work presented here. Subsumptions are therefore studied as pertaining to a specific context: "Football: $\texttt{offside} \sqsubseteq \texttt{violation}$". In a nutshell, institutions can be thought of contextual terminologies, and counts-as statements as their basic building blocks.

As to organizations, the thesis focuses on their structural dimension. The notion of organization presupposes a notion of organizational structure, i.e., the structure specifying how the roles of the organization are related to one another (e.g., whether an authority relation holds between role r and s). On the grounds of foundational literature on the theory of organizations, the work presented stresses two essential aspects of organizational structures which are then both addressed from a formal point of view.

First, the structure of an organization is always multiple. That is to say, roles are connected by a number of different relations (who obeys whom? who communicates to whom? etc.), and not by just one as it is usually the case in network-based or chart-based representations of organizations. Such structures display, therefore, several different types of connections which we represent and study as multi-graphs, i.e., graphs containing links, or edges, of several different types. Second, the structure of an organization has a precise impact on the activities that the agents' taking part in the organization can engage in. In other words, the graph-theoretical dimension of an organizational structure has, so to say, a meaning in terms of the agents' activities it makes possible. It is shown that the formal machinery introduced for the analysis of institutions, constitutes also a viable formal tool for representing this semantic dimension of organizational structures. This combined perspective allows us to provide both quantitative methods based on graph-theory to compare different organizations from a structural point of view, and qualitative ones based on logic to address the types of interaction which different organizations put in place.

Finally, on the grounds of these results, a comparison of the two notions of institution and organization is provided, which makes explicit the different aspects that each of these two notions stresses in the conceptualization of social interaction between agents.

Ontwerpen van Onzichtbare Handboeien—*Samenvatting*

Het hier gepresenteerde werk begint met de opmerking dat de moeilijkheid om sociale noties te begrijpen komt door hun intrinsieke onzichtbaarheid. Net zoals de "onzichtbare hand" van Adam Smith, zijn instituties en organisaties iets waarnaar niet kan worden verwezen in de buitenwereld. Desondanks zijn ze er overal aanwezig. De eerste onderzoeksvraag van dit proefschrift betreft dus de grondslagen van een wetenschappelijke benadering van deze noties. Als wij naar een wetenschap van instituties en organisaties streven, hoe moeten wij dan tegen die noties aan kijken? Wat voor begrippen moeten we gebruiken om instituties en organisaties te representeren en er over te redeneren? Of, om verder met de metafoor te gaan, hoe kunnen wij ze "zichtbaar"maken?

Interessant genoeg, hebben recente ontwikkelingen in het veld van multi-agent systemen (MAS) ook de behoefte naar een antwoord voor zulke vragen benadrukt. In MAS interacteren een aantal autonome stukken software —de agenten— om complexe taken uit te oefenen. In zulke systemen is de crux van het verhaal het ontwerpen van interactie patronen tussen agenten zodanig dat, aan de ene kant, agenten autonoom blijven en, aan de andere kant, dat het systeem in het geheel aan bepaalde gewenste eigenschappen voldoet. In menselijke maatschappijen zijn instituties en organisaties juist ontworpen om dit doel te bereiken. Zij leggen onzichtbare grenzen —"handboeien"— aan de activiteiten van de individuën in de maatschappij. Als zulke "handboeien" moeten worden ontworpen om software agenten te coördineren, dan is een formele theorie nodig die de grondslagen voor dit ontwerp proces kan bieden. Het doel van dit proefschrift is dus om een dergelijke formele theorie voor te stellen.

Wat de methodologie betreft kiezen wij eerst voor een precieze visie van instituties en organisaties die in de literatuur van sociale theorie en rechtsleer kan worden gevonden. Deze visie wordt daarna geformaliseerd en de resulterende theorie wordt uiteindelijk bediscussieerd in zowel haar formele aspecten als de theoretische gevolgen die zij voor de analyse van instituties en organisaties heeft.

Het hier gepresenteerde werk beschouwt instituties als systemen van constitutieve regels. Volgens Searle zijn constitutieve regels zinnen van het soort "X telt als Y in context C-–de zogenaamde counts-as regels— en zij liggen onder de hele constructie van de institutionele werkelijkheid. Door middel van deze regels worden er institutionele eigenschappen aan de werkelijkheid gekoppeld die normen beschrijven voor de activiteiten van de agenten. Een typisch voorbeeld uit de voetbalwereld is "een buitenspel situatie telt als een overtreding van de regels van het voetbalspel". Instituties kunnen dus worden gezien als het opleggen van institutionele conceptualisaties die uitgedrukt worden in de vorm van counts-as regels. Vanuit een formeel standpunt suggereert dit idee het representeren van instituties als terminologieën of taxonomische "dozen" ("taxonomical boxes", TBoxes). TBoxes zijn verzameling terminologische subsumpties die de logische relaties beschrijven tussen de begrippen die de institutie gebruikt om de werkelijkheid te conceptualiseren. Een terminologische subsumptie die de buitenspelregel formaliseert is de volgende: "$\mathtt{buitenspel} \sqsubseteq \mathtt{overtreding}$".

Het spreekt vanzelf dat meerdere instituties tegelijkertijd kunnen bestaan die op verschillende manieren tegen het zelfde domein aankijken. Dit motiveert de formele analyse van de notie van context en van de gerelateerde notie van contextuele terminologie die, vanuit het standpunt van de formele machinerie die hier wordt gebruikt, het hele werk onderbouwt. Subsumpties worden dus gekoppeld aan bepaalde contexten: "Voetbal: $\mathtt{buitenspel} \sqsubseteq \mathtt{overtreding}$". Kort gezegd, kunnen instituties worden gezien als contextuele terminologieën, en counts-as regels als hun basale elementen.

Wat organisaties betreft focust dit proefschrift over hun structurele dimensie. De notie van organisatie veronderstelt de notie van organisatiestructuur, i.e., de structuur die specificeert hoe de rollen binnen de organisatie aan elkaar gebonden zijn (e.g., of er een autoriteit relatie bestaat tussen rol r en rol s). Op de basis van literatuur over de grondslagen van organisatietheorie behandelt het proefschrift twee essentiële aspecten van organisatiestructuren vanuit een formeel perspectief.

Ten eerste is de structuur van een organisatie altijd meervoudig. Met andere woorden rollen zijn altijd aan elkaar gebonden via meerdere verschillende relaties (wie gehoorzaamt wie? Wie communiceert met wie?, etc.), en niet door een unieke relatie zoals het in netwerkengebaseerde analyses normaal wordt gedaan. Zulke structuren tonen dus verschillende soorten connecties die wij als multi-grafen representeren, i.e, grafen die verbindingen bevatten van verschillende typen. Ten tweede heeft de structuur van een organisatie een specifieke invloed over de activiteiten van de agenten die deel van de organisatie uitmaken. Anders gezegd heeft de structurele dimensie van een organisatie altijd een betekenis in termen van de activiteiten die zij mogelijk maakt voor de agenten. Wij laten zien dat de formele machinerie die voor de analyse van instituties is geïntroduceerd, van toepassing is voor de representatie van deze semantische dimensie van organisatiestructuren. Dit gecombineerde perspectief biedt ons zowel kwantitatieve methoden, die gebaseerd zijn op grafentheorie, voor de structurele analyse van organisaties, als kwalitatieve methoden, die op logica zijn gebaseerd, voor de analyse van de interactiepatronen die verschillende organisaties tot stand brengen.

Tot slot worden, op basis van deze resultaten, de noties van institutie en organisatie vergeleken om expliciet te maken wat voor specifieke aspecten elk van deze twee noties benadrukt voor de conceptualisatie van sociale interactie.

Progettare Manette Invisibili—*Riassunto*

Il lavoro presentato in questa tesi muove suggestivamente dalla considerazione che ciò che rende le nozioni sociali di difficile comprensione risiede nella loro intrinsica "invisibilità". Come la notoria "mano invisibile" di Smith, le istituzioni e le organizzazioni sono qualcosa di esistente nel mondo esterno cui semplicemente non si può puntare il dito. Ciononostante sono in qualche modo là fuori, e sono onnipresenti. La principale domanda cui il lavoro presentato vuole dare una risposta è dunque di natura fondazionale. Se si punta allo sviluppo di una scienza delle istituzioni e delle organizzazioni, in che termini devono essere pensate? O, per continuare la metafora, come possono esser rese "visibili"?

Il bisogno di trovare risposte a tali domante è stato poi sottolineato anche da recenti sviluppi nel campo dei sistemi multi-agente (multi-agent systems, MAS). Nei MAS un numero di diversi software autonomi —i cosiddetti agenti— interagiscono in modo da eseguire compiti complessi. In tali sistemi il problema principale consiste nel progettare l'interazione tra gli agenti in modo tale che, da un lato, gli agenti rimangano autonomi e, dall'altro, che il sistema esibisca certe desiderabili proprietà ad un livello globale. Nelle società umane, istituzioni ed organizzazioni si sono sviluppate esattamente come mezzi per questo fine. Esse impongono limiti invisibili —"manette"— alle azioni dei singoli individui costituenti la società. Se simili "manette" devono essere progettate per consentire la appropriata coordinazione di agenti software, allora una teoria formale delle istituzioni e delle organizzazioni diventa necessaria per fondare tale progettazione. Essenzialmente, scopo del presente lavoro è quello di avanzare proposte precise per lo sviluppo di una tale teoria formale.

Dal punto di vista metodologico. il lavoro presentato procede, in prima istanza, attraverso la selezione di precise concezioni di istituzioni ed organizzazioni che possono essere riscontrate nella letteratura delle scienze legali e sociali. Tali concezioni sono poi formalizzate e le teoria che ne risulta è infine discussa nei suoi aspetti formali, come nelle conseguenze teoretiche che comporta per la comprensione delle nozioni di istituzione ed organizzazione cosí analizzate.

Le istituzioni vengono qui concepite come sistemi di regole costitutive. Seguendo Searle, le regole costitutive consistono in enunciati del tipo X conta come Y nel contesto C —i cosiddetti enunciati di counts-as— e sottendono l'intera costruzione della realtà istituzionale. È nostra tesi che, attraverso questo tipo di enunciati, qualificazioni istituzionali vengono *imposte* sul dominio che una certa istituzione intende regolare, fornendo cosí descrizioni di norme di condotta per gli agenti interessati. Un tipico esempio tratto dal gioco del calcio è l'enunciato "la situazione di fuorigioco conta come una violazione delle regole del gioco del calcio". Le istituzioni possono dunque essere viste come l'imposizione di concettualizzazioni complesse specificate attraverso enunciati di counts-as. Da un punto di vista formale questa concezione suggerisce la rappresentazione di istituzioni come terminologie in logica descrittiva o scatole tassonomiche (taxonomical boxes, TBoxes), cioè insiemi di sussunzioni descriventi le relazioni logiche che l'istituzione considera sussistere tra i concetti da essa utilizzati per concettualizzare il dominio che intende regolare. La sussunzione corrispondente all'enunciato di counts-as menzionato sopra è il seguente: $\mathtt{fuorigioco} \sqsubseteq \mathtt{violazione}$.

Ovviamente, istituzioni diverse possono coesistere e al contempo non concordare sulla concettualizzazione di uno stesso dominio. Ciò motiva l'analisi formale della nozione di contesto e di terminologia contestuale che, dal punto di vista del macchinario formale utilizzato nella nostra analisi, fonda l'intero lavoro qui presentato. Le sussunzioni vengono cosí studiate in quanto pertinenti uno specifico contesto: "Calcio: $\mathtt{fuorigioco} \sqsubseteq \mathtt{violazione}$. Riassumendo, le istituzioni possono essere pensate come terminologie contestuali, e gli enunciati di counts-as come come ciò da cui tali terminologie sono costituite.

Quanto alle organizzazioni, la tesi si concentra sulla loro dimensione strutturale. Il concetto di organizzazione presuppone quello di struttura organizzativa, ovverosia, la struttura che specifica come i ruoli della organizzazione sono tra loro collegati (ad esempio, se sussiste una relazione di autorità tra il ruolo r e il ruolo s). Sulla base di lavori fondazionali sulla teoria delle organizzazioni, vengono qui affrontati formalmente due aspetti essenziali delle strutture organizzative.

Primo, la struttura di un'organizzazione è sempre multipla. Ciò sta a dire che i ruoli dell'organizzazione sono connessi da svariati tipi di relazioni (chi obbedisce a chi? chi comunica con chi? etc.), e non esclusivamente da uno come di solito accade nelle rappresentazione delle organizzazioni in forma di reti o diagrammi. Tali strutture esibiscono quindi tipi diversi di connessioni che vengono qui rappresentate e studiate come multi-grafi, ovverosia grafi contenenti, appunto, diversi tipi di nodi. Secondo, la struttura di un'organizzazione possiede un preciso impatto sulle attività degli agenti prendenti parte all'organizzazione che può essere studiato in logica attraverso sistemi di transizioni (transition systems). Tale prospettiva combinata consente sia lo sviluppo di metodi quantitativi, basati sulla teoria dei grafi, per la comparazione dal punto di vista strutturale di diverse organizzazioni, che, al contempo, lo sviluppo di metodi qualitativi, basati sulla logica, per l'analisi dei diversi tipi di interazione realizzati da organizzazioni diverse.

Da ultimo, e sulla base dei risultati presentati, il lavoro offre anche un confronto tra le due nozioni di istituzione ed organizzazione sottolineando i diversi aspetti che ciascuna di queste nozioni enfatizza nella concettualizzazione dell'interazione sociale tra agenti.

Curriculum Vitae

Davide Grossi

20-04-1979	Born in Udine, Italy.
1993 - 1998	High school studies, *"Liceo Scientifico Nicolò Copernico"* (Scientific gymnasium), Udine, Italy.
1998 - 2003	University studies in Philosophy, *"Università di Pisa"*, Pisa, Italy; Thesis on Logic and Ethics.
1998 - 2003	University studies in Philosophy, *"Scuola Normale Superiore"* (University center for teaching and research), Pisa, Italy; Thesis on Argumentation Theory.
2003 - 2007	Ph.D. student in Artificial Intelligence, *"Universiteit Utrecht"*, Utrecht, The Netherlands.

SIKS Dissertation Series

1998

Johan van den Akker, *DEGAS - An Active, Temporal Database of Autonomous Objects*, CWI, 1998-1

Floris Wiesman, *Information Retrieval by Graphically Browsing Meta-Information*, UM, 1998-2

Ans Steuten, *A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective*, TUD, 1998-3

Dennis Breuker, *Memory versus Search in Games*, UM, 1998-4

E.W. Oskamp, *Computerondersteuning bij Straftoemeting*, RUL, 1998-5

1999

Mark Sloof, *Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products*, VU, 1999-1

Rob Potharst, *Classification using decision trees and neural nets*, EUR, 1999-2

Don Beal, *The Nature of Minimax Search*, UM, 1999-3

Jacques Penders, *The practical Art of Moving Physical Objects*, UM, 1999-4

Aldo de Moor, *Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems*, KUB, 1999-5

Niek J.E. Wijngaards, *Re-design of compositional systems*, VU, 1999-6

David Spelt, *Verification support for object database design*, UT, 1999-7

Jacques H.J. Lenting, *Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation.*, UM, 1999-8

2000

Frank Niessink, *Perspectives on Improving Software Maintenance*, VU, 2000-1

Koen Holtman, *Prototyping of CMS Storage Management*, TUE, 2000-2

Carolien M.T. Metselaar, *Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief*, UVA, 2000-3

Geert de Haan, *ETAG, A Formal Model of Competence Knowledge for User Interface Design*, VU, 2000-4

Ruud van der Pol, *Knowledge-based Query Formulation in Information Retrieval*, UM, 2000-5

Rogier van Eijk, *Programming Languages for Agent Communication*, UU, 2000-6

Niels Peek, *Decision-theoretic Planning of Clinical Patient Management*, UU, 2000-7

Veerle Coup, *Sensitivity Analyis of Decision-Theoretic Networks*, EUR, 2000-8

Florian Waas, *Principles of Probabilistic Query Optimization*, CWI, 2000-9

Niels Nes, *Image Database Management System Design Considerations, Algorithms and Architecture*, CWI, 2000-10

Jonas Karlsson, *Scalable Distributed Data Structures for Database Management*, CWI, 2000-11

2001

Silja Renooij, *Qualitative Approaches to Quantifying Probabilistic Networks*, UU, 2001-1

Koen Hindriks, *Agent Programming Languages: Programming with Mental Models*, UU, 2001-2

Maarten van Someren, *Learning as problem solving*, UvA, 2001-3

Evgueni Smirnov, *Conjunctive and Disjunctive*

Version Spaces with Instance-Based Boundary Sets, UM, 2001-4

Jacco van Ossenbruggen, *Processing Structured Hypermedia: A Matter of Style*, VU, 2001-5

Martijn van Welie, *Task-based User Interface Design*, VU, 2001-6

Bastiaan Schonhage, *Diva: Architectural Perspectives on Information Visualization*, VU, 2001-7

Pascal van Eck, *A Compositional Semantic Structure for Multi-Agent Systems Dynamics*, VU, 2001-8

Pieter Jan 't Hoen, *Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes*, RUL, 2001-9

Maarten Sierhuis, *Modeling and Simulating Work Practice BRAHMS: a multiagent modeling and simulation language for work practice analysis and design*, UvA, 2001-10

Tom M. van Engers, *Knowledge Management: The Role of Mental Models in Business Systems Design*, VUA, 2001-11

2002

Nico Lassing, *Architecture-Level Modifiability Analysis*, VU, 2002-01

Roelof van Zwol, *Modelling and searching web-based document collections*, UT, 2002-02

Henk Ernst Blok, *Database Optimization Aspects for Information Retrieval*, UT, 2002-03

Juan Roberto Castelo Valdueza, *The Discrete Acyclic Digraph Markov Model in Data Mining*, UU, 2002-04

Radu Serban, *The Private Cyberspace Modeling Electronic Environments inhabited by Privacy-concerned Agents*, VU, 2002-05

Laurens Mommers, *Applied legal epistemology; Building a knowledge-based ontology of the legal domain*, UL, 2002-06

Peter Boncz, *Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications*, CWI, 2002-07

Jaap Gordijn, *Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas*, VU, 2002-08

Willem-Jan van den Heuvel, *Integrating Modern Business Applications with Objectified Legacy Systems*, KUB, 2002-09

Brian Sheppard, *Towards Perfect Play of Scrabble*, UM, 2002-10

Wouter C.A. Wijngaards, *Agent Based Modelling of Dynamics: Biological and Organisational Applications*, VU, 2002-11

Albrecht Schmidt, *Processing XML in Database Systems*, UVA, 2002-12

Hongjing Wu, *A Reference Architecture for Adaptive Hypermedia Applications*, TUE, 2002-13

Wieke de Vries, *Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems*, UU, 2002-14

Rik Eshuis, *Semantics and Verification of UML Activity Diagrams for Workflow Modelling*, UT, 2002-15

Pieter van Langen, *The Anatomy of Design: Foundations, Models and Applications*, VU, 2002-16

Stefan Manegold, *Understanding, Modeling, and Improving Main-Memory Database Performance*, UVA, 2002-17

2003

Heiner Stuckenschmidt, *Onotology-Based Information Sharing In Weakly Structured Environments*, VU, 2003-1

Jan Broersen, *Modal Action Logics for Reasoning About Reactive Systems*, VU, 2003-02

Martijn Schuemie, *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*, TUD, 2003-03

Milan Petkovic, *Content-Based Video Retrieval Supported by Database Technology*, UT, 2003-04

Jos Lehmann, *Causation in Artificial Intelligence and Law - A modelling approach*, UVA, 2003-05

Boris van Schooten, *Development and specification of virtual environments*, UT, 2003-06

Machiel Jansen, *Formal Explorations of Knowledge Intensive Tasks*, UvA, 2003-07

Yongping Ran, *Repair Based Scheduling*, UM, 2003-08

Rens Kortmann, *The resolution of visually guided behaviour*, UM, 2003-09

Andreas Lincke, *Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture*, UvT, 2003-10

Simon Keizer, *Reasoning under Uncertainty in Nat-*

ural Language Dialogue using Bayesian Networks, UT, 2003-11

Roeland Ordelman, *Dutch speech recognition in multimedia information retrieval*, UT, 2003-12

Jeroen Donkers, *Nosce Hostem - Searching with Opponent Models*, UM, 2003-13

Stijn Hoppenbrouwers, *Freezing Language: Conceptualisation Processes across ICT-Supported Organisations*, KUN, 2003-14

Mathijs de Weerdt, *Plan Merging in Multi-Agent Systems*, TUD, 2003-15

Menzo Windhouwer, *Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouses*, CWI, 2003-16

David Jansen, *Extensions of Statecharts with Probability, Time, and Stochastic Timing*, UT, 2003-17

Levente Kocsis, *Learning Search Decisions*, UM, 2003-18

2004

Virginia Dignum, *A Model for Organizational Interaction: Based on Agents, Founded in Logic*, UU, 2004-01

Lai Xu, *Monitoring Multi-party Contracts for E-business*, UvT, 2004-02

Perry Groot, *A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving*, VU, 2004-03

Chris van Aart, *Organizational Principles for Multi-Agent Architectures*, UVA, 2004-04

Viara Popova, *Knowledge discovery and monotonicity*, EUR, 2004-05

Bart-Jan Hommes, *The Evaluation of Business Process Modeling Techniques*, TUD, 2004-06

Elise Boltjes, *Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes*, UM, 2004-07

Joop Verbeek, *Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politiële gegevensuitwisseling en digitale expertise*, UM, 2004-08

Martin Caminada, *For the Sake of the Argument; explorations into argument-based reasoning*, VU, 2004-09

Suzanne Kabel, *Knowledge-rich indexing of learning-objects*, UVA, 2004-10

Michel Klein, *Change Management for Distributed Ontologies*, VU, 2004-11

The Duy Bui, *Creating emotions and facial expressions for embodied agents*, UT, 2004-12

Wojciech Jamroga, *Using Multiple Models of Reality: On Agents who Know how to Play*, UT, 2004-13

Paul Harrenstein, *Logic in Conflict. Logical Explorations in Strategic Equilibrium*, UU, 2004-14

Arno Knobbe, *Multi-Relational Data Mining*, UU, 2004-15

Federico Divina, *Hybrid Genetic Relational Search for Inductive Learning*, VU, 2004-16

Mark Winands, *Informed Search in Complex Games*, UM, 2004-17

Vania Bessa Machado, *Supporting the Construction of Qualitative Knowledge Models*, UvA, 2004-18

Thijs Westerveld, *Using generative probabilistic models for multimedia retrieval*, UT, 2004-19

Madelon Evers, *Learning from Design: facilitating multidisciplinary design teams*, Nyenrode, 2004-20

2005

Floor Verdenius, *Methodological Aspects of Designing Induction-Based Applications*, UVA, 2005-01

Erik van der Werf, *AI techniques for the game of Go*, UM, 2005-02

Franc Grootjen, *A Pragmatic Approach to the Conceptualisation of Language*, RUN, 2005-03

Nirvana Meratnia, *Towards Database Support for Moving Object data*, UT, 2005-04

Gabriel Infante-Lopez, *Two-Level Probabilistic Grammars for Natural Language Parsing*, UVA, 2005-05

Pieter Spronck, *Adaptive Game AI*, UM, 2005-06

Flavius Frasincar, *Hypermedia Presentation Generation for Semantic Web Information Systems*, TUE, 2005-07

Richard Vdovjak, *A Model-driven Approach for Building Distributed Ontology-based Web Applications*, TUE, 2005-08

Jeen Broekstra, *Storage, Querying and Inferencing for Semantic Web Languages*, VU, 2005-09

Anders Bouwer, *Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments*, UVA, 2005-10

Elth Ogston, *Agent Based Matchmaking and Clustering - A Decentralized Approach to Search*, VU, 2005-11

Csaba Boer, *Distributed Simulation in Industry*, EUR, 2005-12

Fred Hamburg, *Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen*, UL, 2005-13

Borys Omelayenko, *Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics*, VU, 2005-14

Tibor Bosse, *Analysis of the Dynamics of Cognitive Processes*, VU, 2005-15

Joris Graaumans, *Usability of XML Query Languages*, UU, 2005-16

Boris Shishkov, *Software Specification Based on Re-usable Business Components*, TUD, 2005-17

Danielle Sent, *Test-selection strategies for probabilistic networks*, UU, 2005-18

Michel van Dartel, *Situated Representation*, UM, 2005-19

Cristina Coteanu, *Cyber Consumer Law, State of the Art and Perspectives*, UL, 2005-20

Wijnand Derks, *Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics*, UT, 2005-21

2006

Samuil Angelov, *Foundations of B2B Electronic Contracting*, TUE, 2006-01

Cristina Chisalita, *Contextual issues in the design and use of information technology in organizations*, VU, 2006-02

Noor Christoph, *The role of metacognitive skills in learning to solve problems*, UVA, 2006-03

Marta Sabou, *Building Web Service Ontologies*, VU, 2006-04

Cees Pierik, *Validation Techniques for Object-Oriented Proof Outlines*, UU, 2006-05

Ziv Baida, *Software-aided Service Bundling – Intelligent Methods & Tools for Graphical Service Modeling*, VU, 2006-06

Marko Smiljanic, *XML schema matching – balancing efficiency and effectiveness by means of clustering*, UT, 2006-07

Eelco Herder, *Forward, Back and Home Again – Analyzing User Behavior on the Web*, UT, 2006-08

Mohamed Wahdan, *Automatic Formulation of the Auditor's Opinion*, UM, 2006-09

Ronny Siebes, *Semantic Routing in Peer-to-Peer Systems*, VU, 2006-10

Joeri van Ruth, *Flattening Queries over Nested Data Types*, UT, 2006-11

Bert Bongers, *Interactivation – Towards an e-cology of people, our technological environment, and the arts*, VU, 2006-12

Henk-Jan Lebbink, *Dialogue and Decision Games for Information Exchanging Agents*, UU, 2006-13

Johan Hoorn, *Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change*, VU, 2006-14

Rainer Malik, *CONAN: Text Mining in the Biomedical Domain*, UU, 2006-15

Carsten Riggelsen, *Approximation Methods for Efficient Learning of Bayesian Networks*, UU, 2006-16

Stacey Nagata, *User Assistance for Multitasking with Interruptions on a Mobile Device*, UU, 2006-17

Valentin Zhizhkun, *Graph transformation for Natural Language Processing*, UVA, 2006-18

Birna van Riemsdijk, *Cognitive Agent Programming: A Semantic Approach*, UU, 2006-19

Marina Velikova, *Monotone models for prediction in data mining*, UvT, 2006-20

Bas van Gils, *Aptness on the Web*, RUN, 2006-21

Paul de Vrieze, *Fundaments of Adaptive Personalisation*, RUN, 2006-22

Ion Juvina, *Development of Cognitive Model for Navigating on the Web*, UU, 2006-23

Laura Hollink, *Semantic Annotation for Retrieval of Visual Resources*, VU, 2006-24

Madalina Drugan, *Conditional log-likelihood MDL and Evolutionary MCMC*, UU, 2006-25

Vojkan Mihajlovic, *Score Region Algebra: A Flexible Framework for Structured Information Retrieval*, UT, 2006-26

Stefano Bocconi, *Vox Populi: generating video documentaries from semantically annotated media repositories*, CWI, 2006-27

Borkur Sigurbjornsson, *Focused Information Access using XML Element Retrieval*, UVA, 2006-28

2007

Kees Leune, *Access Control and Service-Oriented Architectures*, UvT, 2007-01

Wouter Teepe, *Reconciling Information Exchange and Confidentiality: A Formal Approach*, RUG, 2007-02

Peter Mika, *Social Networks and the Semantic Web*, VU, 2007-03

Jurriaan van Diggelen, *Achieving Semantic Interoperability in Multi-agent Systems: A Dialogue-based Approach*, UU, 2007-04

Bart Schermer, *Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance*, UL, 2007-05

Gilad Mishne, *Applied Text Analytics for Blogs*, UVA, 2007-06

Natasa Jovanoviç, *To Who It May Concern - Addressee Identification in Face-to-Face Meetings*, UT, 2007-08

Mark Hoogendoom, *Modeling of Change in Multi-Agent Organizations*, VU, 2007-09

David Mobach, *Agent-Based Mediated Service Negotiation*, VU, 2007-09

Huib Aldewereld, *Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols*, UU, 2007-10

Natalia Stash, *Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System*, TUE, 2007-11

Marcel van Gerven, *Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty*, RUN, 2007-12

Rutger Rienks, *Meetings in Smart Environments; Implications of Progressing Technology*, UT, 2007-13

Niek Bergboer, *Context-Based Image Analysis*, UM, 2007-14

Joyca Lacroix, *NIM: a situated computational memory model* , UM, 2007-15

Davide Grossi, *Designing Invisible Handcuffs: Formal Investigations in Institutions and Organizations for Multi-agent Systems*, UU, 2007-16