A Behavioural Decision-Making Framework For Agent-Based Models

Een kader voor besluitvorming op basis van gedrag voor agentgebaseerde modellen
(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. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 20 februari 2023 des middags te 12.15 uur

door

Khoa Dang Nguyen

geboren op 9 juli 1993
te Ho Chi Minh City, Vietnam

Promotoren:

Prof. dr. F.P.M. Dignum
Prof. dr. R. Schumann

Beoordelingscommissie:

Prof. dr. L. Braubach
Prof. dr. M.M. Dastani
Prof. dr. Y. Demazeau
Prof. dr. F. Klugl
Prof. dr. L. Vanhee

Khoa Dang NGUYEN | CV

Experienced researcher with a demonstrated history of working in the higher education/research industry. Interested in Agent-based Modelling, Data Science and Machine Learning Technology.

Education Qualifications

- **PhD Candidate** — **Sierre, Switzerland**
  HES-SO Valais-Wallis and Utrecht University — *2017 – 2022*
  - Thesis title: *'A Behavioural Decision-Making Framework For Agent-Based Models'*.
- **MEng Computing** — **London, UK**
  Imperial College London — *2012 – 2016*
  - Specialised in Machine Learning, Data Mining and Large Scale Data Management.
- **A-level** — **London, UK**
  Ashbourne College — *2011 – 2012*
  - Awarded 50% entrance scholarship for outstanding achievement.

Academic Works

- **Doctoral project:** *Behavioural-Driven Demand Model (BedDeM) - PROBOUND*
  - Creating an extension for the Multiagent System simulation - BedDeM - to investigate the effects of bounded rationality imply on the purchase of vehicles and mobility services.
  - Coupling with the Swiss TIMES Energy System Model developed by the Paul Scherrer Institute to provide recommendations for energy policy design at both systematic and consumer levels.
- **Doctoral project:** *Behavioural-Driven Demand Model (BedDeM) - Mobility*
  - Developing a Multiagent System simulation for mobility and tourism demand in Switzerland based on Triandis' Theory of Interpersonal Behaviour.
  - Managing data directory and pipeline between models of different research teams in the SCCER-CREST Joint Activity: The evolution of mobility - A socio-economic analysis.
  - Publication chair for the DACH+ Conference on Energy Informatics 2020.
- **Master project:** *A Cognitive-driven System for Guiding The Learning of Genetic Regulatory Networks*
  - Created a framework to apply a new cognitive system - Watson to solve some of the problems with text mining supporting systems biology.
  - Developed a proximity search function to extract yeast gene names and their relationship from a set of paragraphs.
- **3rd year group project:** *'Spyder - Social Networking tool in Real life with Proximity Sensing'*
  - The team created an end-to-end solution using bluetooth sensors on user's smartphones to sense proximity to other users.
  - Responsible for providing a bridge between the user's smartphones and the website displaying the information in real time.

- **2nd year group project:** *'Puzzlin - Android Mobile Puzzle Application'*
  - Led the group to produce a multi-user Android application that encourages people to discover London while solving puzzle questions, maintaining a consistent internal state via use of a databases containing user, question and hint information.
  - In charge of setting up the back-end database and ensuring that the front-end and the back-end do not diverge from each other during the implementation of the above mentioned requirements.
- **Personal project:** *Dimension Jumper - Android Mobile Game*
  - Designed a game app using knowledge of Java and Android graphic library (OpenGL ES).
  - Available on Google Play: http://play.google.com/store/apps/details?id=com.kns.androidgames.dimensionjumper

Main Publications

K. Nguyen and R. Schumann, "A socio-psychological approach to simulate trust and reputation in modal choices," in *PRIMA 2020: Principles and Practice of Multi-Agent Systems* (T. Uchiya, Q. Bai, and I. MarsáMaestre, eds.), Lecture Notes in Computer Science, (Cham), pp. 39–54, Springer International Publishing, Feb. 2021.

K. Nguyen and R. Schumann, "A socio-psychological modal choice approach to modelling mobility and energy demand for electric vehicles," in *Proceedings of the 9th DACH+ Conference on Energy Informatics* (K. Nguyen, R. Roman, and R. Schumann, eds.), vol. 3 of *Energy Informatics*, Springer, Oct. 2020.

K. Nguyen and R. Schumann, "On developing a more comprehensive decision-making architecture for empirical social research: Lesson from agent-based simulation of mobility demands in switzerland," in *Multi-Agent-Based Simulation XX* (M. Paolucci, J. S. Sichman, and H. Verhagen, eds.), vol. XX of *International Workshop on Multi-Agent Systems and Agent-Based Simulation*, (Cham), pp. 39–54, Springer International Publishing, Nov. 2020.

K. Nguyen and R. Schumann, "An exploratory comparison of behavioural determinants in mobility modal choices," in *Advances in Social Simulation* (P. Ahrweiler and M. Neumann, eds.), (Cham), pp. 569–581, Springer International Publishing, 2021.

A. Bektas, K. Nguyen, V. Piana, and R. Schumann, "People-centric policies for decarbonization: Testing psycho-socio-economic approaches by an agent-based model of heterogeneous mobility demand," in *CEF 2018 24th Annual Conference on Computing in Economics and Finance*, June 2018.

A. Bektas, K. Nguyen, and R. Schumann, "Can agent-based computational economics mimic neoclassical demand curve?." Conference poster and presentation, presented at "Fifth International Symposium in Computational Economics and Finance", Apr. 2018.

Technical and Personal skills

- **Computing and Statistical skills:**
  - Agent-based modelling in **RePast** and **NetLogo**.
  - Programming experience in **Java**, **C++**, **Python**, **R**.
  - Proficient knowledge of **Object-Oriented Design**, **Data Management**, **GIT**, **Linux**, **MacOS** and **Window environment**.
  - Familiar with **Javascript** and **Web development**.
- **Languages**: English, Japanese, basic French

"Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves."

Herbert A. Simon

UTRECHT UNIVERSITY

English Abstract

Intelligent Software Systems
Information and Computing Sciences

Doctor of Philosophy

A Behavioural Decision-Making Framework For Agent-Based Models

by Khoa Dang NGUYEN

In the last decades, computer simulation has become one of the mainstream modelling techniques in many scientific fields. In particular, social simulation with Agent-based Modelling (ABM) allows users to capture higher-level system properties that emerge from the interactions of lower-level subsystems. This notion of emergence has introduced an alternative approach to studying the complexity inherited from social phenomena.

Several social simulation approaches were initially developed in the areas of physics and Artificial Intelligence. In fact, ABM is itself an area of application of Distributed Artificial Intelligence and Multiagent Systems (MAS). Despite that, researchers using ABM for social science studies do not fully benefit from the development in the field of MAS. It is mainly because the MAS architectures and frameworks are built upon cognitive and computer science foundations and principles, creating a gap in concepts and methodology between the two fields.

Building agent frameworks based on behaviour theory is a promising direction to minimise this gap. It can provide a standard practice in interdisciplinary teams and facilitate better usage of MAS technological advancement in social research. From our survey of different socio-psychology theories, Triandis' Theory of Interpersonal Behaviour (TIB) was chosen due to its broad set of determinants and inclusion of an additive value function to calculate utility values of different outcomes. As TIB's determinants can be organised in a tree-like structure, we utilise layered architectures to formalise the agent's components. The additive function of TIB is then used to combine the utilities of different level determinants.

The framework is then applied to create models for different case studies from various domains to test its ability to explain the importance of multiple behavioural aspects and environmental properties. The first case study simulates the mobility demand for Swiss households. We propose an experimental method to test and investigate the impact of core determinants in the TIB on the usage of different transportation modes. The second case study presents a novel solution to simulate trust and reputation by applying subjective logic as a metric to measure an agent's belief about the consequence(s) of action, which can be updated through feedback. By performing an experiment set up in the mobility domain, we demonstrate the

framework's ability to capture the ground truth of a service's reputation at different simulation scales and highlight the effects of these concepts on the figure of yearly rail kilometres travelled. The third case study investigates the possibility of simulating bounded rationality effects in an agent's decision-making scheme by limiting its capability of perceiving information. We demonstrate the functionality of this model in the context of purchasing vehicles in Switzerland's households. In the final study, a model is created to simulate migrants' choice of activities in centres by applying our framework in conjunction with Maslow's hierarchy of needs. The experiment can then be set up to test the impact of different combinations of core determinants on the migrants' activities.

Overall, the design of different components in our framework enables adaptations for various contexts, including transportation modal choice, buying a vehicle or daily activities. Most of the work can be done by changing the first-level determinants in the TIB's model based on the phenomena simulated and the available data. Several environmental properties can also be considered (e.g. static/dynamic, know/unknown, fully/partly observable, deterministic/stochastic) by extending the core components or employing other theoretical assumptions and concepts from the social study. With these procedures, the framework can serve the purpose of theoretical exposition and allow the users to assess the causal link between the TIB's determinants and behaviour output. This thesis also highlights the importance of data collection and experimental design to capture better and understand different aspects of human decision-making.

UNIVERSITEIT UTRECHT

Nederlands Abstract

Intelligente Software Systemen
Informatie- en computerwetenschappen

Doctor in de Filosofie

Een kader voor besluitvorming op basis van gedrag voor agentgebaseerde modellen

door Khoa Dang NGUYEN

In de afgelopen decennia worden computersimulaties steeds vaker gebruikt in vele wetenschappelijke gebieden. Op het gebied van de sociologie betreft dit met name sociale simulatie met Agent-based Modelling (ABM). Deze simulaties stellen gebruikers in staat om op hoog niveau systeemeigenschappen te beschrijven die voortkomen uit de interacties van subsystemen op een lager niveau. Het concept van emergentie heeft een alternatieve benadering geïntroduceerd voor het bestuderen van de complexiteit die inherent is aan sociale verschijnselen.

Verschillende aanpakken van sociale simulatie komen oorspronkelijk uit de natuurkunde en de kunstmatige intelligentie. Meer gebaseerd op toepasbare formele technieken dan de sociologische theorieen die ze proberen te onderbouwen. ABM is zelf een toepassingsgebied van Gedistribueerde Kunstmatige Intelligentie en Multiagent Systemen (MAS). Desondanks profiteren onderzoekers die ABM gebruiken voor sociologie niet ten volle van de ontwikkeling op het gebied van MAS. Dat komt vooral omdat de multi-agent systemen gebaseeerd zijn op informatica en individuele cognitieve theorieen. Hier door ontstaat er een kloof tussen de concepten en methodologie in ABM en sociologie.

Om deze kloof te overbruggen hebben we agent architecturen ontwikkeld die zijn gebaseerd op (sociale) gedragstheorie. Deze architecturen zorgen voor een gemeenschappelijke standaard in interdisciplinaire teams en maakt het makkelijker om nieuwe ontwikkelingen in MAS in de sociologie te gebruiken. Uit ons onderzoek bleek dat de Theorie van Interpersoonlijk Gedrag (TIG) van Triandis hier een goede kandidaat voor was vanwege zijn brede reeks aan invloeden en omdat het een additieve waardefunctie bevat om de gevolgen van verschillend gedrag te berekenen. Aangezien de invloeden van TIG in een hierarchische boomstructuur kunnen worden georganiseerd, kunnen we dezelfde structuur gebruiken om de beslis componenten van de agent in lagen te organiseren. De additiefunctie van TIG wordt dan gebruikt om het nut van de verschillende determinanten per laag te combineren.

We passen de ontwikkelde architectuur toe om modellen te creëren voor casestudies in verschillende domeinen. De eerste casestudy simuleert de mobiliteitsvraag van Zwitserse huishoudens. Wij stellen een experimentele methode voor om het effect

van kerndeterminanten in het TIG op het gebruik van verschillende vervoerswijzen te testen en te onderzoeken. De tweede casestudy presenteert een nieuwe oplossing om vertrouwen en reputatie te simuleren door subjectieve logica toe te passen als een metriek om de overtuiging van een agent over het (de) gevolg(en) van een actie te meten, die via feedback kan worden bijgewerkt. Door het uitvoeren van een experiment in het mobiliteitsdomein tonen wij aan dat het raamwerk de reëele basis van de reputatie van een dienst op verschillende simulatieschalen kan vastleggen. In de derde casestudy wordt de mogelijkheid onderzocht om de effecten van beperkte rationaliteit in de besluitvorming van een agent te simuleren door zijn vermogen om informatie waar te nemen te beperken. Wij demonstreren de functionaliteit van dit model in de context van de aankoop van voertuigen in Zwitserse huishoudens. In de laatste studie wordt een model gecreëerd om de keuze van activiteiten van migranten in centra te simuleren door ons kader toe te passen in combinatie met de behoeftenhiërarchie van Maslow. Vervolgens voeren we een experiment uit om de invloed van verschillende combinaties van kerndeterminanten op de activiteiten van de migranten te testen.

Door het gebruik van een modulaire agent architectuur kunnen we makkelijk aanpassingen maken voor verschillende contexten, zoals de keuze van de vervoerswijze, de aankoop van een voertuig of dagelijkse activiteiten. Het meeste werk kan worden gedaan door de invloeden op het hoogste niveau in het TIG-model te wijzigen op basis van de gesimuleerde verschijnselen en de beschikbare gegevens. Verschillende omgevingskenmerken kunnen ook in aanmerking worden genomen (bv. statisch/dynamisch, bekend/onbekend, geheel/gedeeltelijk waarneembaar, deterministisch/stochastisch) door de kerncomponenten uit te breiden of andere theorieën en concepten uit de sociologie te gebruiken. De uitkomsten van deze casestudies bewijzen dat de architectuur kan dienen als theoretische onderbouwing van de modellen en de gebruikers in staat stelt het causale verband tussen de determinanten van de TIG en het uiteindelijke gedrag te beoordelen. Dit proefschrift benadrukt ook het belang van gegevensverzameling en experimenteel ontwerp om verschillende aspecten van menselijke besluitvorming beter vast te leggen en te begrijpen.

Acknowledgements

The work of this thesis is based on different projects at HES-SO Valais-Wallis in Sierre, Switzerland. The topic of my work, a framework for agent decision-making, started when I took my position as a PhD student. Since then, it has taken longer than I expected to complete this thesis, and many significant life events have happened over the last five years, including a pandemic. Consequently, I must thank many people I have met during my research.

First of all, I would like to take this opportunity to express my deepest gratitude to my main superior Prof. Dr. René Schumann (HES-SO Valais-Wallis). He was the one who gave me my first job as PhD student after graduation. Thanks to this opportunity, I was able to move to Switzerland. It was not only a geographical change but also a new perspective on research. I am very grateful for all the valuable advice and encouragement during our meetings as well as outside of work.

I want to extend my gratitude to Prof. Dr. Frank Dignum (Umeå University). I met him during my first conference as a PhD student and several others afterwards. He offered me the opportunity to do this thesis at Utrecht and has since supervised me with great enthusiasm. I am grateful for his continuing advice and encouragement.

I also want to thank Prof. Dr. Lars Braubach (Universität Hamburg) for his support during the first phase of this thesis. Although I could not complete it due to unforeseen circumstance, his hospitality during my visit and continued support was helpful for many years of my endeavour.

I want to thank Valentino Piana (HES-SO Valais-Wallis) for all his work and encouragement during my time at HES-SO. His knowledge and kindness are, without saying, the best in the field.

A special thank you goes to my old colleague Alperen Bektas for his cooperation in our joint projects and, subsequently, my case studies.

I want to thank Dr. Davide Calvaresi for many insightful discussions concerning academic work and career.

A special thank you goes to my current and old colleagues in the SILab group, including Caroline Taramarcaz, Michael Barry, Audrey Dupont and Geovanny Poveda. They have supported me during my work as a PhD.

Also, I would like to thank my friends and colleagues at Techno-Pôle, who worked with me at different times of their careers. Our daily coffee conversation and invitations to various occasions were something I enjoyed during my PhD work. In addition, their willingness to listen to my endless racing conversation is not to be underappreciated.

My family in Vietnam and USA are my daily inspiration to work hard and appreciate what I have. I want to express my deepest gratitude to my parents for their ongoing encouragement and readiness to listen to my stories through the weekly phone calls. Their advice and encouragement throughout this process will go with me for the rest of my life. I would like to thank my brother for the first lesson in programming. His doctoral achievement is exemplary work for me to follow.

All people above have influenced me one way or another, and consequently, how this thesis looks. However, there would be no complete thesis without the love and support from Kevin. During the most challenging time, he showed me the path when I got lost and was the shoulder for me to learn whenever I was tired. For this, I am forever grateful for all the time we shared and hope for more to come.

Contents

List of Figures

List of Tables

List of Abbreviations

ABM	Agent-based Modelling
ABSS	Agent-based Social Simulation
ACT-R	Adaptive Control of Thought-Rational
AI	Artificial Intelligence
AUML	Agent Unified Modeling Language
BDI	Believe-Desire-Intention
BedDeM	Behaviour-Driven Demand Model
BOID	Beliefs-Desires-Obligations-Intentions
BR	Bounded Rationality
CLARION	Connectionist Learning with Adaptive Rule Induction ON-line
DAI	Distributed Artificial Intelligence
DCM	Discrete Choice Modelling
eBDI	Emotional-Beliefs-Desires-Obligations-Intentions
IAC	Integrative Agent-Centered
KQML	Knowledge Query and Manipulation Language
MABS	Multiagent-Based Simulation
MAS	Multiagent System
MEU	Maximum Expected Utility
ML	Machine Learning
MoHuB	Modelling Human Behaviour
MTMC	Transport Microcensus
NoA	Normative Agent
PRS	Procedural Reasoning System
SEU	Subjective Expected Utility
SHEDS	Swiss Household Energy Demand Survey
TIB	Theory of Interpersonal Behaviours
TPB	Theory of Planned Behaviour
TRA	Theory of Reasoned Action
UML	Unified Modeling Language

Chapter 1

Introduction

1.1 Motivation

As a society, we face many social and political challenges, such as changes in the population structure (due to birthrate and migration), financial and economic instability and threats against health (e.g. epidemics and environment hazard) [Con+12]. To provide a better understanding of the social processes underlying these challenges, scientists have begun to classify and categorise them into complex systems [Con+12; HB11a]. These systems are collections of many individuals that interact with each other, motivated by their own beliefs and personal goals together with the circumstances of their social environment. They are examples of non-linear systems, in which the change of the output is not proportional to the change of the input. Hence, they are difficult to be studied using traditional statistical or mathematical techniques [GT00], such as regression analysis [DS98], probability distributions [GKO05] or statistical inferences [Sil17]. In many of these instances, it is the case where no analytically traceable set of equations can be devised to describe the system without making assumptions or estimations based on incomplete knowledge of the phenomenon [GT00; PD02].

Social simulation has introduced a new way of thinking about complex systems, based on the ideas about the *emergence of social patterns* from relatively simple activities [Sim96]. These simulations often involves the usage of Agent-Based Model (ABM) to build an artificial society with software *agents*. The main idea is to simulate the operations and interactions of multiple agents following specific behavioural rules (micro-level) to search for explanatory or exploratory insights into social phenomena (macro-level). These models have helped shift analyses away from structural and aggregate factors to the role of individuals [MW02; BS15].

As one of the forerunners work in this field, Schelling's work in 1971 was among the first models to show how global properties may emerge from local interactions between individuals that have tendencies to refer to their neighbours [Sch71]. It was especially useful for the study of residential segregation of ethnic groups where agents represented householders who relocate to the city. In the 1980s, Robert Axelrod held a computer tournament where people were invited to submit strategies for the Prisoner's Dilemma and had them play in an agent-based manner to determine a winner [AH81]. The author continued to develop several other agent-based models in the field of political sciences [Axe97]. Another model developed by Craig Reynold in the 1980s was one of the first biological agent-based models that contained social characteristics [Rey87]. In addition, Joshua M. Epstein and Robert Axtell created one of the notable large-scale ABM, Sugarscape, to investigate the role of social phenomena, including pollution, the transmission of disease, sexual reproduction and seasonal migration [EA96].

In recent decades, social and life science researchers have become increasingly interested in ABM due to its potential assistance in discovering and analysing social mechanisms [CP14]. For example, Cohen et al. studied the role of direct reciprocity (i.e., a form of conditional cooperation between related or unrelated individuals) on an iterated Prisoner's Dilemma [CRA01]. To investigate social convention, Hodgson and Knudsen modelled an agent population randomly located in a cell ring that had to decide whether to drive clockwise or counter-clockwise around a ring to avoid collision[HK04]. Similarly, Epstein built a model to investigate the link between the strength of a convention and the cognitive costs that individuals have to pay to decide what to do [Eps01]. Hedström and Åberg built an empirically calibrated ABM to examine how social influence mechanisms can explain aggregate youth unemployment rates [HÅ05]. Interested readers can find many more examples in surveys such as [MW02; BS15; BA15].

Beyond academic interest, there is a practical need for models that allow for contingent behaviour of individuals and feedback between micro and macro levels of analysis, most notably in the fields of epidemiology and public policies (e.g. [AG13; BS14; HMNK17]). ABM can become a laboratory for researchers to manipulate the micro-level parameters in a controlled way (according to an experimental procedure in which the baseline condition is calibrated on empirical observations) and allow counter-factual or *if* conditions to be explored in alternative/artificial scenarios. Therefore, ABM can decrease the costs of quantitative manipulations to explore hypothetical scenarios and time-spaces that would be significantly difficult to study or observe empirically [BA15].

However, the most popular and highly-cited method of ABM often employ ad-hoc, simple *condition-action* rules based on theoretical assumptions or derived over statistical distributions, mainly due to their objective of giving insights on possible explanations for general social patterns (see surveys such as [Gro+17; DD19] and discussion in [Gil04; MLT09; Sun+16]). This highly abstract design leads to one of the main criticisms of ABM: unrealistic decision-making process, which provides limited understanding of the causal mechanisms in the agent' actions [WW13; Man14; Nap18]. To illustrate this point, we introduce an example in the field of mobility, which is based on our first case study presented in Chapter 5. Our main objective is building a model for the simulation of daily mobility choices. We are interested in the roles of different decision-making determinants (e.g. attitude, emotion, habit). An agent represents an individual or a household that has to choose between a complete set of exclusive alternatives (e.g. which vehicle or service to use). Each option can be given a utility value from an agent's evaluation process based on its parameters and the decision-making preferences.

In Figure 1.1, an agent has to go from home to work. It has several options, including (1) driving a car, (2) taking a bus or (3) cycling. Combinations of different modes of transport are also possible, e.g. (4) driving or (5) cycling to a station and then taking a train. These potential options can be evaluated under different determinants, such as time, cost, enjoyment, environmental friendliness, sociability, etc. An agent can make a decision based on the following aspects or strategies:

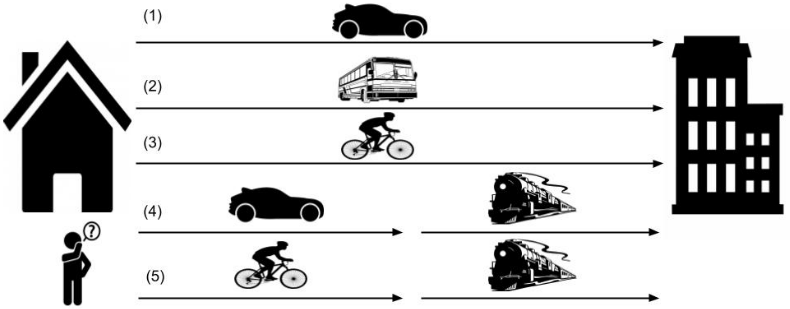


FIGURE 1.1: A mobility choice example

- **Economic attitudes**: Classical economists typically assume that behaviours are motivated primarily by material incentives and that decisions are governed mainly by self-interest and rationality [KS02]. In other words, agents are expected to maximise certain profits, revenue, or rate of profit while not violating any constraints. According to this theory, our agent prioritises the cost and time determinants. If the fuel price is low, driving alone (1) and a combination of car and train (4) would be the top options. Another possibility is applying prospect theory, which also introduces another important aspect from cognitive psychology to the rational actor model - people's willingness to seek or avoid risk influences their decisions [Kah79]. In this case, agents have a degree of risk aversion, whereby they bias decisions towards avoiding loss. So, for example, if our agent wants to avoid being late due to a traffic jam, it could choose the option to take public transport (e.g. options (4) or (5)) for a more accurate journey time. A more elaborate model can have the agent deliberating to update its belief about the environment.

- **Affective**: We could consider the amount of leisure the transportation mode can provide according to the agent's perception [BDF16]. If the agent believes driving a car is the most comfortable way to get to work, (1) is the best alternative.

- **Social**: There are different social biases that we can consider. For example, bandwagon effect (also known as herding) refers to the tendency of people to do (or believe) things because many other people do (or believe) the same things [SB15]. In our model, the agent can believe in the climate urgency and want to choose an environment-friendly mode of transport. Hence, it can use the bus (2) or biking (3) to go to work. Another type of bias is framing, where people may make different decisions based on the same information, depending on how that information is presented [Dru01]. For example, our agent can choose the car (1) since it is presented as the most trendy and comfortable option.

- **Norm Consideration**: Observing (subconsciously or consciously) the behaviours of others can have an impact on a person's behaviour [CRK90; CKR91]. For example, our agent can start by using the car (1) to go to work. However, it finds that other agents are using the bus or the train on the way. After a few more trips, its preferences for the bus (2) and the train (4-5) are updated to be higher than using the car option.

- **Learning**: The agent can start by experimenting with all options over a certain period (e.g. one week) before deciding on the best alternative. This decision strategy is suitable for an unknown environment for the agent [Mit13].

- **Habit**: It is a behaviour we often do, almost without thinking [Gra08, p. 359]. Without a significant change in our agent context, it can repeat the previous action without deliberation. In a sense, it provides a heuristic for decision-making.

The agent decision-making process needs to identify and distinguish their effects on the agent's behaviours. At the same time, the agent can combine these different aspects for a more complex decision-making process. For example, it has a habit of cycling to work (3), but because many neighbours start to take the bus, it starts to follow them and take the bus (2). In another scenario, although an agent receives a discount to use the bus (2), it still chooses the car (1) due to habit. Hence, our model needs to be able to represent these different decision-making determinants.

We can now consider some options to simulate these aspects. A simple design involves agents follow some sets of behaviour rules (i.e. decision-tree or production-rule systems), which apply both in information-gathering stage and making a final choice. It is typically used in conjunction with a set of assumed preferences for the agent to rank outcomes by desirability order. Examples include heuristics that update agent's behaviours according to the accumulated experience (e.g. [TBS05]) or pick the next option that satisfies the qualities identified from empirical data analysis (e.g. [HB99]). In this setup, modellers have a straightforward job to trackback any changes in agents' behaviour but have to face a significant increase in computational complexity when a new rule is introduced [RN10, pp. 46-48]. Alternatively, ABM projects in the mobility domain often use *random parameters logit* [Anw+14] to assign predicted probabilities to outcomes of a set of alternative options. Examples include [CBA08; Anw+14]. These models incorporate empirical data (such as observed choices, survey responses to hypothetical scenarios or administrative records) becomes a flexible framework to estimate the parameters of choice behaviour. However, they cannot distinguish the mentioned aspects of decision-making and their impact on behaviour, so we do not consider them suitable options for our purpose. As the economic interpretation of rationality often influences researchers within the field of mobility mode choice, there is a limited commitment to create a general agent-based approach that can highlight multiple dimensions in decision-making (see surveys such as [Sha+13; Sch+20a]).

We can look for alternatives in the field of Multiagent System (MAS), which uses agents for the primary purpose of solving engineering problems [DKJ18]. MAS decision-making architectures are often implemented based on the principle of cognitive theories, such as the theory of human practical reasoning [Bra87] or unified theories of cognition [New94]. Therefore, they can be sophisticated and have cognitive realism. By familiarising themselves with these agent architectures, ABM modellers can refine their views of the agent as a complex although computable entity. It could further enable more recognition of the role of the mind as a necessary intermediate between social structures and social behaviours [Min88]. However, the benefits above are currently limited in practice due to the difference in methodologies and concepts between MAS and social research. MAS architectures are often derived from computer science concepts and have limited expressive power in other domain's ontologies [Dig17]. In addition, the number of architectures from MAS is significantly large[1]. Therefore without a sufficient technological background, it is difficult for a modeller to select and implement one architecture accurately to cover the multitude of factors affecting decision-making.

In terms of our mobility example, MAS agent architectures can be used to reproduce a more elaborate decision-making process by assigning agents with beliefs, values or world views that correspond to observation from ethnographic data or stakeholder's assessment. Examples include Belief-Desires-Intentions (BDI) architecture and its derivatives (e.g. BOID [Bro+02], eBDI [Per+05], BRIDGE [DDJ08]). Each of these models focuses on an aspect of decision-making, e.g. emotions (eBDI), norms (BRIDGE). Therefore, they have limited capability to capture other aspects. Similarly, normative models, such as NoA and EMIL-A, do not consider the affective or learning dimensions. Another class is cognitive architectures, including CLARION [Sun06], ACT-R [TLA06], SOAR [Lai12] and PECS [Urb00]. However, they do not have the ability to take into account personality differences, emotion and social variables[2].

The variety of the aforementioned aspects and strategies makes it challenging to use only one of the mentioned architectures to represent them correctly. Consequently, creating causal links between individual aspects and the agent's behaviour can be a difficult task. Moreover, as they are developed based on specific cognitive and computer science theories, there is a conceptual gap that limits their application in social science fields.

One promising direction to minimise this gap is to build a framework from a theory of human decision-making that covers a broad set of decision-making aspects, or so-called behaviour determinants. As highlighted in [Sch+17], users can expect manifold benefits from developing models based on this framework:

[1] A survey of all these architectures is provided in Chapter 3

[2] More details of these architectures can be found in Chapter 3.

- It facilitates the reuse and comparison of models since a theory could serve as a standard reference. This process can save time, accelerate scientific advancements, and foster the development of new theories [Bel+15; CCB08].
- Integrating such theories into the decision process can help advance the modelling of behavioural aspects and processes that may yield insights into the possible social phenomena to different events and policy measures or changes in the environment.
- At the same time, modelling an agent's decision-making based on a theory can help limit the enormous options of what aspect could be included in the model to only those deemed relevant by the theory [Edm17b].
- This approach may provide a standard protocol in interdisciplinary teams and facilitate communication between modellers and social scientists [DEB07].

To create such a framework, we perform a literature review. As the results, a candidate for the theoretical foundation of our framework is the Theory of Interpersonal Behaviours (TIB) [Tri77]. TIB states that behaviour is primarily based on the intention to engage in the act, habit and facilitating conditions. In addition, it includes a meaningful set of determinants covering all decision-making aspects mentioned. TIB also provides a function to derive the utility value (or preference of an agent) for an action. We formalise the decision-making framework following the process in [Sch+20b]. More information about the search and formalisation process can be seen in Chapter 4.

Next, by implementing this framework in different decision-making contexts, this thesis can provide an assessment of its functionality and potential application in future study (see chapters 5 to 8).

1.2 Thesis objectives

As mentioned above, we aim to create and formulate an agent framework based on a theory that allows modellers to think more systematically about how to simulate human decision-making in their models and use the diversity of concepts from the field of social sciences. We acknowledge that there is a significant amount of available behaviour determinants that can be considered in one model[3]. As a first step, we decide to include the five high-level dimensions that are often applied in an agent's decision-making, according to [BG14]:

- **Cognition**: The agents have some form of deliberation to choose between different alternatives.
- **Affection**: The agents include an explicit representation of emotion and account for its impact on agent's behaviour.

[3]Interested readers might refer [Die10] and [Bau03] for concise psychology surveys of human decision-making theories.

- **Social aspects**: The agents are capable of distinguishing social network relations and status.
- **Norm**: The agents include an explicit reasoning about social norms, institutions and organizational structures.
- **Learning**: The agents can adapt their behaviour to the change in environment.

These dimensions cover a wide range of topics mentioned in ABM study (see surveys such as [Gil04; MLT09; HB11b]). They will also be used to compare the related work in Chapter 3 and to ensure our framework can cover a sufficient amount of aspects that are relevant for social studies.

To allow better reusability of the agent design for different contexts, the components of our framework are expected to be extensible or can be replaceable by aspects from other theories. It also enables more than five of the dimensions of decision-making above to be considered.

Since research in ABM is often rooted in realism, data patterns arising from agent's behaviours will need to be able to compare and contrast, in selective and sometimes subtle ways, with a range of corresponding empirical results. This capability will partly allow the assessment of a specified level of external validity, but it will also support the comparative evaluation of diverse decision-making aspects and their impacts on the behaviour of the agents.

We list below a summary of the criteria required for our framework:

- The framework shall have a sophisticated decision-making mechanism, moving away from ad-hoc, oversimplified behavioural rules.
- The framework shall allow expression of assumptions, postulates and concepts explicitly drawn from social sciences. At the minimum, it should include the following five dimensions: cognition, affective, social factors, norm and learning.
- The framework shall have a extensible mechanism, which allows to reflect a variety of decision-making aspects.
- The framework shall offer a mechanism to incorporate empirical data.
- The framework shall be applicable in different decision-making contexts and domains, e.g. mobility mode choice, health care, and public policy.

1.3 Research questions

As mentioned, we use TIB as the theoretic foundation to create an agent behavioural decision-making framework. The framework is then applied to create agent-based models in different domains and contexts. These models can provide evidence for the efficiency of our proposed framework to help to close the conceptual gap mentioned in Section 1.1.

In addition, we can assess the practicability of the developed models by identifying their purposes for real-world application. This part will be based on the definitions provided by Edmond in his paper [Edm17a]:

- *Prediction*: The model has the ability to reliably anticipate well-defined aspects of data that are not currently known to a useful degree of accuracy via computations using the model.
- *Explanation*: The model aids the understanding of why something occurs, such as complex social phenomena.
- *Description*: The model is used to record, in a coherent way, a set of selected aspects of the phenomena under observation.
- *Theoretical exposition*: The model allows one to explore the consequences of theoretical assumptions and properties using mathematics and computer simulation.
- *Illustration*: The model makes an idea of a complex system clear or shows it is possible by demonstrating it in a concrete example that might be more readily comprehended.
- *Analogy*: The model of a process is used as a way of thinking about something in an informal manner.
- *Social learning*: The model can capture a shared understanding (or set of understandings) of a group of people.

Hence, the following questions will be addressed in this thesis:

- **Question 1**: What does an implementation of an agent decision-making framework with TIB offers? Does it help to close the conceptual gap between MAS and ABM?
- **Question 2**: When developing models for different case studies, what are the limitations of our agent decision-making framework with TIB?
- **Question 3**: For which research purposes are the models based on our agent decision making framework especially useful?

1.4 Structure of the thesis

To create a socio-psychology framework for agent decision-making, we will first look into the necessary components of an agent-based model, specifically in the field of single decision-maker. Next, a literature review will provide the current state-of-the-art and why a new agent decision-making framework is needed. We will then formulate the framework and apply it to different practical projects to test its functionality in various fields and contexts. Further experiments can then be performed to provide insight into the interchangeability of our determinants set. The rest of this thesis is structured as follows:

- Chapter 2 provides the foundations of this study. In particular, it first introduces the ABM methodology and its components, including agents, the environment and their interactions. Section 2.1 in this chapter discusses the definition, types and general architectures of an agent. In addition, the mathematical and theoretical background of the agent's decision-making is provided. The next two sections summarise the properties of an environment and the type of interactions that should be considered in ABM.

- Chapter 3 further discusses the related work in agent decision-making architectures and socio-psychology frameworks. The categories include BDI and its derivatives, normative architectures, cognitive models and inspired psychological frameworks. Combining with the foundations in the previous chapter, Chapter 3 explain why these state-of-the-art approaches are unsuitable for our aforementioned research purposes and the current research gap for a new framework.

- After the fundamentals of this research have been laid out, the specification of the new architecture is presented in Chapter 4. In the first Section 4.1, we survey and provide the reasons for which TIB is chosen among some recently developed behaviour theories. The proposed decision-making framework is then detailed in Section 4.2. It also gives an overview of an agent's core components, utility function and running examples. Next, the description of the classes and their pseudo-code are provided, which is translated into Java code in Appendix B. The steps to implement the framework are listed in the next section. Finally, we conclude the chapter with some words about our framework documentation.

- The next four chapters 5 to 8 show the implementation of our framework to develop agent-based models in four case studies in the domain of mobility, trust and reputation, vehicle purchasing and public health. Chapter 5 focuses on simulating modal choices of daily transportation, representing short-term decision-making. Trust and reputation of train services are the focus of the following study (Chapter 6). The agent's decision component is updated to consider uncertainty in the environment. The third case study centres around purchasing vehicles and aims to extend the agent's perception component to account for bounded rationalities (Chapter 7). Lastly, we add a case study about the migrants' behaviours in centres that enable the spread of COVID-19, where a different theory is utilised to resolve the problem of missing socio-psychological data (Chapter 8). In each case study, we provide the context, state-of-the-art, data mapping, decision-making mechanism, calibration and an experiment for the framework's functionality. These sections are presented following the framework implementation steps provided in Chapter 4.

- Chapter 9 summarises the previous experience in the process of designing and applying our framework to answer the research questions above. First, it looks at the advantages of building an agent framework using the TIB's and whether it helps to reduce the gap between MAS and ABM (Question 1). The following section assesses the issues of using the framework for model development (Question 2). Finally, we generalise the modelling purposes applicable to our generated models and experiments (Question 3).

- Finally, a conclusion is drawn in Chapter 10. We then outline future research and development directions for the framework.

Chapter 2

Foundations and principles

In this chapter, we present the foundations and principles for this thesis. We first introduce the background of Agent-Based Model (ABM). One of the most recent and prominent definitions was given by Nigel Gilbert:

> "Formally, an Agent-based Modelling is a computational method that enables a researcher to create, analyse, and experiment with models composed of agents that interact within an environment." [Gil19, p. 2]

This definition centres around the functionality of ABM and the contexts where it is beneficial. Focusing more on the components of ABM, Wilensky and Rand gave the following definitions:

> "The core idea of Agent-based Modelling is that many (if not most) phenomena in the world can be effectively modelled with agents, an environment, and a description of agent-agent and agent-environment interactions." [WR15, p. 32]

They highlight the three main components that are necessary for an ABM: 1) agents, 2) an environment, and 3) the interactions. The following three sections discuss them in details. In addition, we summarise them as the technical requirements that apply to our framework.

2.1 The agent

There are many definitions that have been given to *agent* in the literature (e.g. [Gre+97, p. 2], [Woo09, p. 21], [BZW12, p. 19], [Gil19, p. 5]). We are also aware that there is currently no universal agreement in the literature on a precise definition. In this thesis, we follow the one proposed by Russell and Norvig, which is well-know within the fields of MAS and ABM:

> "An *agent* is anything that can be viewed as perceiving its *environment* through *sensors* and acting upon that environment through *actuators*." [RN10, p. 34]

The components of this definition can be seen in Figure 2.1. In simple words, an agent can be described as a function that receives percepts from its environment and produces action(s) as output. The component with a question mark denotes an internal computing process in which the agent decides on the most appropriate action(s). Its design ranges from simple mapping rules to complex processes, which will be further discussed in Section 2.1.1.

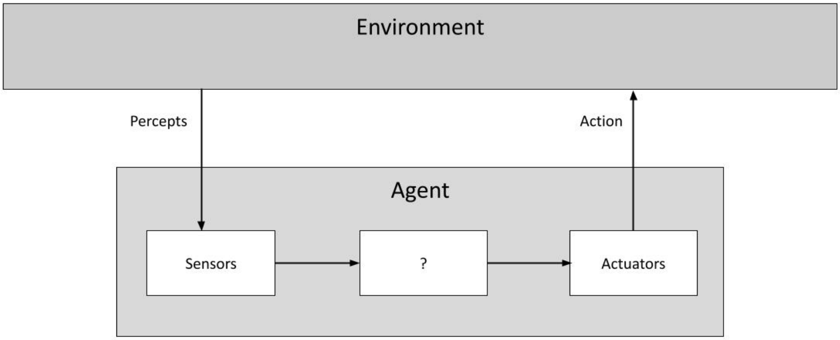


FIGURE 2.1: An agent with its component, taken from [RN10, p. 35]

In the following subsections, we first discuss the mathematical and theoretical background necessary to formalise and reason about the agent decision-making process. The agent's decision-making process can be categorised into different agent types based on the implementation principles and required components. We will introduce four of them in the first subsection, including simple reflex, model-based reflex, goal-based and utility-based agents. In practice, these agent types can be built upon several architectures, each of which is a blueprint depicting the arrangement of components and flow of information. The description of four general classes of agent architectures is provided in the final subsection, including logic-based, reactive, BDI and hybrid architectures.

2.1.1 Mathematical and theoretical background

This section provides the foundations of decision-making process in an agent using a combination of utility theory and probability theory. In the rest of this thesis, we follow the notions and concepts mentioned by Russell and Norvig in [RN10, pp. 610-636] to deal with choosing among actions based on the desirability of their immediate outcomes. The authors proposed a definition of *rational agent*:

> "For each possible percept sequence, a rational agent should select an action that is expected to maximise its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has." [RN10, chap. 2, p. 37]

In a real-world situation, we often have to make decisions in an environment that is stochastic and partially observable, in which the same action performed twice may produce different results or may even fail completely. Russell and Norvig defined a random variable, $RESULT(a)$, whose values are the possible outcome states. The probability of outcome s', can then be written as:

$$P(RESULT(a) = s'|e) \tag{2.1}$$

where e is evidence observations of events that action a is executed.

A *utility function*, $U(s)$, can be use to capture the agent preferences. It assigns a single number to express the desirability of a state and can take a cardinal form (specific numeric value) or an ordinal form (ranking choice by order). Given the evidence ($EU(Do(a)|e)$), the *expected utility* of an action is defined as follows:

$$EU(a|e) = \sum_{s'} P(RESULT(a) = s'|a, e)U(s') \tag{2.2}$$

A rational agent can follow the principle of *Maximum Expected Utility (MEU)*, i.e. it simply chooses the action that maximizes the agents expected utility [RN10, pp. 611]:

$$action = arg \max_a EU(a|e) \tag{2.3}$$

The MEU principle formalises the general notion that the agent should "do the right thing". However, it does not go to the full operationalisation of that advice [RN10, p. 611]. It is because computing $P(RESULT(a)|a, e)$ requires a complete causal (i.e. possible outcomes) model of the environment and probabilistic inferences of the agent's beliefs, which can have NP-hardness (i.e. there are no polynomial-time algorithms to solve this problem) [RN10, pp. 510-551]. In addition, $U(s')$ requires search or planning because an agent needs to know the possible future states in order to assess the worth of the current state.

The MEU principle is not the only rational way to make decision. The agent can also maximises the weighted sum of utilities or minimises the worst possible loss. However, it provides a clear relation to the idea of performance measures for a possible action [RN10, p. 611]. The following notations can be used to describe agent's preferences:

- $A \succ B$: The agent prefers A over B.
- $A \sim B$: The agent does not prefer one over another.
- $A \succeq B$: The agent prefers A over B or considers no difference between them.

In an uncertain environment, we can think of the set of outcomes for each action as a lottery ticket. Russell and Norvig define a lottery L with multiple outcomes $S_1, ..., S_n$ that occur with probabilities $p_1, ..., p_n$ as follows [RN10, pp. 612]:

$$L = [p_1, S_1; p_2, S_2; ...p_n, S_n] \tag{2.4}$$

We list six constraints that are required in any reasonable preference to obey (also known as known as the *axioms of utility theory*) according to Neumann [NVJ44]:

- **Orderability**: Exactly one of $(A \succ B)$, $(B \succ A)$, or $(A \sim B)$ holds.
- **Transitivity**: Given any three lotteries, if an agent prefers A to B and prefers B to C, then the agent must prefer A to C.

$$(A \succ B) \bigwedge (B \succ C) \Rightarrow (A \succeq C) \tag{2.5}$$

- **Continuity**: If some state B is between A and C in preference, then there is some probability p for which the agent is indifferent between getting B for sure and the lottery that yields A with probability p and C with probability 1 - p.

$$A \succ B \succ C \bigwedge \exists p[p, A; 1-p, C] \sim B \tag{2.6}$$

- **Substitutability**: Simpler lotteries can be replaced by more complicated ones, without changing the indifference factor.

$$A \sim B \Rightarrow [p, A; 1-p, C] \sim [p, B; 1-p, C] \tag{2.7}$$

- **Monotonicity**: If an agent prefers the outcome A, then it must also prefer the lottery that has a higher probability for A.

$$A \succ B \Rightarrow (p > q \Leftrightarrow [p, A; 1-p, B] \succ [q, A; 1-q, B]) \tag{2.8}$$

- **Decomposability**: Compound lotteries can be reduced to simpler ones using the laws of probability. An agent should not automatically prefer lotteries with more choice points.

$$[p, A; 1-p, [q, B; 1-q, C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C] \tag{2.9}$$

From these constraints, we can derive the following consequences (for the complete proofs, see the work of Von Neumann and Morgenstern in [NVJ44]):

- **Existence of Utility Function**: If an agents preferences obey the axioms of utility, then there exists a function U such that $U(A) > U(B)$ if and only if A is preferred to B,and $U(A) = U(B)$ if and only if the agent is indifferent between A and B.

$$\begin{aligned} U(A) > U(B) \Leftrightarrow A \succ B \\ U(A) = U(B) \Leftrightarrow A \sim B \end{aligned} \tag{2.10}$$

- **Expected Utility of a Lottery**: The utility of a lottery is the sum of the probability of each outcome multiplied by the utility of that outcome.

$$U([p_1, S_1; ...; p_n, S_n]) = \sum_i p_i U(S_i) \tag{2.11}$$

In general, each outcome S_i of a lottery can be either an isolated state or another lottery. Once the probabilities and utilities of the possible outcome states are specified, the utility of a compound lottery involving those states is entirely determined. Because the outcome of a non-deterministic action is a lottery, an agent can act rationally, i.e. consistently with its preferences, only by choosing an action that maximises expected utility according to Equation 2.3 [RN10, p. 614].

The preceding theorems establish that a utility function exists for any rational agent but do not establish its uniqueness. An agent's behaviour would not change if utility function $U(S)$ were transformed according to an affine transformation:

$$U'(S) = aU(S) + b \tag{2.12}$$

where a and b are constants and $a > 0$.

The existence of a utility function that describes an agent's preference behaviour does not necessarily mean that the agent is explicitly maximising that utility function in its deliberations. Rational behaviour can be generated in several ways. However, by observing a rational agent's preferences, an observer can construct the utility function representing what the agent is trying to achieve. Combining with Von Neumann and Morgenstern's theorem and constraints [NVJ44], it can be used for measuring the strength of a rational agent's preferences over sure options. We refer the interested reader to Gilboa [Gil09] for a comprehensive discussion of the theoretical, philosophical and mathematical properties of decision-making under uncertainty.

The notion of utility can be modified to capture several concepts in different research fields. For example, classical economists typically assume that an agent's behaviour is motivated primarily by material incentives and that decisions are governed mainly by self-interest and rationality [SS15; McF01]. In this context, decision-makers use all available information logically and systematically to make the best choices given the alternatives and the objectives to reach [KS02]. In a typical ABM design, agents make decisions to maximise certain profit, revenue, or rate of profit while not violating any constraints. However, in many other designs, more abstract utility functions (e.g., the CobbDouglas utility function [CW04]), which sometimes includes ecological indicators (e.g., [NK09]), or consumption, aspiration (e.g., [GPL03; Sim55]), are used instead of monetary values. These functions often take an additive or exponential form of a weighted linear combination of many criteria under consideration (e.g., [Le+08; Chu+09; Zel+08; Bro+04; BT06]). With such utility definitions, it is possible to calculate the probability of an agent's choosing one option (e.g., one site or one opportunity) as the probability that the utility of that option is more than or equal to that of any other option. Whichever method is used, the agents are often assumed to make rational choices. Another alternative is the usage of Prospect Theory, proposed by Daniel Kahneman and Amos Tversky [TK79]. Their theory assumes that people derive utilities from *gain and loss* which are measured relative to some *reference points*, rather than from the resulting outcome of the decision. In this case, the utility function follows the loss-aversion bias; namely, the "pain" of losing α dollars should outweigh the "pleasure" of gaining α dollars.

Decision-making can be high stakes and involve multiple aspects, such as in public policy decisions. All similar problems, in which outcomes are characterised by two or more attributes, are handled by *multi-attribute utility theory*. The attributes can be written as $X = X_1, ..., X_n$. A complete vector of assignments can be presented as $x = x_1, ..., x_n$, where each x_i is either a numeric value or a discrete value on an assumed order. When all other things are equal, higher values of an attribute correspond to higher utilities. The multi-attribute utility theory's main goal is to calculate a utility function $U(x_1, ..., x_n)$, which represents the person's preferences on lotteries of bundles. In other words, lottery A is preferred over lottery B if and only if the expectation of the function U is higher under A than under B:

$$E_A[U(x_1, ..., x_n)] > E_B[U(x_1, ..., x_n)] \tag{2.13}$$

As one of our research objectives is to cover multiple dimensions of decision-making (see Section 1.2), we need to represent the $U(x_1, ..., x_n)$ function precisely. For this purpose, the strongest independence property, i.e. *additive independence*, is considered: two attributes X_1 and X_2 are preferentially independent of a third attribute X_3 if the preference between outcomes x_1, x_2, x_3 and x'_1, x'_2, x_3 does not depend on the particular value x_3 for attribute X_3. This property holds true for our decision-making dimensions. According to Keeney et al. [KRM93, p.295], the n-attribute utility function can take an additive form:

$$U(x_1, ...x_n) = \sum_{i=1}^{n} k_i * U_i(x_i) \tag{2.14}$$

where U and U_i are normalized to the range $[0, 1]$, and the k_i is a normalization constant. This function implies that assessing an additive value function of n attributes can be done by assessing n separate one-dimensional value functions, reducing an exponential in the number of preference experiments. It is a natural way to describe agents' preferences and is valid in many real-world situations [RN10, p. 525]. Even when the additive independence does not strictly hold, the Equation 2.14 still provides a good approximation to the agents preferences [RN10, p. 525]. When uncertainty is present in the simulation domain, the structure of preferences between lotteries and the resulting properties of utility functions also need to be considered. Interested readers can find a more detailed discussion in [KR76].

In this thesis, we utilise the notion of the utility function and its additive form for multiple attributes as foundations to compare an agent's options in both certain and uncertain environments. In terms of selecting a behaviour theory suitable for our research purposes, it is essential that this theory can provide a way to capture these notions effectively.

2.1.2 Agent types

In this section, we describe different agent types underlying how an agent takes an input from the environment and derives an action output. This section follows the work of Russell and Norvig [RN10, pp. 46-54], starting from rather simple reflex agent, then discuss more complex ones.

Simple reflex agent

This agent type works on predefined *condition-action* rules, which means it performs its tasks by simply mapping the current state to corresponding action. The structure of this agent type can be seen in Figure 2.2, showing how the *condition - action* rules allow the agent to make the connection from percept to action.

FIGURE 2.2: The typical model of a simple reflex agent, taken from [RN10, p. 49]

This type of agent has a simple reasoning process. Considering that our research framework needs to move away from the ad-hoc, oversimplified representation of decision-making, this design is not a good fit. In addition, it only works correctly if the decision can be made in a fully observable environment [RN10, p. 49]. As one of our research objectives is to work in multiple decision-making contexts, which can include a partly observable environment, this agent type can limit the usage of our framework.

Model-based reflex agents

Model-based reflex agents can work in a partially observable environment by providing an *internal state* to keep track of the part of the world that is either visible or invisible to the agent. Hence, it requires two kinds of knowledge for its perception sequence: 1) information on how the environment evolves separately from the agent, 2) information on how the agent's actions influence the environment [RN10, p. 50]. A set of *condition-action* rules can then be used to choose the appropriate action(s). These processes can be seen in Figure 2.3.

As the model-based reflex agent still mainly based its decision-making on the *condition-action* rules, this agent type does not fit our objective of having a system closer to human deliberation.

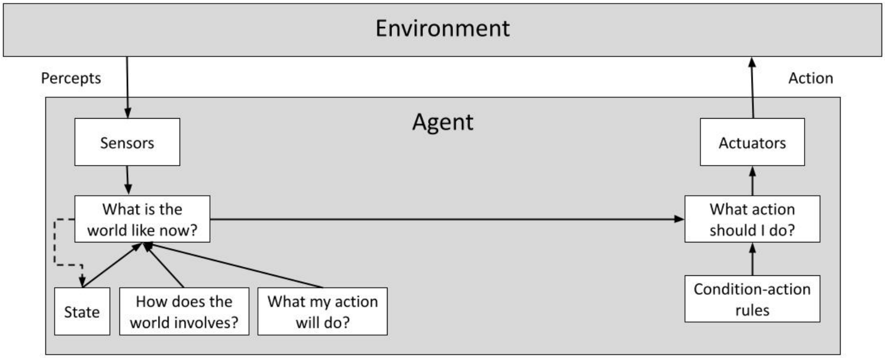


FIGURE 2.3: The typical model of a model-based reflex agent, taken from [RN10, p. 51]

Goal-based agents

Goal-based agent uses a set of *goals* to specify desirable situations. Its main objective is performing actions to reduce its distance from the goal. Hence, it may have to consider a vast list of a possible sequence of actions, which requires search and planning[4]. Figure 2.4 shows the typical structure of such agent.

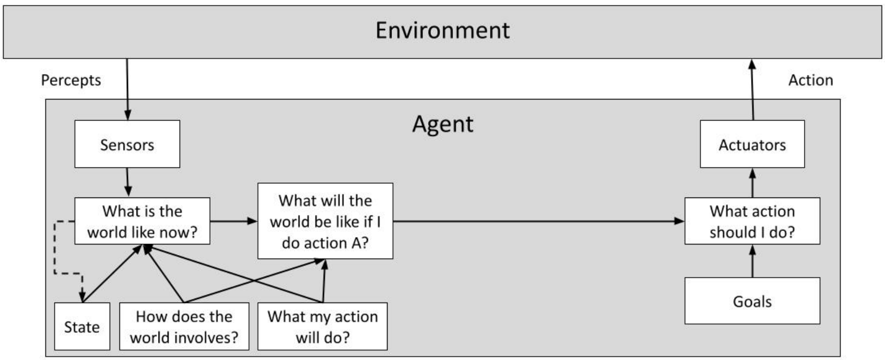


FIGURE 2.4: The typical model of a simple goal-based agent, taken from [RN10, p. 52]

Compared to our thesis objective, a goal-based agent has a binary distinction between good (goal) and bad (non-goal) states, while our target has a continuous measure of outcome quality. In addition, the main objective of this agent type is

[4]Interested readers can find a full discussion on AI perspective from the work of Kolp et al. [KGM01].

solving satisfaction problems and not expressing how the agent deliberates between options. Therefore, this agent type is not suitable for our framework.

Utility-based agent

In several contexts, having a goal is not enough because the agent may have several actions that all satisfy this goal. Compared to goal-based agent, *utility-based agent* is a more complex form that applies a *utility function* mapping a state into a real number. This utility function can be derived from the mathematical and theoretical foundations mentioned in Section 2.1.1 to measure the agent performance. Even if the goal is not satisfied by all possible actions, this function also provides a way for the agent to choose the most appropriate action (e.g. by applying the MEU principle). Figure 2.5 describes a typical design for a utility agent.

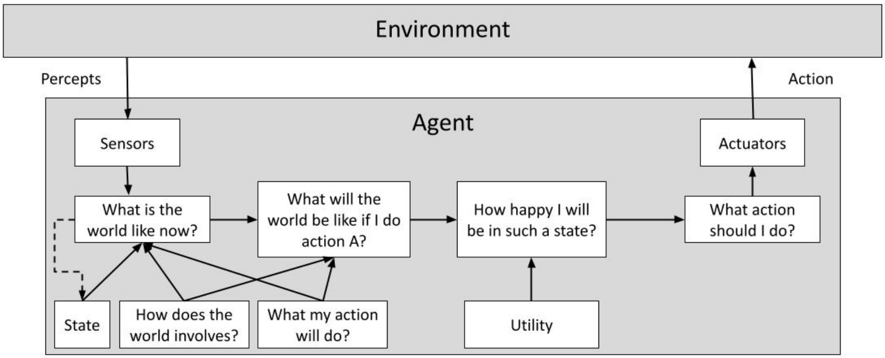


FIGURE 2.5: The typical model of a utility agent, taken from [RN10, p. 54]

This agent type is useful when there are multiple possible alternatives, and an agent must determine the optimum action. Compared to reflex agents, it not only allows a more elaborate decision process but also applicable environments with different properties. In addition, the utility function provides a way to incorporate and measure the impacts of different decision-making attributes on the agent's action (e.g. economic attitudes, social factors, emotion, habits). Considering our research objectives in Section 1.2, this type of agent is the suitable classification.

2.1.3 General architectures

The main components of the agent's implementation and how they interact are defined in an architecture. The subsection below follows the work of Wooldridge in [Woo99, pp. 42-66], which introduces four general architectures:

- **Logic-base agent architecture**: The agent makes decision based on logical deduction.

- **Reactive agent architecture**: The agent makes decisions based on the direct mapping from situation to action.
- **BDI agent architecture**: The agent is characterised by the implementation of beliefs, desires and intentions concepts in the agent's reasoning process.
- **Layered (hybrid) architecture**: The agent combines reactive and deliberative components to form a hierarchy of interacting layers, which have different levels of abstraction.

As we identified our targeted agent type as the utility-based agent, the reactive architecture is simply unsuitable for its implementation. Hence, its introduction is omitted in this section but can be found in [Woo99, pp. 48-54].

Logic-based architecture

The logic-based architecture is based on the traditional artificial symbolic approach by representing the agent behaviour and the environment with symbolic representations [Woo99, p. 42]. Hence, the agent action is based on pattern matching and the syntactical manipulation of these symbolic representations, which correspond to logical deduction or theorem proving[5]. The agent is implemented with a set of deduction rules and a database - a set of logic formulae representing the current belief about the environment. In action selection, the agent will try to derive a predicate $Do(a)$ or falsify $\neg Do(action)$, where a represents an action. If it is successful, the action a will be performed.

Because of its pure logical structure and resulting precise semantics, this model is interesting for theoretical investigations. However, there are some disadvantages arising from the fact that their inference can become computationally complex due to the automated theorem proving process. Decision-making in this architecture is made on the assumption that the world will not change significantly while the agent is deciding what to do and that an action which is rational when decision-making begins will be rational when it concludes. Thus, if agents have to keep a time constraint, logic-based architecture can become a problem, significantly when the environment can change during the inference process, limiting its application in real-world contexts.

BDI architecture

The BDI architecture provides a robust standard framework for any agent-based simulation that wants to consider a more realistic human decision-making process. It is a software architecture that implements the principal aspects of Michael Bratman's theory of human practical reasoning [Bra87] and has been formalised for the usage of Distributed Artificial Intelligence (DAI) by Rao and Georgeff [RG+95]. It allows an agent to have an internal representation of the world to perform their planning

[5] An introduction and examples of deductive agents can be found in [Woo09, pp. 49-55]

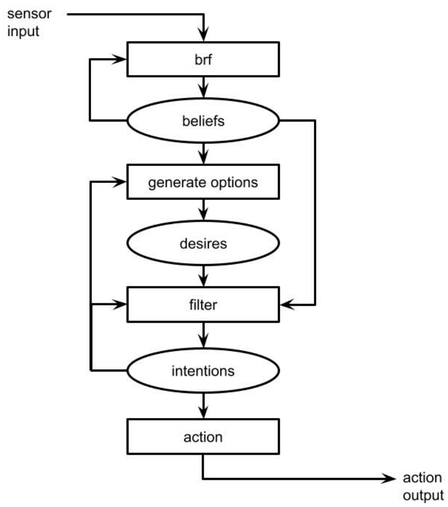


FIGURE 2.6: BDI agent architecture, taken from [Woo99, p. 59]

processes. The full reasoning cycle of a BDI agent architecture can be found in Figure 2.6. It includes three core components:

- **Belief** is the informational state of the agent, in other words, its beliefs about the world. What an agent believes may not necessarily be true (and may change with time).
- **Desire** is the objectives or situations that the agent would like to accomplish or bring about.
- **Intention** is the deliberative state of the agent - what the agent has chosen to do. Intentions are desires to which the agent has, to some extent, committed.

In addition, there are four functions:

- **brf - Belief revision function** takes a perceptual input and the agent's current beliefs and determines a new set of beliefs.
- **Generate options** allows the agent to calculate the options available to satisfy its desires based on its current beliefs about its environment and its current intentions.
- **Filter** represents the agent's deliberation process, which determines its intentions based on its current beliefs, desires, and intentions.

- **Action selection** determines an action to perform based on the updated intentions.

The BDI model is attractive for two main reasons. First, it is intuitive since its reasoning cycle resembles the kind of practical reasoning that we appear to use in our everyday life. Second, the basic components (i.e. beliefs, desires, and intentions) and their functions indicate what subsystems might be required to build an agent. The main difficulty is knowing how to efficiently implement these functions to strike a balance between committed and overcommitted to one's intentions during the deliberation process. A detailed description of this architecture can be found in [Woo99, pp. 54-61]. Several extensions have been derived from this paradigm to cover different dimensions of human behaviour, whose details can be found in Section 3.1.

Layered architecture

The main idea of this architecture is to create an agent that can reflect both reactive and deliberate behaviours. They are represented by separate subsystems, which are organised into layers of a hierarchical structure. We can divide this architecture into two categories based on the flow of information between the layers, i.e. horizontal and vertical layered architectures.

In horizontal layering, each layer is directly connected to the perceptual input and action output (see Figure 2.7). In a sense, this structure represents different agents generating suggestions for a central control. An example of this architecture is the TouringMachine [Fer92], which consists of three layers: a reactive layer, a planning layer, and a modelling layer. Each has its own internal process and can operate concurrently and independently from others.

According to Wooldridge [Woo99, pp. 61-62], a major advantage of horizontal layered architectures is their conceptual simplicity. In an agent's design, n different layers can represent *n* different types of behaviour. The problem is that the layers can compete with one another to generate action suggestions. Hence, there is a danger that the final behaviour output of the agent will not be coherent. One solution is the introduction of a *mediator* function, which makes decisions about which layer has control of the agent at any given time. However, it can generate a *bottleneck* if the number of layers and actions are significantly large.

These problems are partly alleviated in vertically layered architectures, in which perceptual input and action output are processed by at most one layer. They can be subdivided into one-pass and two-pass control architectures, which are illustrated in Figure 2.8. In one-pass architectures, the control passes from the first layer (which gets perceptual input) to the final layer (which produces action output). The main difference in two-pass architectures is that information flow back down after being processed by the final layer. In this case, the first layer is responsible for both receiving the perceptual input and generating action output. Prominent examples of a system based on vertical layered architectures are the InterRRap architecture [FMP96, p. 403-405] and 3T architecture [PB+97].

In these architectures, the complexity of interactions between layers is reduced at the cost of some flexibility. It is because the control must pass between each different

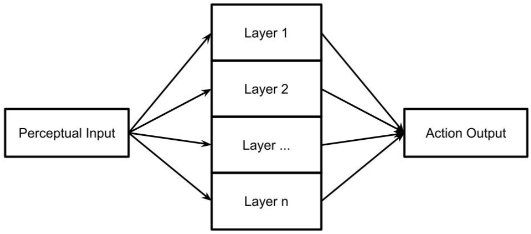
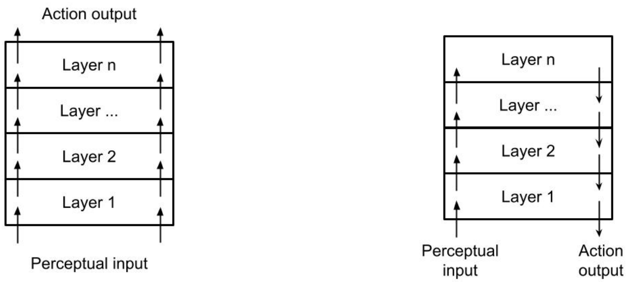


FIGURE 2.7: Information and control flow of horizontal layered architecture, taken from [MPT95, p. 263]



(A) Vertical one-pass control architecture

(B) Vertical two-pass control architecture

FIGURE 2.8: Information and control flow of vertical architectures, taken from [MPT95, p. 263]

layer to make a decision. This design is not fault-tolerant: failures in any layer are likely to affect agent performance seriously.

Considering our research objectives, these layered architectures specify how information flows between different components of the agent, which is useful for us to consider and organise multiple aspects of decision-making.

2.2 The environment

The environment defines the conditions in which the agents exist in the system [Ode+02]. In addition, it provides the space in which agents interact and operate [CH12]. The environment could also refer to the infrastructure in which agents are deployed and, thus, be studied as a first-order abstraction from an agent-oriented software engineering perspective, as thoroughly discussed in [Wey+04].

Regarding the characteristics and dynamics of the environment, Russell and Norvig propose to consider several properties [RN10, p. 42-46]:

- **Single-agent vs multiagent**: An environment with only one agent is a single-agent environment; otherwise, it is a multiagent environment.

- **Episodic vs sequential**: In episodic environments, the agent percepts and performs a single action in each episode. In addition, there is no dependency between current and previous episodes. In sequential environments, the previous decisions can influence all future decisions.

- **Discrete vs continuous** A discrete environment has a finite number of actions that can be deliberated to obtain the output. On the contrary, the number of actions remains unknown in a continuous environment.

- **Static vs dynamic**: An environment is dynamic if it can change while an agent deliberates. Otherwise, an environment is static.

- **Fully observable vs partially observable**: An environment is effectively fully observable if the agent's sensor can detect all aspects relevant to the choice of action. On the contrary, it can be partially observable because of noisy and inaccurate sensors or because parts of the state are missing from the sensor data.

- **Known vs unknown**: The outcomes (or their probabilities) for all actions are given in a known environment. Otherwise, the environment is unknown. A known environment can still be partially observable.

- **Deterministic vs stochastic**: An environment is deterministic if the next state of the environment is entirely determined by the current state and the action executed by the agent(s). Otherwise, it is stochastic.

As the nature of ABM, we are working in *multiagent environments*, where multiple agents are required to give insights into their social patterns. Other environment properties can be decided based on the modelling context. Therefore, more discussions are provided in chapters 5 to 8 to select the most appropriate ones for each case study.

2.3 The interaction

From the ABM perspective, Wilensky and Rand categorise the agents and environments interactions into five different types [WR15, pp. 257-262]. This categorisation is one of the complete overviews of this topic:

- **Agent-Self interactions**: The agent decides what to do depending only on its current internal state. An example in demography is an agent giving birth or dying after a certain amount of time passes.

- **Environment-Self interactions**: Several elements of the environment can be updated themselves without the influence of agents' actions. For example, in a car purchasing model, new types of cars become available in the market in a specific year.

- **Agent-Agent interactions**: An agent can influence other agents using their direct actions. For example, in a driving car simulation, the agent can decide to slow down or speed up depending on the car in front. Agents can also communicate with one another using direct (e.g. message passing [Hew77], signalling [Dan+07]) and indirect methods (e.g. black-board based communication [Cra88])[6].

- **Environment-Environment interactions**: Different environmental elements can interact with one another. A typical example is the diffusion of policies, a region can adopt a successful campaign from its neighbourhood.

- **Agent-Environment interactions**: This type of interaction is in line with the description provide in Section 2.1. The agent examines/percepts and manipulates the environment through its actions. The environment observes the agents and can influence their actions in some ways.

Depending on the modelling context and type of decision-making, one or many of these interactions can be considered. According to Wilensky and Rand, the two types that are most important in ABM are *Agent-Agent* and *Agent-Environment interactions*[WR15, pp. 257-262]. We will consider them in more detail in our framework design (see Chapter 4).

2.4 Summarisation of technical criteria

We summarise below the technical criteria for a suitable framework that can satisfy the research objectives.

- **It includes the basic components of an agent**: The agents need a way to perceive the current state of the environment (e.g. sensors), a way to derive the agent's internal state (e.g. memory), a way to communicate the action to the environment (e.g. actuator) and a decision-making process.

- **It is based on an abstract architecture that can organise the flow of information between the decision-making determinants**: Potential candidates include logic-based, BDI and layered (hybrid) architectures.

- **It provides a way to apply utility functions to compare different alternatives**: The agent's decision-making provides a way to derive the utility of outcomes when an action is performed. The usage of utility function has a broad application in different areas, such as AI, economic, social and psychological studies.

[6] Interested readers can find a more detailed discussion in [CS11]

- **It works in different types of environment**: It might involve modifying the agent's perception and communication/actuators to adapt to the environmental change. In addition, the utility function has to consider the probability of different outcomes from the agent's actions.
- In terms of interaction, the framework should include a mechanism that allows a user to implement agent-agent and agent-environment interactions.

Chapter 3

Related work

This thesis's main objective is to enhance ABM models with more complex decision-making mechanisms. One promising direction is to look at MAS architectures and frameworks. Experts in this field have provided various methods to model people's decision-making based on their experience, theoretical assumptions and knowledge. While this thesis cannot include all existing research and projects relevant to the topic, we aim to cover a wide range of approaches and their examples. The classification below is partly based on the work of Tina Balke and Nigel Gilbert in [BG14]. In Chapter 2, we identified *utility agent* type as the application of our framework. Consequently, in this chapter, we exclude the reactive architecture design and its implementation - production rule systems. As the thesis focuses on building an agent framework based on human behavioural research, another category of socio-psychological inspired frameworks will be discussed, including Modelling Human Behaviour (MoHuB) [Sch+17] and Consumat [JJV99].

In each section, we describe the agent's main components and the decision-making cycle of architecture or framework. In addition, a comparison with our research objectives is provided. We are particularly interested in whether it can express various concepts explicitly drawn from social sciences, especially the five-level dimensions: *cognitive*, *affective*, *social*, *norm* and *learning*. Finally, we summarise why they do not completely fit our research purposes and, therefore, the need to develop a new agent framework based on the existing work but with some features not yet covered.

3.1 BDI and its derivatives

We introduced the Belief-Desires-Intentions (BDI) reasoning cycle in Section 2.1.3. Here, we provide a more complete description of typical BDI agent design.

The BDI model [BIP88; RG91] is a popularly used framework, incorporating beliefs, desires and intentions, to design intelligent autonomous agents. It aims to meet real-time constraints by reducing the time used in planning and reasoning. A BDI agent is designed to be goal-directed, reactive, and social [AG16]. It means a BDI agent is able to react to changes and communicate in its embedded environment as it attempts to achieve its goals. Mechanisms to response to new situations or goals during plan formation for general problem solving and reasoning in real-time processes are also included in BDI systems [GI89; SDSP06]. BDI agents are typically implemented as Procedural Reasoning System (PRS), which is a framework for constructing real-time reasoning systems that can perform complex tasks in dynamic environments [GHFSS04; SDSP06].

The full BDI agent architecture can be found in Figure 3.1. As mentioned in Section 2.1.3, the BDI model is established based on three mental attitudes, including *beliefs*, *desires* (or goals) and *intentions*. To accompany them, there is usually a library of plans, which define procedural knowledge about low-level actions (or steps) that are instructions on how to achieve a goal in specific situations. In its reasoning cycle, a BDI agent updates its beliefs based on its perception. The targeted intentions are pushed onto a stack, which contains all the intentions to be achieved. Using its library, the agent uses the first intention of its intention stack to look for any plans with matching post-condition(s). All options that have their pre-conditions satisfied according to the agent's beliefs are considered. The plan of highest relevance to the agent's beliefs and intentions is selected. Based on its goals and the plan information, the agent computes new intentions, updates old ones and then translates them into executable actions.

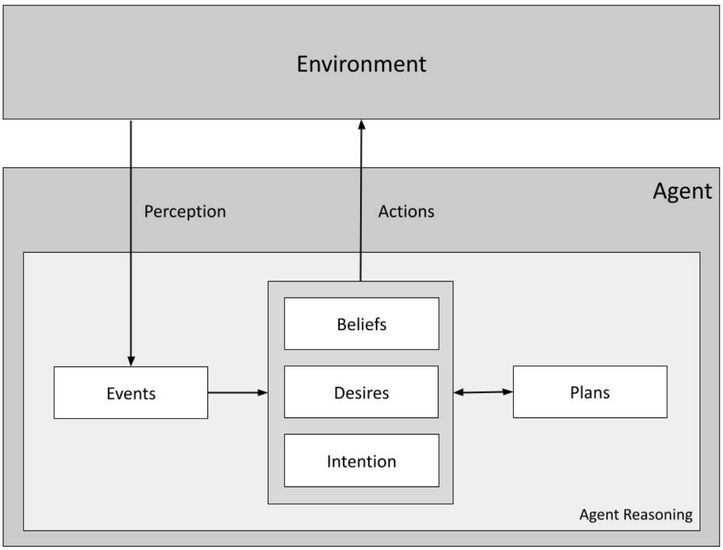


FIGURE 3.1: The BDI agent architecture, adapted from [BG14]

Pure BDI agents still lack any specific learning mechanisms (from past behaviours) and adaptation to new situations [GHFSS04; PWP05]. It is an important feature for agents situated in dynamic environments, which changes can make methods for achieving goals that previously worked well to become ineffective. ABM researchers also question whether the three determinants are sufficient to represent human thought process [RG+95; Her+17]. Several extended architectures address these issues by adding some other elements into the architecture, such as obligation (BOID [Bro+02]), emotion (eBDI [Per+05; JV06]) and social norms (BRIDGE [DDJ08]).

3.1.1 BOID architecture

The Beliefs-Desires-Obligations-Intentions (BOID) architecture is an extension of the BDI, which focuses on normative concepts, in particular *obligations* (i.e. the agent is morally or legally bounded to perform an action) [Bro+02]. It assumes that agents are aware of all social obligations, although the BOID architecture allows them to deliberate about whether or not to follow these obligations and contribute to collective interests. Agents can also drop some obligations in favour of others to deal with conflict among norms.

Its target is not to reach goals, satisfy desires or fulfil obligations, but to decide which desires and obligations it will follow given its beliefs and intention. In other words, it wants to resolve conflicts among its attitudes. According to the work of Broersen et al. in [Bro+02], the order of the four main components of the agent (i.e. *Belief* - B, *Obligation* - O, *Intention* - I, *Desire* - D) can be used to indicate the order of overruling in the case of conflict between them. Therefore, agents can be classified into the following types:

- Realistic: *Belief* overrides all other components.
- Simple-minded: BIDO and BIOD - *Intention* overrules *Desire* and *Obligation*.
- Selfish: BDIO and BDOI - *Desire* overrules *Obligation*.
- Social: BIOD, BOID and BODI - *Obligation* overrules *Desire*.

These classes specify the components of an agent, how they are related, and how the information flows around. Propositional formulas are used to represent the content of informational and motivational attitudes. Each component in the full BOID agent architecture (see Figure 3.2) has set of formulas - or so-called *extension* - as input and output. The platform also includes a *planning process (P)* to decide which actions should be performed to achieve the agent's intentions. The decision-making cycle in this architecture is similar to the BDI one, except for the intention/goal generation process. It starts with receiving input from the environment. The agent then calculates a set of candidate goal sets based on the priority given by its type (e.g. BIDO versus BDIO), elects one goal set, decides which plans should be performed. Next, all components are updated and the cycle starts again.

Most works on BOID are concentrated on the formalisation of the idea, and less on practical implementations [Bro+01; Bro+02; BDT05]. Other issues with this architecture include computational complexity [BDT05] and finding a simulation platform that can cope with BOID complex compositional systems [DT04]. Considering our dimensions, it has the same properties as BDI but allows the modelling of social norms in terms of obligations [BG14].

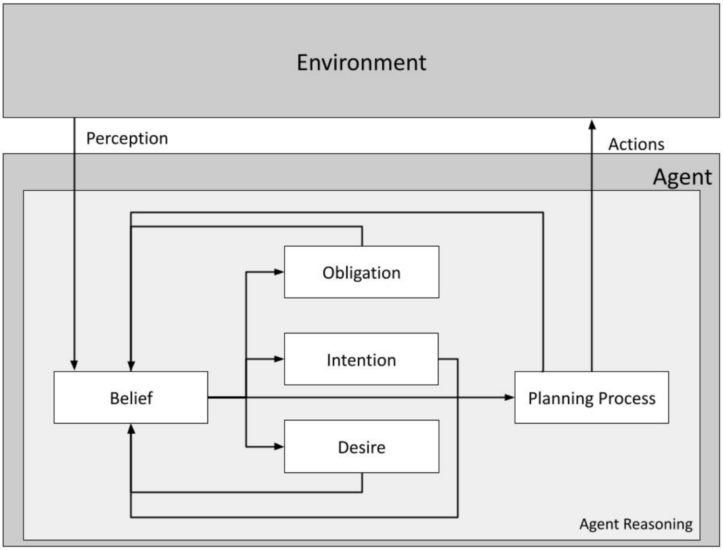


FIGURE 3.2: The general BOID architecture, adapted from [Bro+02]

3.1.2 eBDI architecture

Emotional BDI (eBDI) is one extension of the BDI concept that incorporates emotions as one decision criterion into the agent's decision-making process [Per+05; POM07; JV06]. It is one of the first architectures that account for emotions to control how agents act upon their environment.

Figure 3.3 depicts the main components of an eBDI agent. It utilises the idea of *Capabilities* and *Resources* as the basis for the representation of emotions inside the agent [Per+05]. Capabilities are abstract plans that the agent can use to act. The agent can use its resources (either physical or virtual) to make these plans more specific by matching them against the agent's owned capability and opportunities from the environment. With the limited information about itself and the environment, the eBDI agent might not have the knowledge of all its resources and capabilities. Hence, the agent needs to become "effective" by using an *Effective Capabilities and Effective Resources* revision function (EC-ER-rf) to consider its perceptions and its *Belief*, *Desires* and *Intention* components. The *Sensing and Perception Module* uses semantic association rules to filter and give suitable semantic meaning to the information from all perceptions and other sensor stimuli. Another component is the *Emotional State Manager*, which is responsible to control the capabilities and resources used in the information processing phases. On each level, the agent adds an emotional input to the BDI process that selects the action plan the agent executes. Full description of this architecture is given in [Per+05; JV06].

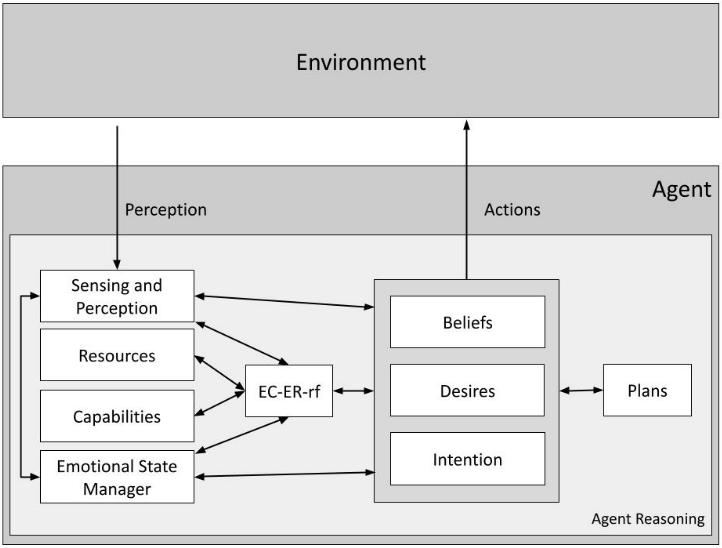


FIGURE 3.3: The eBDI architecture, adapted from [Per+05]

In recent years, Sanchez et al. improved the original model by incorporating a well-known psychotherapeutic model, the ABC model[7], with other affective theories, to support the modelling of the agents' behavioural and affective responses [San+19]. Overall, the eBDI architecture improves on BDI architecture with the recognition of affective dimension but does not consider learning, norms or social relations [BG14].

3.1.3 BRIDGE architecture

Developed by Dignum et al., the BRIDGE architecture extends the idea of the social norms in BOID and aims to provide a model for agent reasoning, which can describe the influence of policies or comparable external influences on the behaviour of agents [DDJ08]. For this purpose, the authors argue that it is essential to provide agents with constructs for social awareness and reasoning update process [DDJ08]. BRIDGE architecture also integrates the bottom levels of Maslow's hierarchy of needs [Mas43] into the agent's decision.

The architecture introduces three new mental components: 1) *Ego*, 2) *Response* and 3) *Goal*. The first describes the filters and ordering preferences that the agent uses. It also defines the agent's personality type, which determines the choice of reasoning mode (e.g. explorative, goal-based, belief-based or evidence-based). *Response* relates to the bottom layers of the hierarchy of needs and implements the reactive behaviour of the agent. It directly influences current goals and can overrule any plans. The *Goals*

[7] Interested readers can find more details in [Ell62].

component derives agent's goals from the agent's current desires derived from its preferences and current state (from *Ego* and *Response*) and the deficiency need.

In the reasoning cycle of a BRIDGE agent (see Figure 3.4), all components and the interactions between them have to work concurrently to allow continuous processing of the input in the form of sensory information (conscious input) and other "stimuli" (subconscious inputs) [DDJ08]. These inputs are first interpreted by adding extra weight or priority to beliefs depending on the agent's personality characteristics (i.e. *Ego*)). Next, these beliefs are sorted and used to filter the agent's desires at any moment. Candidate goals can then be selected based on the selected desires and personality characteristics. The next step is calculating possible plans with the influence of the *Ego* component. Finally, one of the plans is chosen for execution, and the agent's beliefs can be updated. It should be noted that basic urges, the current emotional state, and stress levels in the *Response* component can control the order and choice of current goals.

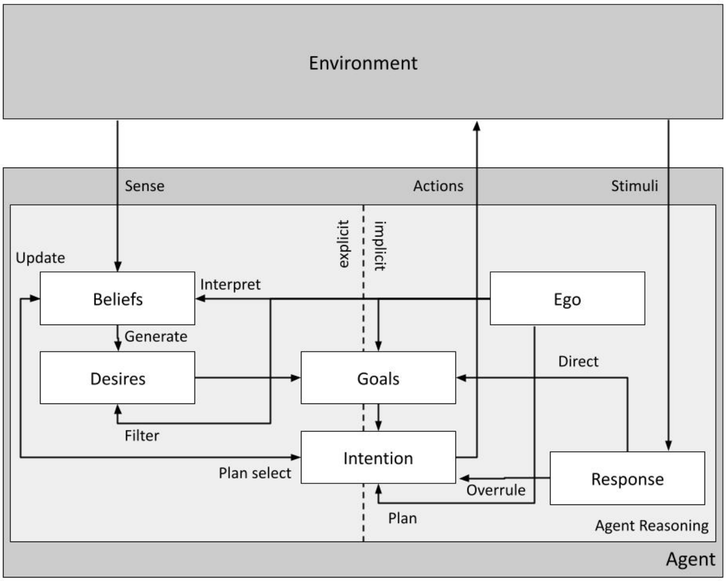


FIGURE 3.4: The BRIDGE architecture, adapted from [DDJ08]

Regarding our dimensions, BRIDGE can cover the emotional aspects by using the *Ego* component to specify types of agents and their different emotional responses to various stimuli. According to the authors, the components they introduced on top of BDI (e.g. *Ego*, *Response*) influence social norms and their internalisation by agents [DDJ08]. Similar to BOID, BRIDGE considers norms solely from an obligation perspective.

3.2 Normative models

Norms can determine behavioural patterns and are used to guide and control the performance of actions within a specific context [Kol05]. A norm can be expressed as an obligation, a prohibition or permission over an individual's behaviour in a decision-making context. It can establish what, when and to whom this regulation is subjected. Therefore, it provides a way to obtain a desirable system behaviour [DHM10].

Using norms as instruments to influence an agent's behaviour has been popular in the ABM community as they are external to the agent and can only be established within the society/environment in which the agent is situated. Many implementations of normative agent architectures thus take BDI and procedural reasoning as a starting point and introduce norms as an influencing factor in this deliberation process. Example includes models of normative systems (e.g. [BTV06; Dig04; GAD06; LL03]) and models of norm-autonomous agents (e.g. [KN03; BTV06; Cas+99; Cri13; And+07b]). The majority of these architectures are abstract framework and have a limited number of implementations.

The BOID architecture can also be classified in this category due to its implementation of social obligation. However, since the structure of its components are similar to the BDI architecture (see Figure 3.2), we listed it in Section 3.1.

Considering our objective of finding the agent architecture that considers not only norms but also the agent's deliberation, we focus our attention the three architectures below:

3.2.1 Deliberate Normative Agents architecture

The Deliberate Normative Agents is a cognitive research-inspired abstract model and is based on an idea that social norms need to be involved in the decision-making process of an agent [CC99; Cas+99; CCD99].

The architecture has six components (see Figure 3.5), which can be further grouped into three layers: 1) an *Interaction Management layer*, 2) an *Maintenance layer*, and 3) a *Process Control layer* where the processing of the information and reasoning occurs. The first layer handles the interaction of an agent with other agents and the environment. Next, the *Interaction Maintenance layer* records information about the other agents, the world and society as a whole. The final layer is where the process of information and reasoning occurs.

At its core, the agent reasoning cycle is similar to the one in BDI (see Section 2.1.3). The consideration of norms in the architecture adds an additional level of complexity since the internalised social norms can affect the generation and the selection of intentions. This process starts with recognising a social norm through observation and communication. It is then evaluated and stored in the *Interaction Maintenance layer*. Using the *Process Control layer*, the agent determines which norms to adopt (or ignore) as well as how they influence the agent's actions. This process results in norm-specific intentions, which are considered in the core reasoning cycle.

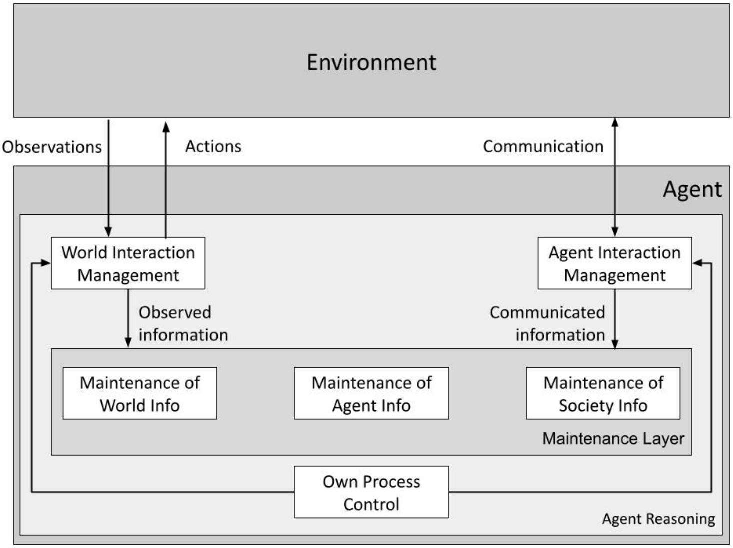


FIGURE 3.5: The Deliberative Normative architecture, according to [Cas+99]

Compared to BDI agents, deliberate normative agents are enhanced on the social and the learning dimensions. It includes an explicit separate norm internalisation reasoning cycle. Concerning the learning dimension, agents have limited capabilities to learn new norm-specific intentions [Cas+99]. To the best of our knowledge, there is no mention of an affective component in this architecture.

3.2.2 NoA architecture

The Normative Agent (NoA) architecture focuses on incorporating norms into agent decision-making while also using extended notions of norms [Kol05]. It is implemented with an explicit representation of a *normative state* - a collection of norms (obligations, permissions and prohibitions) that an agent has at a point in time. Obligations motivate a normative agent to act in order to achieve a goal state. Prohibitions restrict an agent's behaviour, whereas permissions allow an agent to pursue certain activities. NoA agents are equipped with the ability to construct plans to achieve their goals while not violating any of the internalised norms.

A typical decision-making cycle of NoA agent can be seen in Figure 3.6. According to the author in [Kol05], it starts with considering the two sources to change its beliefs: 1) external perception of the environment and 2) internal manipulations resulting from the execution of previous plans by the agent. The agent also looks at external norms to obtain normative specifications from the environment. Next, the reasoning

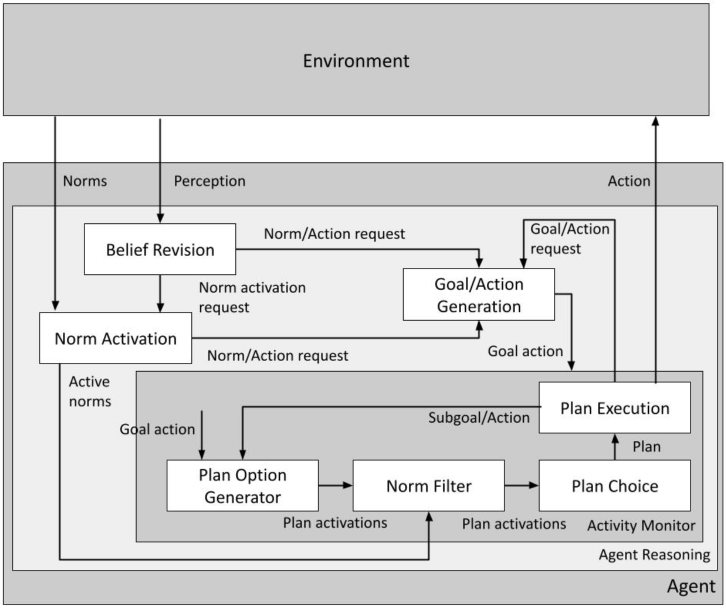


FIGURE 3.6: The NoA Architecture, adapted from [Kol05][p. 80]

cycle performs two distinct operations: 1) the activation and initiation sets of plans, norms and 2) the deliberation process, including the plan selection and execution. The activation mechanism activates and deactivates norms and plans in agreement with the changes of the agent's beliefs. The action generator then selects a plan to be executed, which can include a number of sub-activities to fulfil the main plan and any sub-plans.

Evaluating NoA according to our dimensions, we can see that a deliberation architecture is used. Emotions or other affective elements are not mentioned explicitly. Instead, it focuses on norms with many broad definitions (e.g. obligation, permissions and prohibitions).

3.2.3 EMIL-A architecture

The EMIL-A architecture [And+07a] was developed to model the process of learning, internalisation and usage of social norms in the agents' decision-making.

According to the authors in [And+07a], the EMIL-A model contains the following components:

- four different procedures: 1) norm recognition, containing the mentioned normative frame, 2) norm adoption, containing goal generation rules, 3) decision making, and 4) normative action planning.

- three different mental objects: 1) normative beliefs, 2) normative goals, 3) normative intentions.
- an inventory which includes: 1) a memory to store a repertoire of normative action plans and 2) an attitude module to capture the internalised attitudes and morals of an agent.

EMIL-A include several modules that allow agents to deal with the complexity of social norms (see Figure 3.7). As the name suggested, the *Event Classifier* module perceives, classifies and transfers the events into other specific modules. The *Norm Recognition* module allows agents to interpret new social norms from the interactions with other agents and generate the corresponding beliefs. The *Norm Adoption* module determines which social norms and goals are adopted inside the agent's memory. If agents decide to comply with these norms and goals, the *Norm Compliance* module converts them into normative intentions. The agents use the *Norm Enforcement* module to select the most appropriate sanctions to be implemented when others violate a norm. Finally, the *Norm Salience* uses a utility function to calculate the expected utility obtained from either violating or fulfiling a norm, which becomes a criterion for agents to accept or reject norms.

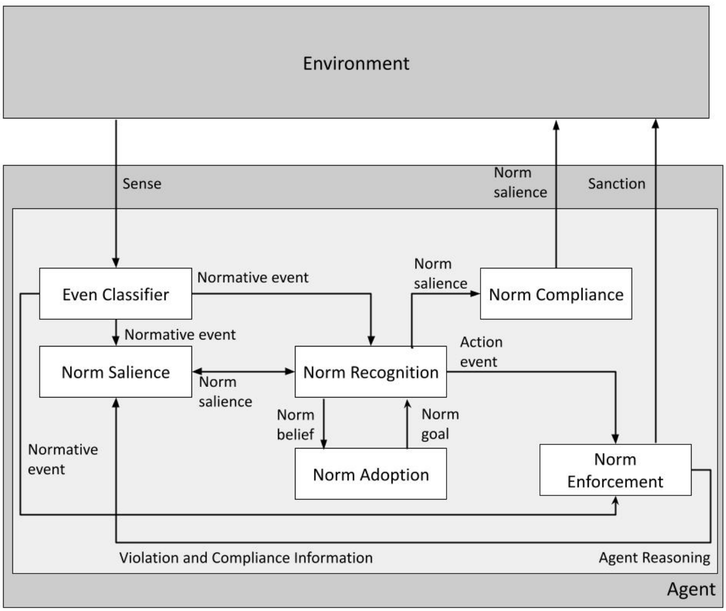


FIGURE 3.7: The EMIL-A architecture, adapted from [And+07b]

Regarding our dimensions, this agent architecture has the same properties as the NoA architecture. It utilises logic-based architecture. In addition, EMIL-A focuses on the social aspects of norms and considers them as a learning method for related change of intentions. However, it lacks a representation of the affective dimensions.

3.3 Cognitive models

A cognitive architecture can be defined as a computational model based on the studies of how the human mind works [New73; Kla+87]. Its main goal is to incorporate the various results of cognitive psychology in agent's reasoning to explain a wide range of human behaviour and to mimic the capabilities of human intelligence [And+04; CTN07]. As a result, a cognitive architecture should possess the following characteristics, according to [CTN07]:

- Long and short-term memories.
- Structures to represent memories and their organisation.
- Functional processes that operate on these structures.

In this section, we list six of the most well-known cognitive architectures that consider the structural properties of the human brain from the field of cognitive sciences: SOAR [LNR87], ICARUS [Lan+91], ACT-R [BA98], CLARION [SMP98] and PECS[Urb00]. They usually consider social theories and focus on different issues that were ignored in the rational agent. It should be noted that BDI architecture also belongs to this category. We separate them due to the conceptual different of their components and a significant number of architecture based on the BDI structure.

3.3.1 SOAR architecture

SOAR [LNR87] is one of the earliest and most extensively developed AI architectures in the history. The main idea is that unifying different or overlapped theories without conflict can produce intelligent behaviours with appropriate learning mechanisms. Figure 3.8 illustrates the SOAR agent architecture. It comprises a *Working Memory* (short-term), a *Long-term Memory*, a *Reasoning module*, a *Perception module*, *Action module* and *Learning module*.

According to Chong et al. [CTN07], knowledge is stored in the *Long-term Memory*, which can be divided into procedural, semantic and episodic memories. Procedural memory provides the knowledge of how to perform a task. Semantic memory contains general facts about the environment. Episodic memory contains specific memory of an event that the agent experienced. Hence, both procedural and semantic memories can be applied universally, whereas episodic memory is contextually specific. When procedural knowledge is insufficient, semantic and episodic memories can be employed to aid the problem-solving process.

The *Working Memory* in SOAR architecture stores all the knowledge that is relevant to the current situation. It contains the goals, perceptions, hierarchy of states, and operators. The goal directs the agent's intention into the desired state. The perception contains the immediate model of the current environment. The states (and sub-states)

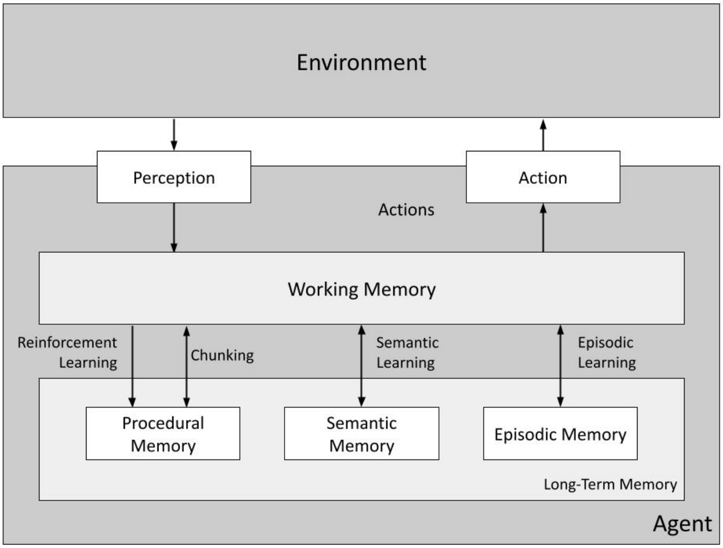


FIGURE 3.8: The SOAR architecture, adapted from [Lai12]

give information on the current situation. The operator provides the procedure for problem-solving. The *Working Memory* also has access to the relevant knowledge from the long-term memory and motor actions.

The decision-making process in SOAR mainly consists of matching and firing rules that are context-dependent representations of knowledge [LRN86]. SOAR architecture allows the firing of rules and retrieving several pieces of knowledge occur simultaneously. It is based on an idea that it is better to rely on as much information as possible in an uncertain situation with limited knowledge. To bring about the knowledge relevant to the current problem, production rules are then matched with the working memory subcomponents. At the same time, references are created as recommendations for selection of appropriate operators. This phase ends when the long term memory finishes the firing of rules, ensuring that all knowledge relevant are considered before a decision is made [LRN86; LLR06]. The decision cycle then continues with the decision phase, wherein the preferences are evaluated. The operators that better satisfy the agent's goals are chosen and applied during the application phase. During the output phase, the selected operator is applied as the agent's action. If there is insufficient information to select and apply an operator, an impasse arises, and a sub-state is created to resolve it. In SOAR, there are four types of impasse: no-change, tie, conflict, and rejection [LRN86; CTN07].

The classical SOAR uses a learning mechanism called *chunking*, which acquires rules from goal-based experience [LRN86]. A more recent version (SOAR 9) extends to reinforcement learning modules. They are also linked with the appraisal component, which is used to capture the affective aspect of decision-making. The agent evaluates the situation along with multiple dimensions such as goal relevance, causality, etc. The appraisal of how the goal is met turns into the agent's emotions. This architecture, however, lacks norms consideration and other social aspects [BG14].

3.3.2 ICARUS architecture

ICARUS shares several core assumptions with other candidate architectures, including SOAR and many other reactive architectures [Lan+91]. According to a survey made by Langley and Choi [LC06], these include claims that:

- Short-term memories are distinguished from long-term memories.
- Memories contain modular components with symbolic structures.
- Long-term structures can be accessed with pattern matching.
- Cognitive processing occurs in retrieval/selection/action cycles and involves a dynamic composition of mental structure.

However, ICARUS also makes different assumptions [LC06], including:

- High-level cognition is grounded in perception and action.
- Relational categories and skills are separate cognitive entities.
- Short-term components are instances of long-term structures.
- Long-term knowledge is organised hierarchically.

Figure 3.9 presents an overview of the ICARUS architecture, which consists of four main components: the *Perceptual Buffer*, the *Conceptual Memory*, the *Skill Memory*, and the *Motor Buffer*. The *Perceptual Buffer* is a temporary storage of percepts from the environment. The *Conceptual Memory* can be divided into *Short-term Conceptual Memory* and *Long-term Conceptual Memory*. The *Short-term Conceptual Memory* stores the set of active inferences about the perceived world. The *Long-term Conceptual Memory* comprises the known conceptual structures describing objects or classes observed from the environment. The *Skill Memory* can also be split into *Long-term Skill Memory*, which stores the possible actions/skills set, and *Short-term Skill Memory*, which contains the chosen skill to be implemented. The skill's signals in the *Motor Buffer* are then executed by the ICARUS agent, which can create changes in its environment.

Considering our dimensions, ICARUS has an underlying symbolic-based reasoning cycle. To the best of our knowledge, its components do not cover the affective, social and norm aspects.

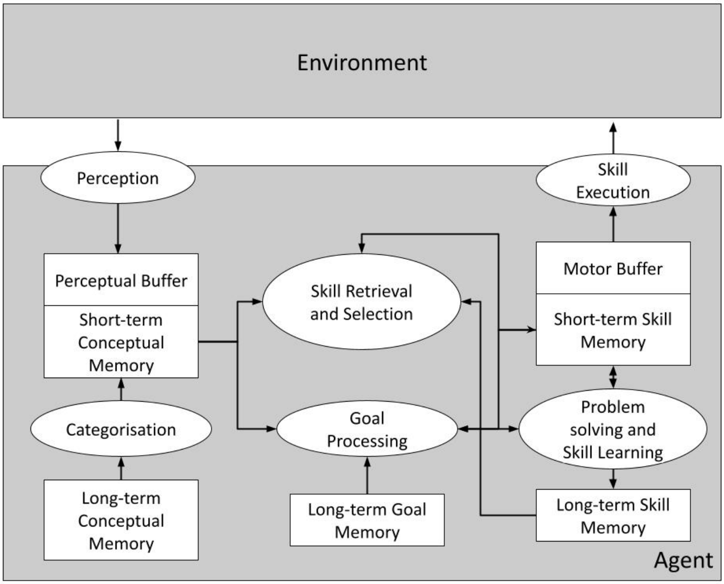


FIGURE 3.9: The ICARUS architecture, adapted from [CL18]

3.3.3 ACT-R/PM architecture

Taking the ideas from Newell and Simon [SN71], Adaptive Control of Thought-Rational (ACT-R) and its extensions ACT-R/PM combine models of cognitive psychology with perceptual-motor modules to create a production system in which all components communicate with each other [BA98; AL98; Byr00].

According to the authors in [And+04], the architecture consists of a *Cognitive Layer*, a *Perceptual/Motor layer* and a *Buffer Immediate Layer*. In the *Cognitive Layer*, agent memory is represented by *procedural modules* (for production rules) and *declarative modules* (for facts and goals) (see Figure 3.10). The former takes a central position connecting all major components and is represented by production rules. The latter is represented by a number of schema-like structures called *chunks* [SS91], which contains pointers to their category and content. Agents also have a working memory in the buffer layer, which responds to the knowledge used when performing a task. This knowledge can be retrieved from both declarative and procedural memories. The architecture also includes several learning mechanisms: 1) declarative knowledge can be modified either from the input of the perceptual layer or as a result of a production rule; and 2) procedural knowledge is altered through inductive inference from existing procedural rules and case studies.

Similar to other cognitive-inspired architectures, ACT-R and ACT-R/PM focus on

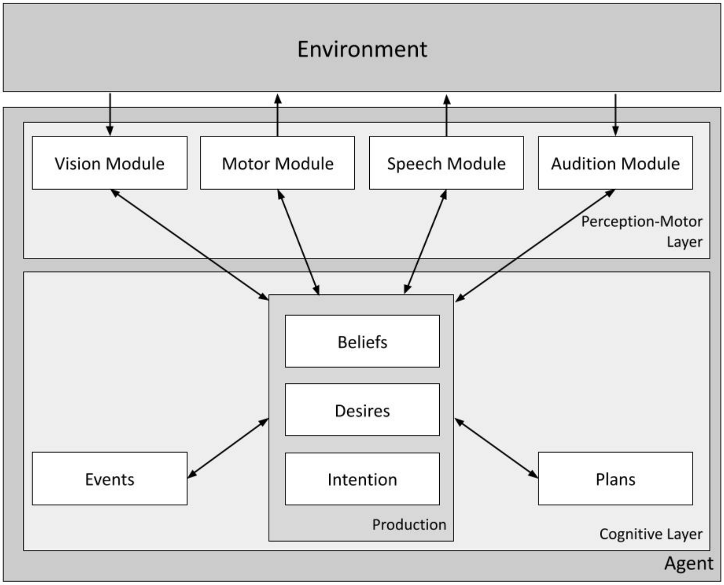


FIGURE 3.10: The ACT-R architecture, reproduced from [TLA06][p. 31]

modelling single human decision-making and learning processes. In its documentation [BA98; AL98; Byr00], there is little mention of the contribution of affective (i.e. emotion) and social aspects in the agent's decision.

3.3.4 CLARION architecture

Connectionist Learning with Adaptive Rule Induction ON-line (CLARION) architecture uses hybrid neural networks[WS98] to simulate cognitive and social psychology tasks in AI applications [SMP98; SPS02]. According to the author, it has several features which distinguish it from other cognitive architectures: 1) containing built-in motivational and meta-cognitive structures, 2) integrated both bottom-up and top-down learning and 3) considering two dichotomies: explicit versus implicit representation and action-centred versus non-action-centred representation [Sun06]. The CLARION architecture consists of two levels: 1) a top level containing prepositional rules and, 2) a bottom level (also know as reactive level) containing procedural knowledge (see Figure 3.11). Procedural knowledge can be acquired from reinforcement learning (i.e. Q-Learning [WD92]) accumulatively over time. Besides, declarative knowledge can be learned from trials and errors. The architecture combines their recommendations in a weighted sum to select the most appropriate action to perform at each step [CTN07].

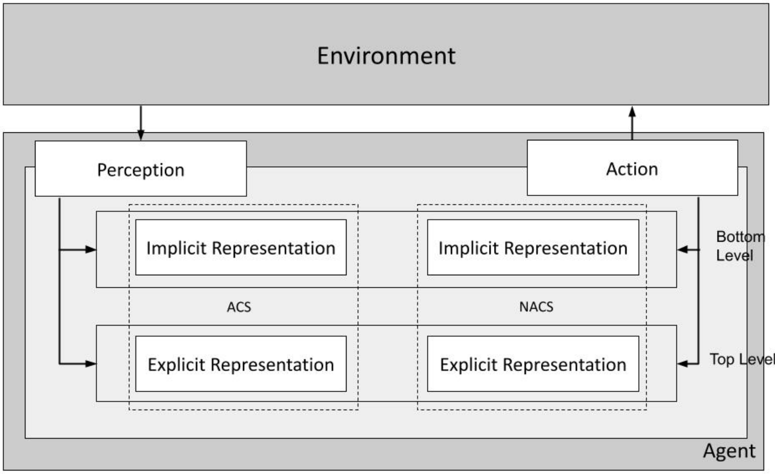


FIGURE 3.11: The CLARION architecture, adapted from [Sun06].

The CLARION architecture includes *Non-Action-Centered Subsystem* (NACS) and *Action-Centered Subsystem* (ACS). NACS contains mostly declarative knowledge, while ACS contains mainly procedural knowledge. In addition, it includes two other subsystems: 1) the *Motivational Subsystem* (MS) uses goals to instruct agent's action and 2) the *Meta-Cognitive Subsystem* (MCS) is utilised as the main controller the operations of all subsystems dynamically.

CLARION includes both similarity-based and rule-based reasoning to mimic human reasoning. It is done by comparing a known chunk (single declarative unit of knowledge) with another chunk. The relations between two chunks can be established when they have certain degree of similarity in their representation. To select the most appropriate course of action for the agent to react to its environment, the reasoning of this architecture can occur iteratively to ensure all possible conclusions are found. During this process, each step's conclusion can be used as a starting point for the next step.

Learning in CLARION can be differentiated between implicit learning (procedural knowledge) or explicit learning (declarative knowledge). Procedural knowledge at the bottom level can be updated through the reinforcement learning paradigm, such as multi-layer neural networks and backpropagation algorithms, which are used to compute Q-values [WD92]. When the agent favours the outcome of an action, the Q-value of the action raises and thus increases the tendency to perform that action. In addition, learning rules at the top level can occur by extracting knowledge from the bottom level. Upon successfully executing an action, the agent extracts a rule corresponding to the action selected by the bottom level and adds it to the top level. It then tries to verify the rules learnt via applying them in the subsequent actions.

Whilst social and affective aspects are not explicitly mentioned, CLARION integrates reactive routine, learning and generic rules to create agents that can learn in a variety of contexts and adapt to different environments. Learning in this architecture is the most complex so far, with the usage of neuron networks [SPS02] or reinforcement learning techniques for implicit knowledge and an one-shot learning technique [SMP98] for explicit knowledge. CLARION focuses on cognition and learning and so, have limited representation of other dimensions (i.e. emotion, social factors and norm consideration).

3.3.5 PECS architecture

The PECS (i.e. Physical conditions, Emotional state, Cognitive capabilities and Social status) architecture's main objective is to allow the consideration of physical, emotional, cognitive and social influences in the agent's reasoning [Urb00].

The main components of PECS is illustrated in Figure 3.12. According to the author [Urb00], this architecture can be divided into three layers:

- The *Input Layer* includes Sensor and Perception and is in charge of processing input data.
- The *Internal Layer* models the internal state of the agent. It is composed of the *Physics, Emotion, Cognition and Social Status* components.
- The *Output Layer* calculates and executes actions. It consists of *Behaviour* and *Actor* components.

The *Sensor* component receives the inputs and passes them to the *Perception* component to perform information filtering mechanisms or other perceptional processes. The agent's internal state can comprise *physical, emotional, cognitive and social* attributes and processes. The *Behaviour and Actor* components have a repertoire of possible actions and can perform the action selection processes. The *Behaviour* component selects the individual actions or sequences of actions connected with the currently triggered activity. An action instruction is generated by the *Behaviour* component and handed over to the *Actor* component, where the action's execution is triggered and output to the environment.

Considering our dimensions, PECS covers many of them mainly because it is only a conceptual model. It utilises reaction-based and deliberative architectures. The affective and the social levels are covered by PECS components, though little details about their actual implementation. It does not represent norms in the architecture. The pre-defined update functions can theoretically be used for learning.

3.4 Socio-psychological inspired frameworks

The following frameworks use one or many socio-psychology theories to allow the user to adapt the agent's design and decision-making based on the context. In the next two subsections, we present two of the well-known frameworks within the

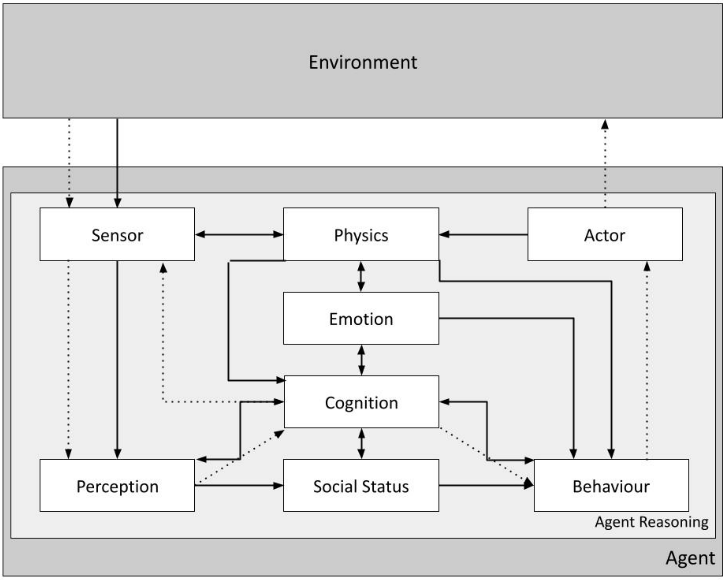


FIGURE 3.12: The PECS architecture, adapted from [Urb00]

ABM community. These two frameworks are selected due to their wide range of applicability to explain different human behaviours.

3.4.1 MoHuB framework

Schlüter et al. propose the Modelling Human Behaviour (MoHuB) framework that aims to support communicating and comparing different theories of human decision-making. It needs to be generic enough to capture the majority of theories and, at the same time, allow for a meaningful distinction between them [Sch+17]. The decision-making process within an individual is divided into three major parts: 1) what comes in (perception), 2) what goes out (behaviour), and 3) what happens in between (i.e., rules and representations that lead to the selection and execution of a behaviour).

Figure 3.13 illustrate MoHuB general agent architecture. Essentially, this architecture enhances Russell and Norvig's utility-based agent diagram [RN10, Chap. 2, p. 52] (see Figure 2.5 in Section 2.1.2) with components from behaviour theories. The agent is represented with several structural elements (i.e. *State* and *Perceived Behavioural Options* and processes involved in decision-making (i.e. *Perception*, *Evaluation*, *Behaviour*, *Selection*). In Figure 3.13, the solid arrows and corresponding ellipses indicate the flow of processes. The boxes represent structural elements. The dashed arrows represent the influence of one element on another. Decision-making involves both conscious and unconscious processes that lie at the interface of the individual,

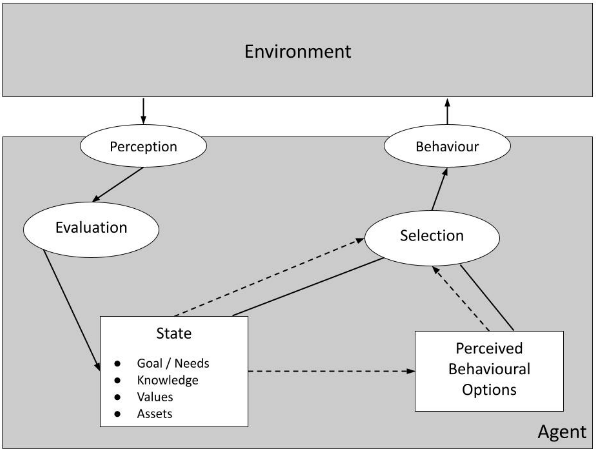


FIGURE 3.13: The MoHuB framework of individual decision-making, adapted from [Sch+17, p. 24]

the environment (perception and behaviour) and internal processes (evaluation and selection) [Sch+17].

An agent perceives the current state of its environment, evaluates the information and possibly updates its internal state. Next, the agent use its *State* and *Perceived Behavioural Options* to select the option that fulfils given goals/needs/satisfaction criteria. Finally, the option is executed and affects the state of the environment and its neighbours.

The *State* of an agent can constrain the original set of *Perceived Behavioural Options* or enhance it with new additional options derived from new knowledge. In addition, the set of *Perceived Behavioural Options* can also be updated over time due to processes of learning, forgetting or changing in intentions. The *State* may impact the *Selection* process by activating a new utility function. Furthermore, the *Perceived Behavioural Options* may influence the *Selection* process by excluding some options in the search routines (e.g. optimisation is not helpful for a set with only one option).

Different behaviour theories are described as alternative configurations (presence or specification) of an individual's structural elements, processes, and context. For example, *Rationality* in the MoHuB framework can be seen on Figure 3.14. It focuses on reflecting the self-interested needs of the homo-economicus paradigm, e.g., maximising expected utility. In this setting, an agent has complete knowledge about the system, and therefore, the *perceived behavioural options* include all possible options. Since the rational agent is *all-knowing*, the *Perception* and *Evaluation* components are excluded in this case. With unlimited cognitive capacity and the knowledge of all possible costs and constraints, an agent can always calculate and select the optimal option, i.e. maximising its utility function.

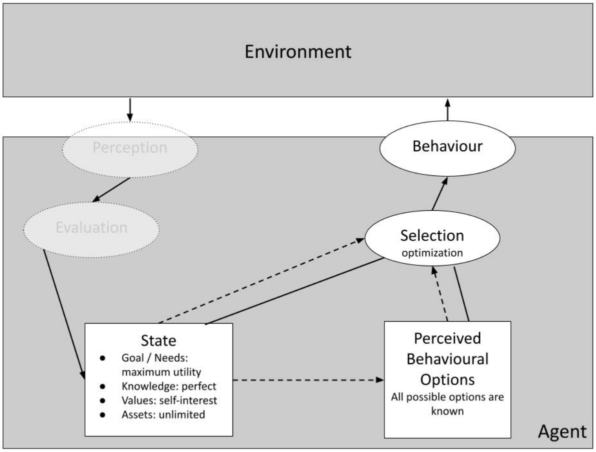


FIGURE 3.14: The MoHuB framework for rationality, adapted from [Sch+17, p. 27]

Similarly, the architecture of MoHuB can be mapped into other theories, including prospect theory, bounded rationality, descriptive norm, and habitual / reinforcement learning. Full descriptions of them can be found in [Sch+17]. With these mappings, the framework has been utilised in the applications of social-ecological systems, including agriculture [OASZ20; Hub+18] and energy system transitions [Koc+19].

Recently, Constantino et al. extended MoHuB with the addition of emotions and emotional states to the set of agent's characteristics [Con+21]. Individuals in this new framework act based on stable latent characteristics and a subset of context-specific situational ones, thus accounting for the influence of social and biophysical context.

Since MoHub is a conceptual framework, there is no direct way to establish the link between the decision-making aspects to the behaviour. In addition, the modeller needs to redefine the agent's components to adopt a new theory. This process can limit the reusability of the framework for different decision contexts.

3.4.2 Consumat framework

The Consumat model of Jager and Janssen was initially developed to model the behaviour of consumers and market dynamics [JJV01; Jag00] research topics, such as bio-economic [Hub+22], flood management [BV03] and transition to electric cars [JJB14]. Since its introduction, the Consumat approach has been used as a generic model of human behaviour based on people's decisions in satisfying their basic needs in various settings. Although the framework is not capable of simulating elaborate cognitive, reasoning processes or morality in agents, it does allow for simulating many key processes that capture human decision-making in a variety of contexts [JJ12].

The Consumat framework builds on three primary considerations: 1) human needs are multi-dimensional, 2) cognitive and time resources are required to make decisions, and 3) that decision-making is often done under uncertainty. Jager and Janssen base their work on the hierarchy of needs from the work of Maslow [Mas43] as well as Max-Neef [MN92]. However, due to the complexity of modelling and their interactions in an agent architecture, the Consumat framework only focuses on three types of needs: personal needs (satisfying one's personal taste, engaging in activities one likes and being different from others), social needs (having interaction with others), and existence needs (having means of existence, food, income, housing etc.). These types of needs may conflict with one another; consequently, an agent has to balance their fulfilment.

Jager and Janssen acknowledge that the resources available for decision-making are limited, thus constraining the number of alternatives an agent can reason about [JJ03]. Hence, they use the idea of *heuristics* to decrease the agent's cognitive effort and simplify complex decision-making problems. The different decision strategies translate into the sets of opportunities taken into consideration. The key rules are 1) the lower satisfaction is, the more involved one is in processing information on behavioural opportunities, and 2) the more significant uncertainty is, the more the behaviour of other people is used to identify attractive behavioural opportunities [JJ12].

An agent's current level of need satisfaction and uncertainty can be used to decide which heuristic can be applied in a specific situation. The Consumat agent has six different heuristics, which are based on the two dimensions of uncertainty and cognitive effort (see Figure 3.15). Starting from the left, when an agent has high need satisfaction (i.e. low cognitive effort required) and low uncertainty, it will simply *repeat* what it has been doing so far. Then, with a medium-low cognitive effort and low uncertainty, the agent compares its alternatives until it finds one that is *improving* its current state (satisfactory problem). If uncertainty is higher, it will try *imitating* the behaviour of agents with similar abilities. In the next step (i.e. medium-high cognitive effort and a low level of uncertainty), the agents apply a strategy where they determine the consequences of choosing an option one by one and stop as soon as they find one that *satisfies* their needs. Then, with the higher uncertainty level, the agent *compares* its performance with those in its networks. Finally, more cognitive effort is required in agents with the lowest level of need satisfaction in order to compute the consequences of all possible actions for a fixed amount of time and choose the action that can optimise its situation.

The satisfaction of a need is based on the utility derived from the current behaviour(s) and expectations of future utilities. The future outcomes can be more or less discounted depending on the type of need and decision. As a result, a discount function can be used to describe the importance of outcomes over time in deriving a level of need satisfaction:

$$N_y(x)t = \sum_{t}^{t=n} f(t) * U_y(x)t$$

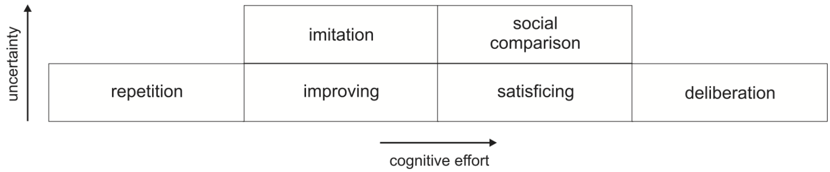


FIGURE 3.15: The six heuristics used in the Consumat Approach, according to [BG14]

where $N_y(x)t$ is need satisfaction of need y (one of the three) for using opportunity x at the current time step t. $t = 1..n$ is the time frame considered. $f(t)$ describes the decay function if decreased weighting of utility over time. $U_y(x)t$ is utility for need y of opportunity x at time t.

In 2012, Jager and Janssen presented an updated version which they refer to as Consumat II [JJ12]. It aims to improve several aspects of modelling, including 1) accounting for the expertise of agents in the social-oriented heuristics, 2) accounting for different agent capabilities in estimating the future, 3) lessening the difference between repetition on the one hand and deliberation on the other, and 4) considering several different network structures.

In terms of our dimensions, Consumat can represent several key processes that capture human decision-making in various situations. It has been used to study the effects of heuristics compared to the extensive deliberation approaches of other architectures [JJV01; BV03; Kan14]. Regarding the affective level, values and morality are considered, but emotion is not represented explicitly. The architecture considers legal norms and institutions but not social norms [Jag00]. However, as the agent compares its success to its peers, we consider it can capture a form of sociality. Finally, Consumat agents also improve their heuristics via learning from imitation of the actions of their peers, though they do not typically go beyond this comparison [BG14].

3.5 Current research gap

We provided an overview of different categories of agent architectures and frameworks in this chapter. Their comparison with our dimensions in Section 1.2 can be seen in Table 3.1.

The first class is BDI and its derivatives. They have a deliberative cycle and utilise formal logic-grounded semantics but require extensive computational resources. In addition, each of them only focuses on one particular dimension, such as emotion (eBDI) and norm (BRIDGE, BOID). The BRIDGE architecture can also cover emotional aspects using its ego component. As each focuses on a dimension of decision-making, no architecture in this category can explicitly represent all dimensions in our objectives, i.e. cognition, emotion, social factors, norm consideration and learning.

TABLE 3.1: Comparing dimensions covered by the related work

Architecture / Framework	**Cognition / Deliberation**	**Emotion**	**Social**	**Norm**	**Learning**
BDI	X				
BOID	X			X	
eBDI	X	X			
BRIDGE	X	X		X	
Deliberative Normative	X			X	X
NoA	X			X	
EMIL-A			X	X	X
SOAR	X				X
ICARUS	X				X
ACT-R/PM	X				X
CLARION	X				X
PECS	X	X	X		X
MoHuB	X	X	X	X	
Consumat	X		X	X	

Normative agent architectures (i.e. deliberative normative agents, NoA and EMIL-A) offer the ability to integrate social and individual factors to provide increased levels of fidelity concerning modelling social phenomena such as cooperation, coordination, organisation, and so on in a society of agents [BVDT07]. Additionally, it provides a tool to examine sociology through the perspective of methodological individualism [Neu08]. Methodological individualism attempts to build the foundations of sociology using individual actors and study the emergent phenomenon. To accomplish this, methodological individualism investigates the feedback mechanisms present in society and the system dynamics. The normative architecture suffers the same problem as the architectures of the BDI class. Their primary focus is on how norms are captured in human deliberation, whilst learning and emotion are limited.

The next category is cognitive architectures, which are based on cognitive theories and basic neurology concepts. Their main purpose is to understand how people organise knowledge, produce intelligent behaviour based on numerous facts derived from psychological experiments, and employ quantitative measures. However, these models do not consider social realism since they do not have the capabilities to incorporate different demographics, personality differences, situational and emotive variables and group dynamics [DDJ08]. Consequently, they do not cover the affective and social dimensions. In addition, neurological-inspired models often use AI technique to mimic the brain, such as neural networks [Par+19]. For ABM research, they lack the explanation power to create a causal link between the implementation of decision components to observed behaviours and, therefore, limit their usage in social studies.

One of the most established works is the MoHuB framework [Sch+17], which aims to include a different set of behavioural theories into formal models. Comparing it with the criteria in Section 1.2, this framework still does not meet all of them sufficiently. Firstly, MoHuB is a conceptual agent architecture framework. It provides a simple way to adapt a standard agent architecture to reflect different theories. Consequently, there is no systemic way to connect the effect of one determinant in decision-making to output behaviour or compare the impacts of multiple determinants. Secondly, MoHuB provides definitions of agent components based on a particular behaviour theory. They have to be redefined to implement a new theory, or additional processes are needed. This process requires a certain level of expertise and effort from the users. Consequently, it limits the reusability of the framework.

Finally, the Consumat framework has been used to study the effects of heuristics compared to the extensive deliberation approaches of other architectures. However, because the Consumat approach is rather complex, its formalisation for a specific domain requires more effort and deliberate choices than a (much simpler) rational actor approach. Especially if the modelled context becomes complex, it will require significant work to formalise the complete framework, the need-satisfying capacities and resource demands of many different opportunities. The researcher willing to make this effort must be convinced of the relevance of sophisticated behavioural dynamics in the system to be modelled [JJ12]. In addition, the concept design of Consumat agents is specific. Users cannot consider a different setup or interpretation. For example, on the social level, Consumat emphasises comparing the agent's own success and that of its peers. The Consumat agents are capable of comparing and reasoning about the success of their actions in relation to the success of their neighbours' actions, which are utilised for learning better behavioural heuristics. Nevertheless, they do not account for other social effects, such as the impact of the behaviour of others on their actions [BG14].

Depending on the theory and its relationship with empirical data, incorporating a theory in ABM models can have different degrees of explanatory power and fulfil different epistemological functions. Models can then be used to check the internal consistency of one or more theories or derive hypotheses that can be further tested through empirical research. In practice, however, it has been very limited to date [Gro+17; Sch+20b]. According to Schlüter et al. [Sch+17] and our surveys above, there are four challenges: 1) a large number of decision factors, 2) the focus of most theories on only a specific aspect of decision-making, 3) their varying degree of formalisation and, 4) the lack of specification of a causal mechanism.

Since the current related works cannot satisfy the objectives of our study (see Section 1.2), especially in terms of our dimensions and research focus, we propose another approach in the next chapter. While keeping a general agent architecture, the decision-making mechanism is redesigned to be based on a behavioural theory that takes into account multitudes of socio-psychology determinants. We then build an agent-based framework by formalising this mechanism and creating a code base to allow re-usability in different domains documented in different case studies in chapters 5 to 8.

Chapter 4

A behavioural decision-making framework for agents

Building an agent decision-making framework from a behavioural theory can provide a standard practice to incorporate multiple aspects relevant to social studies [Sch+17]. It can also provide a reference to facilitate the use of more sophisticated architectures in ABM research . To create a framework that can satisfy our objectives in Section 1.2, this chapter follows the four tasks proposed in [Sch+20b] to integrate human decision-making theories into models, including 1) selecting a theory, 2) formalising it in an agent architecture, 3) translating it into code and 4) documentation. We first perform a literature review of all relevant behavioural theories. They are then filtered, and the theory most suited for our research objective is selected. It is then formalised in the context of an agent-based framework for choice modelling. Two examples are given to show how it has been implemented in our current platform. They also emphasises the modularity potential of our framework. The Unified Modeling Language (UML) diagram and pseudocode of core components are provided in the following section, while the Java implementation can be seen in Appendix B. Next, we outline the steps to utilise our framework to create an agent-based model. Finally, we provide some comments on how the framework is documented.

4.1 Selection of a behavioural theory

To create a system that can mimic a human society, the first question to address is the origin of our behaviours. In socio-psychology, different schools of modelling have attempted to describe this process in the form of theories of human decision-making. Finding theories for one's own modelling project is a challenging aspect, which requires a systematic literature review process. It involves a process of identification and selection of existing studies which are relevant to clearly formulated research questions by using standardised methods [Hig+19; Aia+15].

To conduct this literature review, we follow a protocol from Popay et al. [Pop+06], which is illustrated in Figure 4.1. From Section 1.2, we defined our *research objectives*. Based on these objectives, the *conceptual boundaries* can be identified: theories that reason, explain and outline the mechanism of decision-making and behaviours in

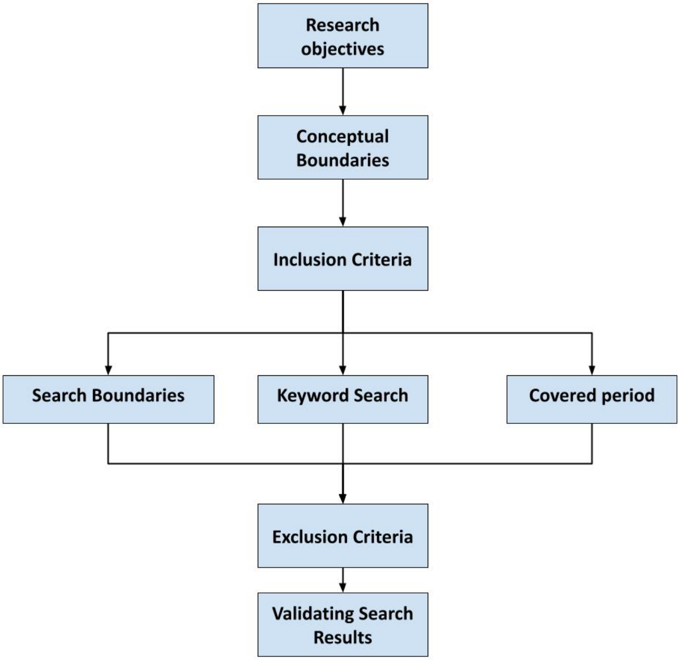


FIGURE 4.1: Literature review protocol, according to [Pop+06]

the field of psychology and social science. We further establish the *inclusion criteria*, which can be divided into three aspects:

- *Search boundaries*: Different types of sources are suitable to find individual decision-making theories, including:
  - Encyclopaedias, text books for decision theory, social psychology.
  - Reviews of theories in social psychology (e.g. [Dav+15; VLKH12; DHH10]).
  - Public database, i.e. Google scholar, Emerald, IEEE Xplore Digital Library, JSTOR, ProQuest Science direct, Wiley, Springer, PUBMED.
  - Informal method (e.g. asking experts, attending context-related conferences).
- *Keywords*: On one hand, modellers who already have knowledge about existing theories and the terminology used in these areas could search for theories addressing the specific modelling context. On the other hand, modellers who are not familiar with the discipline-specific terminology are hindered from searching for theories with appropriate keywords. For instance, "imitation", and "social norm" are keywords for theories dealing with copying behaviour

of others. In order to be comprehensive, all possible synonyms are used to express each of the concepts we aim to cover (i.e. attitudinal, emotion, social, normative, learning, habit). The indicative search strings used were: *social psychology theory, decision-making theory, behaviour theory, the agent-based model applied theory, attitudinal (viewpoint, frame of mind, perspective, position, belief, thoughts, intention) theory, normative (convention, standard, benchmark, point of reference, pattern, guideline) theory, social (role, contribution, community, collective, group, responsibility) theory, emotion (self, feeling, sentiment) theory, habit (trait, pattern, routine, characteristic) theory*.

- Covered period: The starting date is not considered, but the end date is December 2020.

After the initial search with the above criteria, 47 well-defined theories are found, including:

- Action Identification Theory
- Attachment Theory
- Balance Theory
- Broaden-and-Build Theory of Positive Emotions
- Cognitive Dissonance Theory
- Correspondent Inference Theory
- Drive Theory
- Dual Process Theories
- Dynamic Systems Theory
- Equity Theory
- Error Management Theory
- Escape Theory
- Excitation-Transfer Theory
- Implicit Personality Theory
- Inoculation Theory
- Interdependence Theory
- Learning Theory
- Logical Positivism
- Opponent Process Theory
- Optimal Distinctiveness Theory
- Prospect Theory
- Realistic Group Conflict Theory
- Reasoned Action Theory
- Reductionism
- Regulatory Focus Theory
- Relational Models Theory
- Role Theory
- Self-Affirmation Theory
- Self-Categorization Theory
- Self-Determination Theory
- Self-Discrepancy Theory
- Self-Expansion Theory
- Self-Perception Theory
- Self-Verification Theory
- Social Exchange Theory

- Social Identity Theory
- Social Impact Theory
- Sociobiological Theory
- Stress Appraisal Theory
- Symbolic Interactionism
- Temporal Construal Theory
- Terror Management Theory
- Theory of Mind
- Theory of Reasoned Action
- Theory of Planned Behaviour
- Theory of Interpersonal Behaviour
- Threatened Egotism Theory

The theories' main ideas and references are listed in Appendix A.1. From this list, we narrow them down using the following *exclusion criteria*, based on our objectives in Section 1.2 and technological requirements in Section 2.4:

- The theory covers multiple aspects, or so-called determinants, of decision-making.
- The theory has a way to derive a utility function and, potentially, an additive value function (see 2.1.1) to evaluate the outcomes of actions from different determinants.
- The theory focuses on individual behaviours, not on evaluating the actions of others.
- The theory's concepts can be organised in a structure to mark the flow of information between the components and their contribution to the final action output.

The three theories satisfying these criteria are: Ajzen and Fishbein's Theory of Reasoned Action (TRA) [FA75], Ajzen's Theory of Planned Behavior (TPB) [Ajz85] and Theory of Interpersonal Behaviours (TIB) [Tri77]. In the following subsections, we discuss them in detail and provide a reason to choose TIB as the foundation of our framework. In the following section, we use the term *behaviour* to describe the output of a theory, while the term *action* is an output of the agent's decision-making.

4.1.1 Theory of Reasoned Action (TRA)

Developed by Martin Fishbein and Icek Ajzen in 1967, the primary purpose of TRA explains the relationship between attitudes and behaviours within human actions. Its main predictor is a person's intention to perform a behaviour, while the two main determinants of behavioural intention are attitudes and norms (see Figure 4.2). By further examining them, researchers can understand whether or not one will perform the intended action.

A high correlation of *Attitude* and *Subjective Norm* to *Behaviour Intention*, and subsequently to *Behaviour*, has been confirmed in many studies (see survey in [SHW88]).

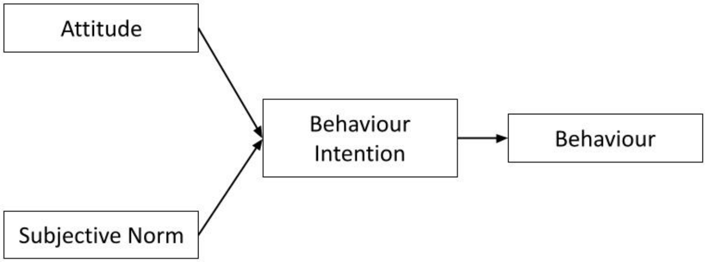


FIGURE 4.2: Theory of Reasoned Action, adapted from [FA75, p. 340]

Hence, according to TRA, *Behaviour Intention* is a function of both *Attitude* and *Subjective Norm*:

$$BI = (AB)W_1 + (SN)W_2 \tag{4.1}$$

where BI is the *Behaviour Intention*, AB is one's *Attitude* toward performing the behaviour, W is empirically derived weights, and SN is one's *Subjective Norm* related to performing the behaviour.

Although its scope is wide, TRA still has a major limitation: it cannot be used to predict behaviours that require access to certain opportunities, skills, conditions, and resources since it does not consider that certain conditions that enable the performance of a behaviour are not available to individuals [EC93]. The authors also acknowledged that:

> "... some behaviours are more likely to present problems of controls than others, but we can never be certain that we will be in a position to carry out our intentions. Viewed in this light, it becomes clear that, strictly speaking, every intention is a goal whose attainment is subject to some degree of uncertainty." [Ajz85]

4.1.2 Theory of Planned Behavior (TPB)

TPB was proposed in [Ajz85] as an extension to TRA. It adds a new component, *perceived behavioural control*, which refers to the degree to which a person believes that they control any given behaviour (see Figure 4.3). TPB suggests that people are much more likely to enact certain behaviours when they feel that they can enact them successfully.

In a simple form, behavioural intention for the Theory of Planned Behavior (TPB) can be expressed as the following mathematical function:

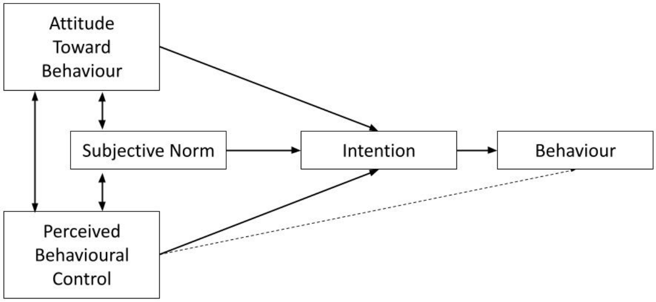


FIGURE 4.3: Theory of Planned Behaviour, adapted from [Ajz05, p. 118]

$$
\begin{aligned}
BI = w_A A + w_{SN} SN + w_{PBC} PBC \\
\text{provided:} \\
A \propto \sum_{i=1}^{n} b_i e_i \\
SN \propto \sum_{i=1}^{n} n_i m_i \\
PBC \propto \sum_{i=1}^{n} c_i p_i
\end{aligned}
\tag{4.2}
$$

where BI is the behavioural intention. A is the attitude toward behaviour. b is the strength of each belief concerning an outcome or attribute. e is the evaluation of the outcome or attribute. SN is the subjective norm. n is the strength of each normative belief of each referent. m is the motivation to comply with the referent. PBC is the perceived behavioural control. c is the strength of each control belief. p is the perceived power of the control factor. w is the empirically derived weight/coefficient.

The perceived behavioural control can, together with intention, be used to predict behaviour:

$$B = w_{BI} BI + w_{PBC} PBC \tag{4.3}$$

where B is the behaviour. BI is the behavioural intention. PBC is the perceived behavioural control. w is the empirically derived weight/coefficient.

There are several limitations of the TPB, including the assumption that behaviour results from a linear decision-making process and does not consider that it can change over time. Besides, it does not account for other variables that influence behavioural intention and motivation, such as emotion or experience.

4.1.3 Theory of Interpersonal Behaviour (TIB)

TIB was proposed by Harry Triandis in 1977 [Tri77]. Compared to the TRA and the TPB, Triandis incorporates another key role of *habit* and affective factors in explaining behaviour (his tri-level model can be seen in Figure 4.4). TIB states that interpersonal behaviour is a complex and multi-aspects phenomenon. In any interpersonal decision, a person's behaviour is determined based on what that person perceives to be appropriate in that particular situation and his/her social pressures. The first level of TIB concerns the way personal characteristics and prior experiences shape personal attitudes, beliefs and social determinants related to the behaviour. The second level explains how attitude, affect, and social factors influence the intentions of a specific behaviour. Finally, the third level states that intentions of a behaviour, prior experience and situational conditions predict whether or not the person will perform the behaviour in question.

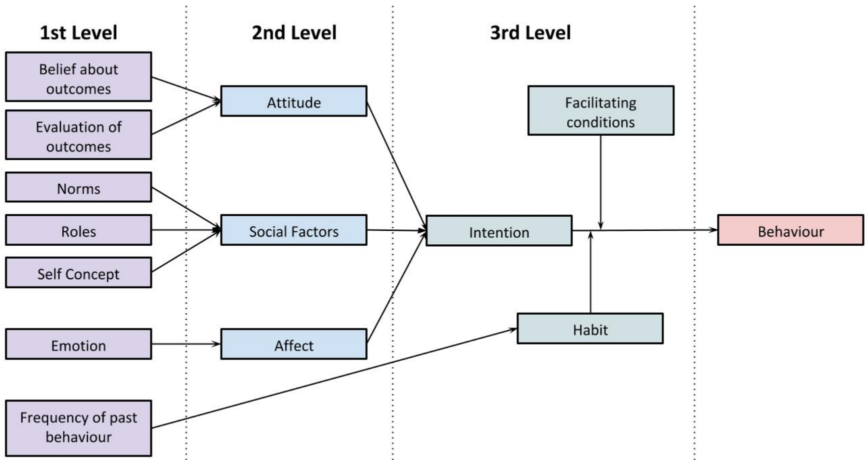


FIGURE 4.4: Triandis' tri-level model, adapted from [Jac05]

According to Triandis, the determinant of a behaviour depends on three major factors: 1) the strength of the habit emitting the behaviour, which is indexed by the number of times the behaviour has already occurred in the history of the individual, 2) the behavioural intention to emit the behaviour, and 3) the presence or absence of the condition that facilitates the performance of the behaviour. We can express this idea using the following equation:

$$P_b = (w_H * H + w_I * I) * F \tag{4.4}$$

where P_b is the probability of an behaviour and varies from 0 to 1. H is the habit and is measured by the number of times the act has already been performed by the person. I is the intention to action and "is a cognitive antecedent of an act" [Tri77, p.

5]. We will extend this definition in the next formula. F refers to Facilitating condition, such as the ability of the person to carry out the act or his/her knowledge. The w_H and w_I refer to the normalised weights of the habit and intention components. These weights can be determined by a statistical procedure called *multiple regression analysis*, which measures the correlation among variables under consideration.

The determinants of behavioural intentions may be expressed in a simple equation:

$$I = w_S(S) + w_A(A) + w_C(C) \tag{4.5}$$

where I is the intention to act. S is social factors. A is the affection attached to the behaviour itself. w_S , w_A and w_C are normalised weights which can be derived similarly to the previous function. The value of perceived consequences of the behaviours, C, depends on the sum of the products of the *subjective probability* that a particular consequence will follow a behaviour (P_c) and the value of that consequence (V_c). Hence, it can be further expressed as follow:

$$C = \sum_{i=1}^{n} P_{ci} V_{ci} \tag{4.6}$$

where n is the number of consequences that a subject perceives as likely to follow a particular behaviour. This equation can be utilised for uncertainty in an agent's perception of a partially observable environment.

From the description above, we can see that the three theories (TRA, TPB and TIB) provide comprehensive understandings of what determines individual behaviour and are helpful in explaining complex human thought processes, which are influenced by social and physical environments. These processes' determinants are organised in a well-defined structure. Each theory also includes an *additive value function* to compute the expected utility of an option. However, these theories cover different dimensions from our research objective, as shown in Table 4.1.

TABLE 4.1: Comparing dimensions covered by TRA, TPB and TIB

Theory	Cognitive / Mental Attitudinal	Emotion	Social	Norm	Learning
TRA	X			X	
TPB	X			X	
TIB	X	X	X	X	

While *cognitive* and *norm* are explicitly mentioned in the TRA and TPB, other determinants are not clearly shown. One of the major advantages of TIB is its diverse combination of behavioural variables that have been neglected in other behavioural theories [FB10]. It provides a sound conceptual and theoretical background to test unique variables such as social pressure, expectations, and habits [BFS10]. Moreover, TIB informs a wide range of research designs and methodologies as it is a flexible tool that identifies a set of potentially relevant factors and their interactions [GN10].

Thanks to its broader set of determinants, TIB is flexible enough to reflect other behaviour theories by exchanging psychology elements and assigning weights to mark their contribution to the agent's decision-making process. For example, we can cut off the determinants of *Habit*, *Social* and *Affect/Emotion* in Figure 4.4 to create a resemblant of TPB model. Therefore, we choose TIB to build the theoretical foundation for our framework.

However, in terms of our dimensions, TIB does not explicitly cover the learning aspect as it does not describe how people internalise or revise their beliefs using the observed facts. One solution for this issue is building an agent decision-making architecture incorporating feedback loops. In other words, the agents review the macro patterns as a result of their earlier action(s) and decide what to do in the future. This design is further discussed in the next section.

4.2 Formalisation of the theory

As presented in Chapter 2, there are several elements that need to be considered in the agent's decision-making, including the architecture, core components and utility function. In this section, we provide a detailed selection of them and illustrate this with two working examples from the mobility and vehicle purchasing domains.

4.2.1 Selection agent architecture

Different architectures can be selected depending on the type of agent that has to be built. Three architecture classes, logic-based, BDI and layered architectures, are relevant to our objectives (see Section 2.1.3). In this subsection, we provide their details and select the most suitable one for the TIB theory.

Logic-based architecture is based on the view that intelligent behaviour can be generated in a system by giving that system symbolic representation and its desired behaviour and syntactically manipulating this representation [Woo99, p. 42]. Hence, their primary purpose is theory proving or logical deduction, i.e. how an agent applies its own plan/theory to derive the goal results. If the theory about the agent behaviours is correct, it should arrive at the end goal. The problem with using this class of architecture for social research is the difficulty in representing the properties of a dynamic, real-world environment. In addition, the computational complexity of theory proving makes it difficult for the agent can operate efficiently in a time-constrained environment. An agent design that moves away from strictly logical representation languages and deduction rules can perform better [Cal+21]. However, this solution can also lose the most significant advantage of this class of agents: simple, elegant semantics [Woo99, p. 47].

BDI architecture provides a general flow of information that resembles how we reason in real life. However, as it also utilises a form of logical semantics, it suffers from the problem of representing complex, dynamic environments. A further weak point of BDI architectures (and consequently of BDI agents) is that their associated agent language is often severely restricted as it consists of literals (i.e., propositional symbols or their negations) based on the theory of human practical reasoning [Bra87]. Therefore, its conceptual design and flow of information do not match our selected theory, TIB. As discussed in Section 3.1, a pure BDI agent lacks the affective, social and norm dimensions of our objective. One attempt is to include all of them in its decision-making cycle. However, it would further make the belief revision functions more complex and not suitable for time-constrained, dynamic environments.

We now consider the layered architectures. In a horizontal layering, the input is processed by several components, and then this information is aggregated to determine the output/action of the agent. In a one-pass layered architecture, the input is processed in each stage and then delegated to the next one. A combination of these two architectures is suitable for our purpose of creating a modular framework that passes through different levels of determinants. As TIB has a tree-like/hierarchical structure, the flow of information can be organised similarly. In particular, determinants from the same level (e.g. norm, role, self-concept) can be put in layers of the horizontal layout. They are the aspects that can be evaluated individually from one another. Hence, the utility functions of these determinants can be applied at the same time. Their results can be passed from one level to another following the one-pass structure. In this case, an additive value function can be used to combine the utilities from the ancestors, i.e. the determinants connected in the previous level. The same process is repeated until it arrives at the *behaviour output*. The flow of the decision-making procedure between the components of the agents can also follow the one-pass architecture. Details are provided in the following sections.

4.2.2 Concepts utilised in the agent framework

Our agent framework utilises the following core concepts, which can also be seen in in Figure 4.5:

- Environment: The (social and biophysical) environment in which the individuals are embedded in.
- Environment state: The current state of an environment perceived by the individual. It can also include the communication signals from other agents.
- Perception: The process by which an individual senses the surrounding social and biophysical environment.
- Perceived options: The set of options the individual perceives and can choose from. This set can be created by the modeller and then filtered by the agent's *Perception*.
- Memory: The storage of an individual internal state as well as history of its actions and feedbacks.
- Evaluated option: The result of decision-making. It is usually in form of a probability that an action will be performed. In case of choice modelling, the individual can output only the top evaluated action.

- Feedback: The results of an action given by the environment after an action by individual. It can be used to update the memory.
- Communication: The action that an individual executes as a result of the decision process. The action impacts the socio-environmental system and, in addition to perception, is the second interface between an individual and its environment. Selected actions may fail to be executed if the action is physically impossible.

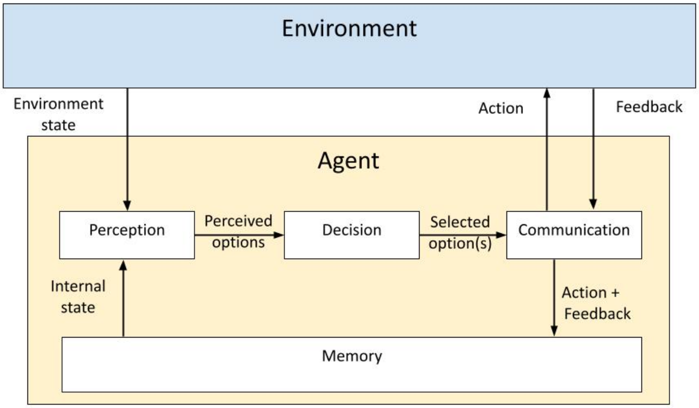


FIGURE 4.5: Overview of agent's architecture

A typical decision-making cycle is as follows: When a task is assigned, the *Perception* observes the current state of the *Environment* and combines them with the agent's internal state to produce a list of perceived options. Then, they are given to the *Decision* unit to be evaluated. Details of this process are described in Section 4.2.3. The *Communication* component then utilises this result to execute the chosen option(s) with *Environment* and other agents. The *Environment* can then provide feedback(s) based on the nature of the system associated with the action. The agent remembers these feedbacks in the *Memory*, which can then be used to modify the probability of expected values in future decision-making.

4.2.3 Utility Function

A full decision-making cycle with TIB determinants is illustrated in Figure 4.6. It should be noted that the description below is modified version of the ones provided in our previous publications [NS19b; NS19a; NS20b; NS20a].

As the theory has a tree structure, in which a node represents a determinant. It can be connected to many children (connected node of the next level) but must be connected to exactly one parent (precedent node), except for the root node, which has no parent. Each determinant in this TIB module represents a utility function to evaluate expected utility from the outcomes of an option.

An agent is given a list of tasks that isolated decision-making task needs to be sequentially executed and a list of actions. To perform a task, the agent first filters the list of actions with the information from its internal state and the external environmental state to generate a set of possible options. As an example, using our case study below (see Section 4.2.4), the agent observes the available modes of transportation around its local environment (e.g. bus/tram/train) and combines them with other modes in its capabilities (e.g. car, bike, walking) to create a list of available options to choose from.

For all determinants (d) in TIB, each option (o) is then given an utility value which comes from comparing its properties with other's ($U_o(d)$). In the first level, this value can be in the form of a cardinal utility measure (for determinants such as price or time) or ordinal utility ranks (for determinants such as emotion). Both of them can be calculated from empirical data (e.g. census, survey) or calibrated with experts' knowledge and stakeholders' assessment. The results for these determinants are then multiplied with an normalised weights (called $w(d)$). This process is captured in the following equation, which is adapted from the additive value function in Section 2.1.1:

$$U_o(d) = \sum_{c=1}^{C} U_o(c) * w(c) \tag{4.7}$$

where $U_o(d)$(opt) is the utility value of an option o at determinant d. C is the set of all children c of d, i.e. determinants connect with d in the previous level. Therefore, $c \in C$ where $connect(d, C)$. $w(c)$ is the normalised weight of child determinant c. This weight represents the importance of a decision-making determinant compared to others at the same level and emphasises the heterogeneity of individuals. It also allows the modeller to express a certain theory by cutting determinants that are not relevant to a case study, i.e. setting their weights to 0. The combination process then continues until it arrives at the behaviour output list, whose utility values can be translated to the probabilities that the agent will perform that option. If the agent is assumed to be deterministic, it picks the option that is correlated to the highest or lowest utility, depending on the modeller's assumptions.

In the first level in Figure 4.6, to translate the determinant *belief* and *evaluation* to a practical application, we use the concept of consequences, i.e. outcomes of an action. Triandis defines *belief* as the chance (or percentage) that a consequence will happen [Tri77]. The *evaluation* gives the expected utility value to that consequence. We will demonstrate the usage of this function in the two following examples. They also show the modularity of the framework by changing the TIB mapping.

4.2.4 A mobility running example

Table 4.2 shows a running example in the mobility domain which follows the TIB determinants mapping in Figure 4.7. An agent needs to make a working trip and has

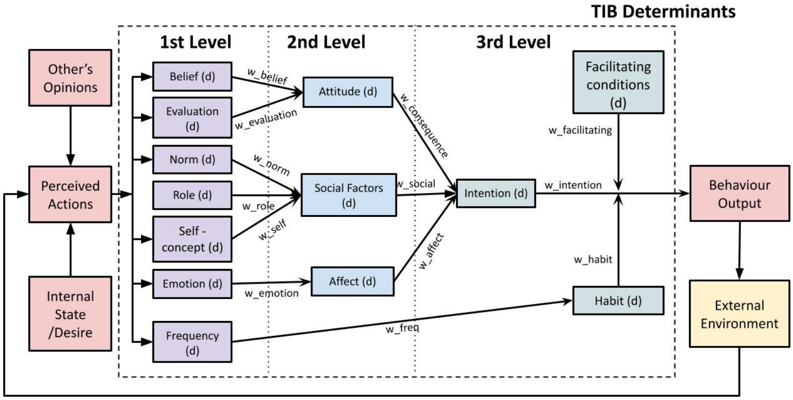


FIGURE 4.6: Agent decision-making mechanism with TIB Module

access to three options: using *car*, taking *train* or *bike*. In addition, U is a cost function, i.e. option that has smaller value is preferred.

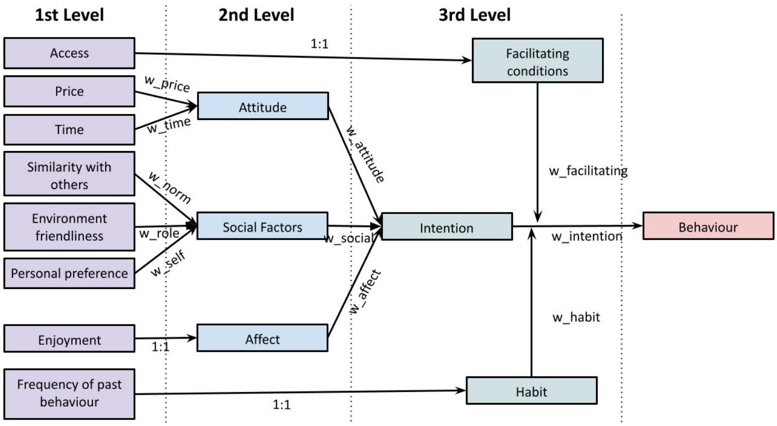


FIGURE 4.7: Example of mobility decision with TIB's determinant

In this case, we interpret determinant *evaluation* as the cost of choosing a mode of transportation in terms of *price* and *duration*. In addition, an agent is assumed to believe completely in its evaluation (*belief* = 100%), i.e. the environment is deterministic and fully observable. Other determinants in the first level can be interpreted as follow:

- Norm: ranking of a mode based on the number of neighbours using it.
- Role: ranking of modes based on their environmental friendliness.

- Self-concept: ranking of the modes based on the agent's personal reference.
- Frequency of past actions: the number of times the agent used the mode previously.

TABLE 4.2: Example of in mobility context

Level	Determinant	w	EU
1st	Evaluation (Price - Swiss franc), Belief = 100%	2	$U_{car} = 4$ $U_{train} = 3$ $U_{bike} = 0$
	Evaluation (Duration - hours), Belief = 100%	4	$U_{car} \approx 0.3$ $U_{train} \approx 0.2$ $U_{bike} \approx 1$
	Norm (similarity with others)	3	$U_{train} = 1$ $U_{car} = 2$ $U_{bike} = 3$
	Role (environmental friendliness)	2	$U_{car} = 3$ $U_{train} = 2$ $U_{bike} = 1$
	Self-concept (personal preference)	3	$U_{car} = 1$ $U_{train} = 2$ $U_{bike} = 3$
	Emotion (enjoyment)	1	$U_{car} = 1$ $U_{train} = 2$ $U_{bike} = 3$
	Frequency (past similar trips - note that lower value means more usage)	3	$U_{car} = 0$ $U_{train} = 0$ $U_{bike} = 1$

Continued on next page

Table 4.2 – continued from previous page

Level	Determinant	w	EU
2nd	Consequence (Evaluation + Belief)	4	U_{car} = 4*2/7 + 0.3*4/1.5 ≈ 1.94 U_{train} = 3*2/7 + 0.2*4/1.5 ≈ 1.39 U_{bike} = 0*2/7 + 1*4/1.5 ≈ 2.67
	Social factors (Norm + Role + Self-concept)	2	U_{car} = 1*3/6 + 3*2/6 + 1*3/6 = 2 U_{train} = 2*3/6 + 2*2/6 + 2*3/6 ≈ 2.67 U_{bike} = 3*3/6 + 1*2/6 + 3*3/6 ≈ 3.33
	Affects (Emotion)	2	U_{car} = 1*1/6 ≈ 0.17 U_{train}= 2*1/6 ≈ 0.33 U_{bike} = 3*1/6 = 0.5
3rd	Intention (Attitude + Social factors + Affect)	4	U_{car} = 1.94*4/6 + 2*2/8 + 0.17*2/1 ≈ 2.13 U_{train} = 1.39*4/6 + 2.67*2/8 + 0.33*2/1 ≈ 2.26 U_{bike} = 2.67*4/6 + 3.33*2/8 + 0.5*2/1 ≈ 3.61
	Habit (Frequency)	3	U_{car} = 0*3/1 = 0 U_{train} = 0*3/1 = 0 U_{bike} = 1*3/1 = 3
	Facilitating conditions (lower mean easier to access)	2	U_{car} = 0 U_{train} = 0 U_{bike} = 0
	Behaviour output		U_{car}= 2.13*4/7.67 + 0*3/3 + 0*2/1 ≈ 1.11 U_{train} = 2.26*4/7.6 + 0*3/3 + 0*2/1 ≈ 1.18 U_{bike} = 3.28*4/7.6+ 3*3/3 + 0*2/1 ≈ 4.73

The agent expects the *car* option would have the price around 4 Swiss Franc, and so $U_{car}(Price) = 4$. Correspondingly, $U_{train}(Price) = 3$ and $U_{bike}(Price) = 0$. Their total value, $\sum U(Price)$, is 7. The estimations for duration are $U_{car}(Duration) \approx 0.3$, $U_{train}(Duration) \approx 0.2$ and $U_{bike}(Duration) \approx 1$; the sum of which is 1.5. In this example, we assume that the agent's $w(Price)$ and $w(Duration)$ are 2 and 4 respectively. By applying Equation 4.7, the new expected value in next level ($U(Attitude)$) of *car* would be = 4*2/7 + 0.3*4/1.5 ≈ 1.94, *train* would be 3*2/7 + 0.2*4/1.5 ≈ 1.39, and *bike* would be 0*2/7 + 1*4/1.5 ≈ 2.67. Hence, according to determinant Consequence, *train* would have the highest chance to be picked, followed by *car* and *bike*.

For non-measurable value such as Norm, the agent uses the concept of reputation (popularity) to rank the options: $U_{train}(Norm) = 1$, $U_{car}(Norm) = 2$, $U_{bike}(Norm) = 3$ (best to worst); the sum of which is 6. The same values are applied for determinant *Self-concept*. On the contrary, the agent might have an environmental consciousness and rank these mode in the opposite order, e.g. $U_{bike}(Role) = 1$, $U_{train}(Role) = 2$ and $U_{car}(Role)$(car) = 3. According to the data in a survey [Web+17], $w(Norm) = 3$, $w(Role) = 2$ and $w(Self) = 3$. By combining these social factors using Equation 4.7, we then have $U_{car}(Social) = 2$, $U_{train}(Social) \approx 2.67$ and $U_{bike}(Social) \approx 3.33$. Because expected utility is a cost function, i.e. agents prefer lower value, inverse values of determinant *Frequency* and *Facilitating Conditions* are used in Table 4.2.

The process continues with other determinants on different levels (see Table 4.2 and Figure 4.6) until the agent reaches its behaviour output, where expected values are $U_{car} \approx 1.11$, $U_{train} \approx 1.18$ and $U_{bike} \approx 4.73$. These utilities indicate that car would be the best option for this agent. We choose this example to highlight the importance of social factors in decision-making because the best choice would have been using the train if the agent only makes an evaluation based on Price and Duration (see 2nd Level in Table 4.2).

4.2.5 A car purchasing running example

Table 4.3 shows a running example of an agent need to make a vehicle purchase, which follows the TIB determinants mapping in Figure 4.8. The agent is assumed to have a choice between the three type of models: *diesel*, *gasoline* and *electric vehicle(EV)*. In this case, U is also a cost function, i.e. option that has smaller value is preferred.

- Role: recommendation from the media and dealers (value from 0 to 1) .
- Norm: recommendation from the neighbours (value from 0 to 1) .
- Self-concept: the car's from the brand preferred by the agent (value 0/1).
- Frequency of past actions: the same model was owned by the agent before (value 0/1).

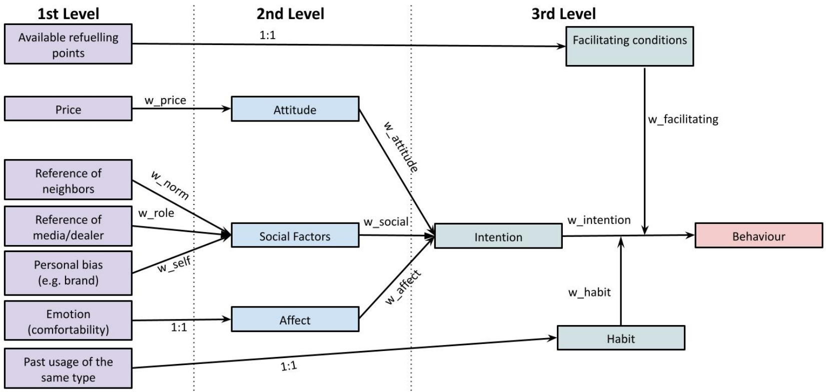


FIGURE 4.8: Example of vehicle purchasing decision with TIB's determinant

TABLE 4.3: Example in vehicle purchasing context

Level	Determinant	w	EU
1st	Evaluation (Price - thousand Swiss franc), Belief = 100%	1	$U_{diesel} = 25$ $U_{gasoline} = 30$ $U_{EV} = 45$
	Norm (recommendation from neighbours)	3	$U_{diesel} = 0.5$ $U_{gasoline} = 0.5$ $U_{EV} = 0.8$
	Role (recommendation from media/dealers)	2	$U_{diesel} = 0.3$ $U_{gasoline} = 0.5$ $U_{EV} = 0.75$
	Self-concept (brand)	3	$U_{diesel} = 1$ $U_{gasoline} = 1$ $U_{EV} = 0$

Continued on next page

Table 4.3 – continued from previous page

Level	Determinant	w	EU
	Emotion (comfortability)	1	$U_{diesel} = 2$ $U_{gasoline} = 1$ $U_{EV} = 1$
	Frequency (past ownership)	3	$U_{diesel} = 0$ $U_{gasoline} = 0$ $U_{EV} = 1$
2nd	Consequence (Evaluation)	4	$U_{diesel} = 25$ $U_{gasoline} = 30$ $U_{EV} = 45$
	Social factors (Norm + Role + Self-concept)	2	$U_{diesel} = 0.5*3/1.8 + 0.3*2/1.55 + 1*3/2 \approx 2.72$ $U_{gasoline} = 0.5*3/1.8 + 0.5*2/1.55 + 1*3/2 \approx 2.98$ $U_{EV} = 0.8*3/1.8 + 0.75*2/1.55 + 0*3/2 \approx 2.3$
	Affects (Comfortability)	2	$U_{diesel} = 1*1/6 \approx 0.17$ $U_{gasoline}= 2*1/6 \approx 0.33$ $U_{EV} = 3*1/6 = 0.5$
3rd	Intention (Attitude + Social factors + Affect)	4	$U_{diesel} = 25*4/100 + 2.72*2/8 + 0.17*2/1 \approx 2.02$ $U_{gasoline} = 30*4/100 + 2.98*2/8 + 0.33*2/1 \approx 2.61$ $U_{EV} = 45*4/100 + 2.3*2/8 + 0.5*2/1 \approx 3.38$
	Habit (Frequency)	3	$U_{diesel} = 0*3/1 = 0$ $U_{gasoline} = 0*3/1 = 0$ $U_{EV} = 1*3/1 = 3$
	Facilitating conditions (lower mean easier to access)	2	$U_{diesel} = 0$ $U_{gasoline} = 0$ $U_{EV} = 0$

Continued on next page

Table 4.3 – continued from previous page

Level	Determinant	w	EU
	Behaviour output		U_{diesel}= 2.02*4/8.01 + 0*3/3 + 0*2/1 $\approx$ 1.01 $U_{gasoline}$ = 2.26*4/7.6 + 0*3/3 + 0*2/1 $\approx$ 1.19 U_{EV} = 3.28*4/7.6+ 3*3/3 + 0*2/1 $\approx$ 4.72

Using a catalogue (e.g. [QE20]), the agent expects the *diesel* option would have the price around 25'000 Swiss Franc, and so $U_{diesel}(Price)$ = 25'000. Correspondingly, $U_{gasoline}(Price)$ = 30'000 and $U_{EV}(Price)$ = 45'000. Since *Price* the only children node of *Attitude*, the utility values of all the models are kept the same, i.e. $U(Attitude) = U(Price)$.

For the *Role* determinant, the agent can calculate the utility of each model as the normalise value of the reviews provided by newspapers, online platform and local dealers. For this example, we assume that $U_{diesel}(Role) = 0.3$, $U_{gasoline}(Role) = 0.5$ and $U_{EV}(Role) = 0.75$. Their total sum is 1.55. Similarly, the utility value of *Norm* determinant can be derived from reviews of its neighbours. We also assume that $U_{diesel}(Norm) = 0.5$, $U_{gasoline}(Norm) = 0.5$ and $U_{EV}(Norm) = 0.8$. Their sum is 1.8. To assess *Self-concept* determinant, the agent can assign value 1 for the model of a brand in its preference and value 0 for the rest. Hence, we have $U_{diesel}(Self) = 1$, $U_{gasoline}(Self) = 1$ and $U_{EV}(Self) = 0$. According to the data in a survey [Web+17], $w(Norm) = 3$, $w(Role) = 2$ and $w(Self) = 3$. By combining these social factors using Equation 4.7, we then have $U_{car}(Social) \approx 2.72$, $U_{gasoline}(Social) \approx 2.98$ and $U_{EV}(Social) \approx 2.3$.

The process continues with other determinants on different levels (see Table 4.2 until the agent reaches its behaviour output, where expected values are $U_{diese}l \approx 1.01$, $U_{gasoline} \approx 1.19$ and $U_{EV} \approx 4.73$. These utilities indicate that the diesel would be the best option for this agent.

4.3 Translating the formalisation into code

To enable reusability and widen the applicability of our framework to various social domains, the concepts and mathematical foundation above need to be translated into code. This section provides an Unified Modeling Language (UML) diagram of relevant classes to implement the core components (i.e. *Perception*, *Memory*, *Communication*, *Decision*) and concepts mentioned above and the general flow of the decision-making cycle. Most of them involve abstract functions that can be implemented with glue code to connect the functionality of the agent to the modelling context. In the *Decision* component, the implementation for leaf node (i.e. the node that has no children) and parent node determinants are also presented. This mechanism allows users to interchange, add or remove the decision factors to present other behaviour theories. All implementations in Java are listed in Appendix B.

4.3.1 Specification of classes in the framework

We illustrate the relationships between the main classes mentioned in Section 4.2 in Figure 4.9:

- Abstract class **Task** describes the work to be done or undertaken by the agent. It specifies a trigger time and potentially some pre/post-conditions that the agent has to satisfy, e.g. some environment externals state or agent's internal state have to be reached.
- Interface **Environment State** is a current model of the environment. Users can implement some getter and setter functions for this information. It can also include communication signals with other agents.
- Interface **External State** stores the information about the current internal state of an agent. Some getter and setter can be implemented for these fields.
- Interface **Option** is a potential action that an agent can do to accomplish the Task.
- Interface **Feedback** is an optional class which provides the data about the reaction of the current environment when the agent turns the Option to action using its communication component.
- Abstract Class **Environment** includes an id and two functions: *getEnvironmentState* and *getFeedback*. The first provides the environment's current state. The second, which is optional, indicates the reaction of the environment in response to the agent's behaviour.
- Abstract Class **PerceptionComponent** stores the address of the environment and has a function - *generateOptions* - that creates a list of options with their properties based on the environment state and internal state. The generated options should satisfy the pre-condition mentioned in the Task.
- Abstract Class **MemoryComponent** has information of an agent's internal state. It is responsible to provide a suitable information to generate options and incorporate feedback into the internal state using two functions, *getInternalState* and *updateInternalState*.
- Abstract Class **DecisionComponent** evaluates the Option from perception component and selects the suitable option(s) for the given Task. The *evaluateOptions* function utilises tree structure to compute the expected utility value, according to Equation 4.7. The default setting is the TIB structure. It starts at a root node *Determinant* represents the behaviour output value. A node with one or more children is a *ParentDeterminant*. On the contrary, it can also have zero children, i.e. a *LeafDetermiant*. In our TIB model, the *LeafDetermiant*s represent the first level determinants, while the *ParentDeterminant*s is the one on levels two and three. To implement this structure, users only have to specify how the Option can be evaluated in terms of the *LeafDetermiant*.

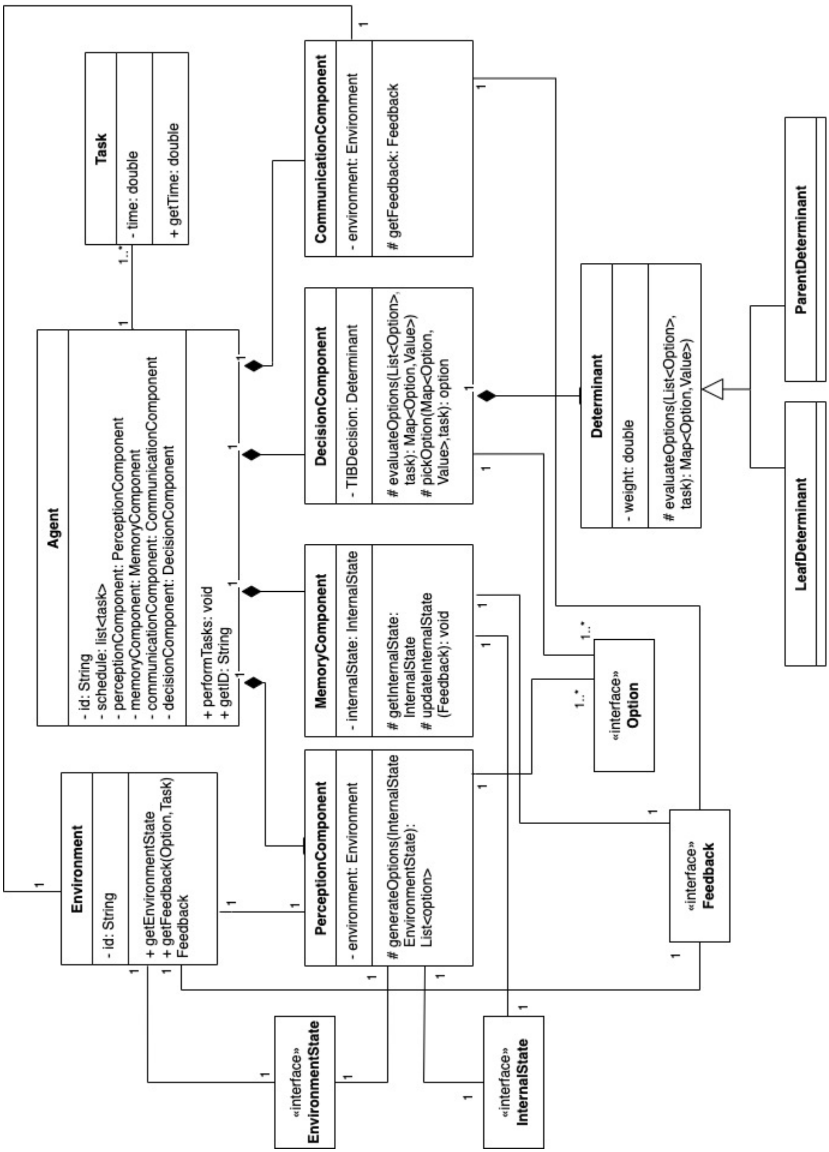


FIGURE 4.9: UML diagram of classes

- Abstract Class **CommunicationComponent** gathers *Feedback* from the *Environment* with regard to the action chosen, which can be specified in the *getFeedback* function by users.

4.3.2 General flow of the decision-making process

Using the classes and interfaces above, we provide the description of a decision-making cycle of an agent as follows:

- When the simulation starts, agents are assigned a number of **Tasks** that have different start times and a list of possible actions.
- At the time of triggering, agents use the **PerceptionComponent** to *filterOptions*, which in turn utilizes *getEnvionmentalState* function from **Environment** and *getInternalState* function from **MemoryComponent**.
- The list of possible **Options** (and their properties) is passed to **DecisionComponent** to be evaluated (given a double value). It is performed by *evaluateOption* function, which calls the decision tree recursively (default TIBDecision) to give a value to each Option depending on its property and the Task. Users often only need to define utility functions *LeafDeterminant* (1st level of TIB tree).
- The **DecisionComponent** selects the most suitable *Option* in terms of 1) the best evaluated (if utility represents the gained benefit) / lowest evaluated option (if utility represents the cost), i.e. deterministic mode, 2) an option based on the percentage composition of all the options' utility values, i.e. stochastic mode.
- The selected **Option** is passed to the **CommunicationComponent** to be communicated to the **Environment**, which will produce a **Feedback**. It is passed to **MemoryComponent** to be used in the function *updateInternalState*.

4.3.3 Pseudo code

This section describes the Pseudo code for the agent's main agent class and its core components. The functions in *Perception, Memory and Communication* components depend on the modelling context and have to be defined in the glue code, so we leave them as abstract functions.

Agent

This class defines all the general components and processes of an agent (see Listing 4.1).

Listing 4.1: Pseudocode of Agent

```
class Agent:
  DEFINE id AS String
  DEFINE schedule AS Task[]
```

```
4   DEFINE perceptionComponent AS PerceptionComponent
5   DEFINE memoryComponent AS MemoryComponent
6   DEFINE decisionComponent AS DecisionComponent
7   DEFINE decisionComponent AS DecisionComponent
8   DEFINE communicationComponent AS CommunicationComponent
9   DEFINE task AS TASK
10  FOR EACH task IN schedule
11    environmentState <- perceptionComponent.getEnvironmentalState()
12    internalState <- memoryComponent.getInternalState()
13    options <- perceptionComponent.filterOptions(environmentState,
          internalState)
14    evaluatedOptions <- decisionComponent.evaluateOptions(options,task)
15    pickedOption <- decisionComponent.pickOption(evaluatedOptions,task)
16    feedback <- communicationComponent.getFeedback(pickedOption,task)
17    memoryComponent.updateInternalState(feedback)
18  END FOR
```

- **Line 1**: Initialise the agent's ID.
- **Line 2-8**: Initialise all agent components.
- **Line 11**: Call *Perception* component to get the current environmental state (including other agent's opinion(s)).
- **Line 12**: Call *Memory* component to get the agent's internal state.
- **Line 13**: The internal state then passed to *Perception* component to derive a set of possible options.
- **Line 14**: This options list is then evaluated by *Decision* component and sorted into a list.
- **Line 15**: The *Decision* component either pick the best evaluated option (deterministic) or turns utility into percentages for agent's selection (probabilistic). It then communicates this information to the environment.
- **Line 16**: The *Perception* component captures feedback(s) of the environment to the agent's action.
- **Line 17**: The *Memory* component stores this feedback for future loop.

Perception

The Perception component initialises the functions that observe the environmental state, combines it with the internal state to produce a set of possible options and captures the feedback(s) from the agent's actions (see Listing 4.2).

Listing 4.2: Pseudocode of Perception component

```
class PerceptionComponent:
  DEFINE environment AS Environment

```

```
  DEFINE function getEnvironmentalState()
  DEFINE function filterOptions(EnvironmentState,InternalState)
  DEFINE function getFeedback(Option,Feedback)
```

Memory

The Memory component stores the internal state of the agent and is responsible for updating it (see Listing 4.3).

Listing 4.3: Pseudocode of Memory component

```
class MemoryComponent
  DEFINE internalState AS InternalState
  DEFINE function getInternalState()
  DEFINE function updateInternalState(Feedback)
```

Communication

As the name suggested, the *Communication* component outputs the agent's decision to the environment and gathers feedback information (see Listing 4.4).

Listing 4.4: Source code of Communication component

```
class CommunicationComponent
  DEFINE environment AS Environment
  DEFINE function getFeedback(Option,Task)
```

Decision Component

The *Decision* component is responsible for evaluating all options in the provided set (see Listing 4.5). It outputs a map of options to their utility. It can then either choose the option which has the highest/lowest utility outcome (deterministic) or turn the utilities into probabilistic functions.

Listing 4.5: Source code of Decision component

```
class DecisionComponent
  DEFINE TIBDecision AS Determinant
  DEFINE function evaluateOptions(Option[],Task)
  DEFINE function pickOption(Pairs(Option, double)[],Task)
```

In this component, the *TIBDecision* variable is the root node of TIB tree structure. Next, we implement the two classes: Leaf determinant (no child node) and Parent determinant (with at least one child node). Each determinant evaluates all options and gives utility values to them.

The *LeafDeterminant* requires user to define an *evaluateOptions* function to assign a utility value to a list of provided *Option*s (see Listing 4.6).

Listing 4.6: Source code of Leaf Determinant

```
class LeafDeterminant
  DEFINE weight AS Double
  DEFINE function Pair(Option,Double)[] evaluateOptions(Option[], Task)
```

The *ParentDeterminant* essentially follows the Equation 4.7 to provide a map between *Option* and its utility value (see Listing 4.7). We describe the code for *evaluateOptions* function as follow:

Listing 4.7: Source code of Parent Determinants

```
1  class ParentDeterminant
2    DEFINE weight as Double
3    DEFINE children AS Determinant[]
4
5    DEFINE function evaluateOptions(options,task):
6      DEFINE results AS Pair(Option,Double)[]
7      DEFINE child AS Determinant
8      FOR EACH child IN children
9        sumValue <- 0
10       DEFINE childEvaluation AS Pair(Option,Double)[]
11       childEvaluation <- child.evaluateOptions(options,task)
12       childrenKeySet <- childEvaluation.keySet()
13       FOR EACH option AS Option IN childrenKeySet
14         sumValue <- sumValue + childEvaluation.getValue(option)
15       END FOR
16       IF sumValue = 0 THEN
17         sumValue <- 1
18       END IF
19       DEFINE option AS Option
20       FOR EACH option IN options
21         childValue <- childEvaluation.getValue(option) * child.getWeight()
                / sumValue
22         IF results.containsKey(opt)
23           results.put(option, results.getValue(option) + childValue)
24         ELSE IF
25           results.put(opt, childValue);
26         END IF
27       END FOR
28     END FOR
29     RETURN results
30   END FUNCTION
```

- **Line 6**: Define the *results* variable as a map between an *Option* and its utility value.
- **Line 7**: Define the *child* variable as a *Determinant*.
- **Line 8**: The loop goes through each child in the list.
- **Line 9**: Initialise the *sumValue* variable, which is used to calculate the sum of all utility of the determinant's children node.
- **Line 10**: Define the *childEvaluation* variable as a map between an *Option* and its utility value.

- **Line 11**: Assign the *childEvaluation* variable to the result of a recursive call to the *evaluateOptions* function in that child.
- **Line 12**: Assign the *childrenKeySet* variable as the key set of the results of the previous call.
- **Line 13-15**: Assign the *sumValue* as the sum utility values of all options in that particular child.
- **Line 16-18**: If *sumValue* is 0, assign it to 1 to avoid division to 0.
- **Line 19**: Define the *option* variable as an *Option*.
- **Line 20**: The loop goes through each option in the input list.
- **Line 21**: Assign the *childValue* variable as utility value of an *Option* in term of the particular child that we are considering. This process is corresponding to the inner part of the summarisation in Equation 4.7.
- **Line 22-26**: If a value corresponding to that option already exist in our *results* map, we increase it by the amount of the *childValue* variable. Otherwise, we simply put the value of *childValue* variable in the *results* map.

4.4 How the framework is used

The framework can be utilised to simulate a decision-making process. To use it, users have to define the parameters of the agents and the environment as well as their interactions. In terms of programming, the modeller can implement the glue code for the aforementioned abstract classes and interface.

1. Specify the targeted behaviours or the interesting features/phenomena that can be simulated through the given modelling context.
2. Parametrise agents and environment and their attributes using the available data. This way, all the fields and variables of the main *Environment* and *Agent* classes should be defined.
3. Design how the agents and environment interact, including *Agent-Self*, *Environment-Self*, *Agent-Agent*, *Environment-Environment*, *Agent-Environment* (see Section 2.3). In particular, the modeller can define the variable fields in the *EnvironmentState* and *Feedback* interfaces.
4. Define how an agent filter and evaluate an option. This process involves the implementation of the *InternalState* and *Option* interfaces. Users also need to extend the main components of the agents, i.e. *PerceptionComponent*, *MemoryComponent*, *DecisionComponent* and *CommunicationComponent*, and specify the functions in these classes. In addition, the function *evaluateOptions* in the *LeafDeterminant* should be defined. The default of the computation is set as TIB

three-layered model, but a user can create their own decision-making structure that is more suitable for the context and available database.

5. Calibrate the agent parameters so that the outputs reflect empirical/historical data.

6. Set up an experiment to demonstrate or test the model's functionality base on the type of decision-making and the model's purpose. Its results can then be interpreted and discussed.

It should be noted that the steps above are suitable for a situation where the model is developed after data collection. Therefore, the modellers cannot influence the data set's content, so they have to adapt the agent and environment's parameters and their behaviours accordingly. This situation applies to all of our case studies in the next four chapters 5 to 8.

In an alternative case, the modellers can incorporate interesting behaviour features into the planning of the data collection process. It implies that correct and enough contextual data is available to create a model. In this case, we suggest that step 2 can be moved back behind steps 3 and 4.

4.5 Documentation

In this section, we list the relevant documentation process for the framework to help the readers to be able to grasp the core aspects and be able to replicate an example model of the framework if they want to understand how it works. Often, there is not a space within the length of a conventional article to describe a simulation sufficiently to enable replication to be carried out. A more radical solution is to publish the code in a public repository. To provide this implementation, we first consider suitable programming languages and platforms. A majority of them can be found in surveys, such as [CF21; KB15; Das14]. We list two of the most popular tools in the education and the ABM research community below:

- NetLogo[8] [TW04] is a multi-agent programmable modelling environment, which allows quick prototyping but sophisticated simulations. It comes with an extensive models library, including models in various domains, such as economics, physics, chemistry, biology and psychology and system dynamics. NetLogo can be used by teachers in the education community and domain experts without a programming background to model related phenomena. Beyond exploration, it also allows the authoring of new models and modifying existing models. However, its features are relatively simple and cannot be used for large-scale simulations.

- The Repast Suite[9] [Nor+13] is a family of advanced, free, cross-platform, agent-based modelling and simulation toolkits. It provides multiple implementations

[8]Main website: `https://ccl.northwestern.edu/netlogo`

[9]Main website: `https://repast.github.io`

in several languages and many built-in adaptive features. Currently, there are two editions of Repast and several ways to write models in each edition to satisfy many different kinds of users and cases. Compared to NetLogo, Repast is potentially more flexible since users have access to the whole range of Java libraries and other major languages, but it requires programming background. RePast is also actively updated for newer Java versions and functionalities.

As our team has experts in the Java language and aims to build a scalable model, we implement our simulation, Behaviour-Driven Demand Model (BedDeM), using RePast. Its source code with mobility example can be seen on GitHub[Git]: `https://github.com/SiLab-group/beddem_simulator`.

Another documentation methods is through publication, adding our findings to the stock of scientific knowledge. There are an ongoing number of articles and abstracts that results from the usage of our framework, including [NS19b; BS19a; BS19b; NS19a; NS20b; NS20a; BPS21; NPS22]. They mainly follow the *Overview, Design concepts and Details* (ODD) document protocol to describe the implemented agent-based model [GPT17], which includes three main categories:

- The *Overview* section creates an outline of the model.
- The *Design Concepts* section describes the general concepts underlying its design.
- The *Details* section provide all information that are needed to re-implement the model (e.g. initialisation, input, calibration).

This thesis is also a form of system documentation, which provides the following formal language descriptions: 1) ontologies and graph that formally describe components and their relationships, 2) pseudo-code combining natural language with programming syntax and 3) mathematical descriptions in the form of mathematical equations.

In addition, the next four chapters 5 to 8 will show that our framework is suitable for different research fields (i.e. domain-independent) and contexts. Our implementation platform - BedDeM - is being developed in Java using the Repast library for agent-based modelling [Nor+13]. We provide four case studies with different focuses: 1) modelling mobility demand, 2) trust and reputation for travelling by train, 3) bounded rationalities in purchasing new vehicles and 4) activities of migrants during a pandemic. Each model focuses on different contexts, determinants and types of decision-making, such as optimal/bounded rationality behaviours and short-term/long-term decisions. The properties of the environment mentioned in Section 2.2 are also considered. The diversity of these case studies tests the flexibility of our framework and its ability to highlight different social phenomena or effects that are not well captured in previous modelling efforts.

In general, each case study follows the below structure adapted from the steps presented in Section 4.4:

1. **Introduction and description of case study**: This subsection gives the information background, the complete scenario description and the potential insights provided by the models and experiments in the case study. It also identifies the type of decision-making (e.g. short-term/long-term, optimal/bounded optimal) and properties of the environment (e.g. fully/partially observable, deterministic/stochastic, static/dynamic, know/unknown). (Step 1 of Section 4.4)

2. **Related work**: This subsection summarises the models and other implementations that have been developed for a similar context, including the applications of the related work in Chapter 3.

3. **Dataset and parametrisation of the environments and agents**: This subsection describes the available data set and how we incorporate it into the model's parameters. We then detail the procedure in which we extend the main *Environment* and *Agent* classes and implement relevant interfaces (e.g. *InternalState*, *Option*) to reflect the given description. (Step 2 of Section 4.4)

4. **Interactions between the environment and agents**: This subsection specifies which information an agent can perceive from the environment and how it communicates its decisions to the environment and other agents. We also summarise how the *EnvironmentState* and *Feedback* interfaces are implemented. (Step 3 in Section 4.4)

5. **Agent's decision-making process**: This subsection describes the information flow between the core components (i.e. the *PerceptionComponent*, *MemoryComponent*, *DecisionComponent*, *CommunicationComponent* classes) to determine the best option from the observed list, in particular. It also provides the details of implementing of the function *evaluateOptions* in the *LeafDeterminant*. (Step 4 of Section 4.4)

6. **Calibration**: This subsection outlines the process of finding a set of agent parameters that is most compatible with the historical data. (Step 5 of Section 4.4)

7. **Experiments and results**: This subsection details experiments highlighting the usage of a socio-psychological decision-making platform. (Step 6 of Section 4.4)

8. **Advantages and limitations of using our framework**: This subsection outlines the advantages as well as disadvantages of our framework compared to the state-of-the-art mentioned in Chapter 3.

Some case studies include the extended components of the previous case. Hence, some sections are omitted. Furthermore, the first and second studies are modified versions of our publications. The third case includes the components of a submitted paper. In these cases, we will specify the publication accordingly.

Chapter 5

Case study: The model of mobility demand

In the first case study (i.e. Chapter 5), agents make decisions on transportation mode as needed for their routine, i.e. a form of *short-term* behaviours. In this case, the environment is *deterministic*, *fully observable*, *static* and *known*. Experiments in this section are based on *optimal behaviours*. In other words, the agents select the best option after performing the evaluation process using determinants in TIB. This chapter contains the modified version of some elements from our publication [NS19b; NS19a], including the related work, BedDeM's design, calibration and the experiment.

5.1 Introduction and description of case study

Our framework's first application is in the mobility domain, which aims to generate yearly demands at the individual household level that can be interpreted at the granularity of historical *evolution of transportation* for Switzerland's private households. In this case study, we demonstrate the procedure of adopting our framework in the context of mobility modal choice and the available data set. The simulation is performed to test our framework capability to capture the effects of different aspects or strategies mentioned in the example of Section 1.1 (e.g. economic attitude, emotion, habit) at the micro level. The experiment in this study involves changing the weights of second and third-level determinants in TIB. Comparing the results with a calibrated scenario provides a sensitive analysis of the associated weights for different transportation modes.

After considering some of the related work in simulating mobility-related decision-making, we describe the statistical data and technical details of the environment and agent population. The following sections detail their interactions and the agent's decision-making process. Next, an experiment demonstrates how to use the framework to capture the effect of individual decision-making determinants on the emerged collective phenomena. Finally, we identify the advantages and limitations of our framework, especially compared to those mentioned in Chapter 3.

5.2 Related work

In this section, we provide related work on models that deal with mobility modal choice. In this case, the agent's objective is to select the most appropriate option from a set of alternatives. From the decision-making process, the utilities or probabilities of all options can be derived from different methods.

One approach incorporates statistical discrete choice models with the agent's preferences, strategies and likelihood of choosing a particular action [BAB99]. Many projects in the mobility domain use *random parameters logit* [HG03] as a way to assign predicted probabilities to outcomes of an action. Examples include [CBA08; Anw+14]. Empirical data (e.g. observed choices, surveys, hypothetical scenarios, administrative records) can then be used to estimate the parameters of the agent's preferences. However, this approach does not allow users to distinguish the different socio-psychological aspects of decision-making (e.g. economic attitudes, habits, emotion) and their impact on the agent's behaviours.

To produce a more elaborate decision-making process, another agent class is assigned with beliefs, values or world views corresponding to observation from demographic data or stakeholder assessments. It is often implemented using the Belief-Desires-Intentions (BDI) [RG91] architecture. Practical examples include the work of Padgham et al. [Pad+14], Bazzan et al. [BWK99], and Balmer et al. [Bal+04]. However, this agent class is often criticised for the lack of experimental grounding [DMJ06] and the agent choice of being homogeneous, selfish and focused only on economic drivers [RG91]. To the best of our knowledge, there is no project in the mobility domain that utilises more complex cognitive architectures, such as ACT-R, CLARION [Sun06], [TLA06] or SOAR [Lai12]. Regardless, as mentioned in Chapter 3, they do not have the components to cover all major aspects of human decision-making (i.e. cognitive, affective, social, norm and learning) [BG14].

5.3 Parametrisation of the environment and agents

In this case study, the Transport Microcensus (MTMC) [Mic] and the Swiss Household Energy Demand Survey (SHEDS) [Web+17] are utilised. In this section, we describe their characteristics and the process of capturing them in the environment and agent's parameters.

5.3.1 Dataset

The MTMC is a statistical survey of the travel behaviour of the Swiss population conducted every five years by the Federal Office for Spatial Development (ARE) and the Federal Statistical Office (FSO). Most recent data were published in 2015, containing information about:

- The socioeconomic characteristics of households and individuals' mobility tools

- Daily mobility and occasional journeys
- Attitudes towards transport policy in Switzerland.

The SHEDS is designed to collect a comprehensive description of the Swiss households' energyrelated behaviours, their longitudinal changes and the existing potential for future energy demand reduction. The survey was planned in five annual waves (2016-2020), thus generating a rolling panel dataset of 5,000 respondents per wave. The household characteristics collected can be classified into three main disciplines:

- Economics: number of people per household, average age, income, education, type of residence, energy literacy, etc.
- Psychology: environment attitudes, risk attitudes, emotions, values, life satisfaction, etc.
- Sociology: abstract of context and performance, life events, etc.

We are particularly interested in questions that compared the criteria for mobility mode choices, whose answers can be interpreted as the weights(w_i) for different psychological determinants in TIB. A typical example can be observed in Figure 5.1. We can interpret them as follow:

Please rate how important the following aspects are for choosing this mode of transportation.

	Not important 1 (1)	2 (2)	3 (3)	4 (4)	Extremely important 5 (5)
Choosing the cheapest option [_**1**]	❍	❍	❍	❍	❍
I am used to taking this means of transport [_**2**]	❍	❍	❍	❍	❍
Travelling as safely as possible [_**3**]	❍	❍	❍	❍	❍
Travelling as fast as possible [_**4**]	❍	❍	❍	❍	❍

FIGURE 5.1: An example question from the Swiss Household Energy Demand Survey, taken from [She]

1. Choosing the cheapest option - Weight for *cost*.
2. I am used to taking this means of transport - Weight for *habit*
3. Travelling as safely as possible - Weight for *emotion*
4. Travelling as fast as possible - Weight for *time*

5.3.2 Environment parametrisation

The environment includes 26 entities which represent different *Cantons*, which are the major administrative divisions in Switzerland. This study will refer to them as *regions*. Each region has a list of mobility agents presenting its population. It also contains the list of available public transport modes for the region, e.g. bus, tram, train. In addition, they provide a way to monitor and report the activities of different geographic groups of agents for the results. The environment here is considered to be *deterministic*, *fully observable*, *static* and *known*.

Using the classes and interface provided in Section 4.3, the implementation of the **Environment** interface and **Mode** can be seen in Figure 5.2. Each region has the following properties:

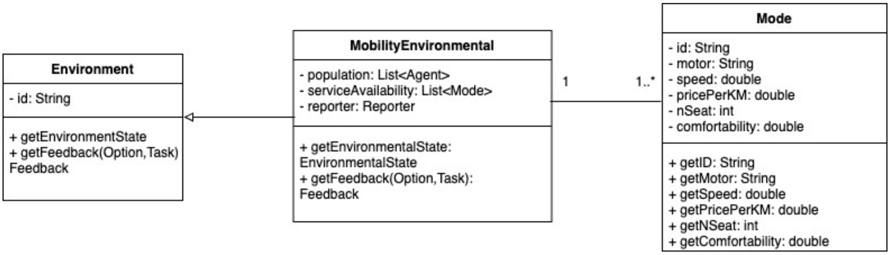


FIGURE 5.2: The extension of the Environmental and Mode class

- Population: list of agents who are living in this particular region.
- Service availability: list of public transport (i.e. bus, tram, train) with their availability in binary format.

The mode can also be defined with the following fields:

- Motor: the type of engine of the vehicle, e.g. gasoline, diesel, electric, hybrid.
- Speed: the speed of the vehicle in terms of kilometres per hour, which can be used to calculate the duration of the trip.
- Price per kilometre: the price per kilometre of the vehicle, which can be used to calculate the cost of the trip.
- Number of seats: the number of available seats on the vehicle.
- Comfortability: a number assigned by our economic collaborator to compare the luxury class of the vehicle, e.g. sports car (50) versus family car (30).

5.3.3 Agent parametrisation

Figure 5.3 illustrates the process of parametrising the household profiles to build a synthetic population. Interested readers can refer to the full procedure in [BS19a]. In short, MTMC's entries (N = 57,091) are placed in a latent space (socio-matrix) that is represented by a symmetric Gower distance matrix [Gow71]. All pairwise distances/dissimilarities are created based on the common features of the two data sources (e.g. age group, gender, region, household size, income level, number of personal vehicles). This matrix also provides a way to calculate the recommendation for agents from the same network (i.e. $U(Role)$). We then find the most similar peers with the lowest distance to each other and join them with SHEDS entries (N=5,515). A random number of representatives for each geographical region in Switzerland are selected to become our agent population (N=3,080).

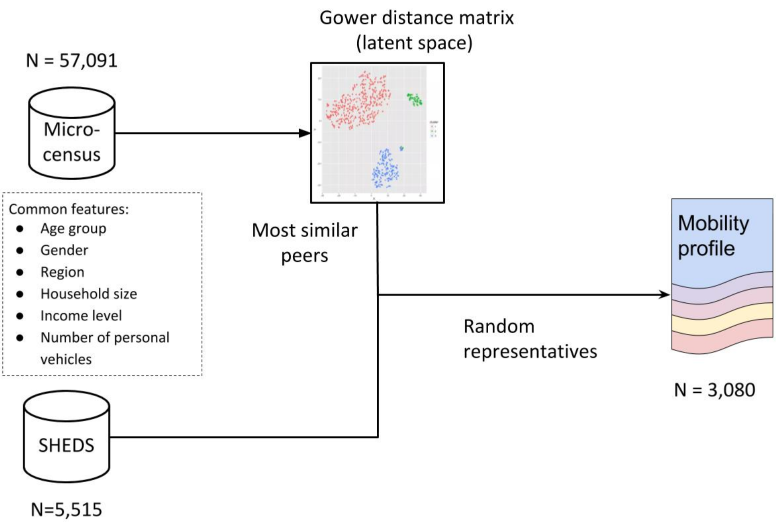


FIGURE 5.3: Building a synthetic population

After the above process, we can define the following components in our agent framework: *Internal state*, *Task* and *Option*. Figure 5.4 is an UML diagram illustrating the implementation with the classes and interface provided in Section 4.3.

- **Agent's internal state**:
  - Location: Region (or *Canton* in Switzerland) where the agent lives.
  - Budget: Weekly travelling budget.
  - Accessibility set: List of available transportation services for the agent. They are allocated according to the MTMC. For each mode, the agent has information about its speed, type of engine, cost per km, comfortability

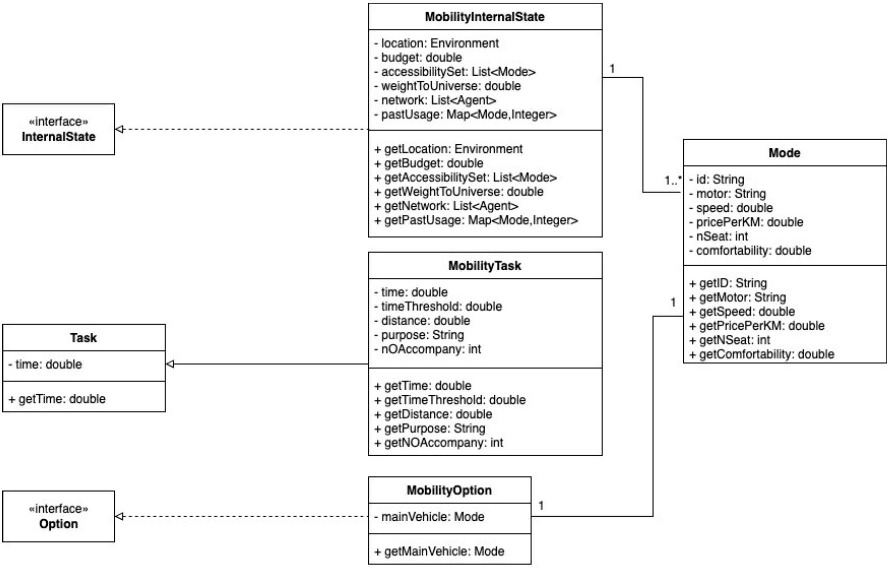


FIGURE 5.4: The implementation of *Mobility internal state*, *Mobility Task*, *Mobility Option*

(assigned according to the vehicle category, e.g. SUV, family car, sports car) and emission.

- Weight to universe: The proportion of the population that the agent represents.
- Network: The list of neighbour agents is created based on the location of the agent profile. The strength of the connection relies on its trust in a different role in society, which is extracted from SHEDS.
- Past usage: The map to record the number of usage times per each mode of transportation.

- **Task**: A weekly schedule is also derived for each agent from MTMC. It includes several trips that the agent has to make. The agent's main purpose is to select a transportation mode to perform a task on its schedule. Each task has the following properties:
  - Start time (in terms of hours and day of the week).
  - Threshold (maximum) time for the trip.
  - Distance between locations.
  - Purpose: go to work, shopping, school, leisure, etc.
  - Number of accompanying people.

- **Option**: A transportation mode (including rail, car, bus, tram, biking, walking and others). There is also an option of not performing the scheduled activity due to the constraints from the agent's internal state or environment (e.g. exhaustion of budget or exceeded travelling time on all available modes).
- **Feedback**: There is no immediate reaction from the environment caused by the agent's choice.

5.4 Interaction between the environment and agents

The mobility model is illustrated in Figure 5.5. The simulation process starts with a central controller creating all the agents with all their attributes and assigning them to their respective regions. Each region contains information about the available public transportation and can reflect the dynamic change in traffic rate concerning the simulating time. The agent then looks at its schedule, creates decision-making events and puts them in the stack to be activated. At the time of simulation, the controller triggers these events simultaneously, waits for them to finish, and then skips to the next scheduled point (i.e. event-driven). After all the tasks are finished, a reporter component in the region collects the final results. Since each agent represents a portion of the total population, these numbers are then multiplied by the weight_to_universe parameter in the agent's internal state to be compatible with MTMC for the calibration process.

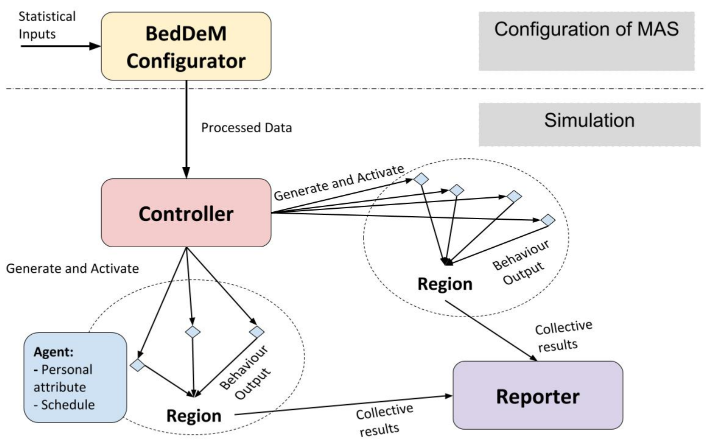


FIGURE 5.5: Overview of BedDeM

In terms of our interfaces provided in Section 4.3, the **EnvironmentState** can be implemented to provide information on whether a public transportation mode (i.e. bus, tram, train) is available at the agent's location (see Figure 5.6). At this stage of development, there is no response that the environment needs to give back to the agent. Hence, the **Feedback** interface is not implemented in this model.

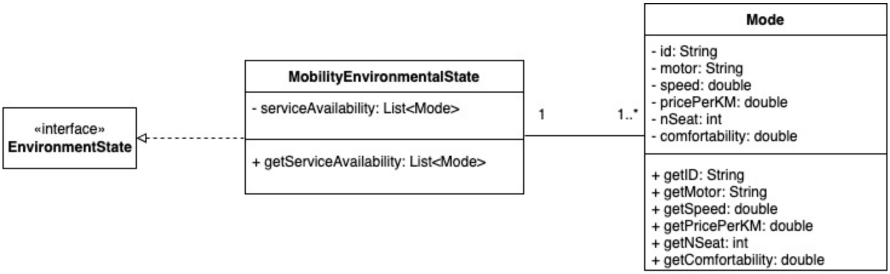


FIGURE 5.6: The implementation of Environmental state interface

5.5 Agent decision-making process

Our decision-making component requires two elements to calculate the expected utility for a set of options: 1) how to specify a ranking order of the option according to a determinant ($U_{opt}(d)$) and 2) the weight of the determinant ($w(d)$). A mapping of the first-level determinants can be seen in Figure 5.7. It is based on some of the past research [Amp04] and the kinds of properties that can be measured or ranked objectively (using common sense). In this case, the determinant *Belief* is omitted since the system assumes that the knowledge/perception of agents is always correct.

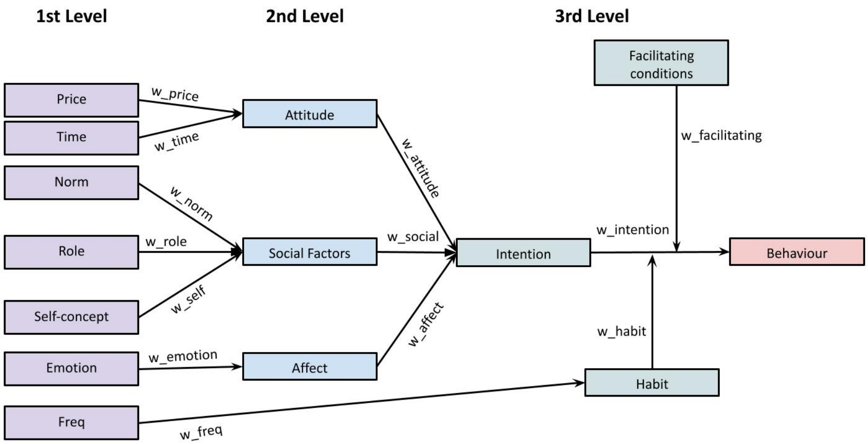


FIGURE 5.7: TIB's determinants mapping for the mobility case study

In terms of our framework implementation, we extend the **LeafDeterminant** to represent the determinants in the first level of TIB (see Figure 5.8). The *evaluateOptions* function can then be implemented with as the expected utility of each option $U_{opt}(d)$, which can be calculated or be given a ranking value related to other option:

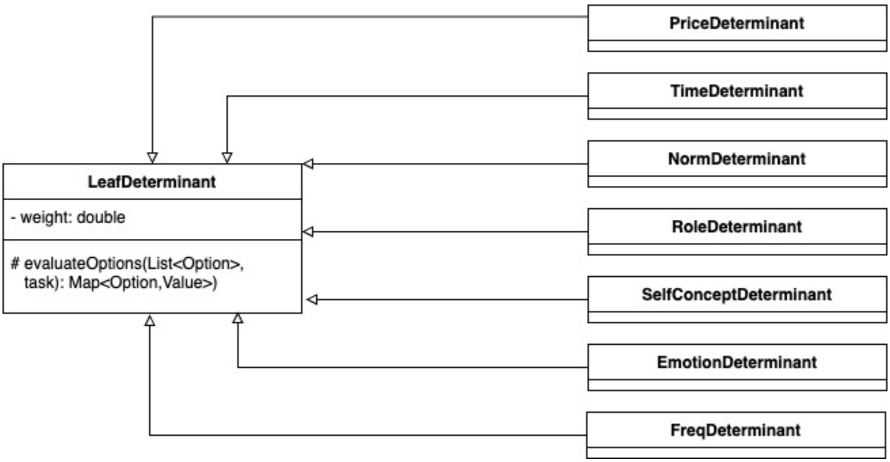


FIGURE 5.8: The implementation of first level

- $U(Price)$ = distance of the trip divided by cost per km of the mode.
- $U(Time)$ = distance of the trip divided by speed of the mode.
- $U(Norm)$ = the number of neighbours in the agent network using the mode.
- $U(Role)$ = ranking of environmental friendliness of the mode: 1 - Walking/Biking, 2 - Electric, 3 - Combust Engine (Gas/Diesel).
- U(Self-concept) has no data available in the data set. We will calibrate it by finding the ranking of the agent's preferred vehicle groups, including walking/biking, car/motorbike, train/tram/train and others (see Section 5.6).
- $U(Emotion)$ = the vehicle luxury / comfortability.
- $U(Frequency)$ = the number of times the agent used the same mode previously.
- $U(Facillitating)$ = whether the destination has convenient parking: value 0 or 1.

The weight of each determinant ($w(d)$) is extracted from the rating provided by a profile in SHEDS (see example in Section 5.3.1), with range from 1-5. The list of questions with matching determinants is as follows:

- $w(Price)$ - Please rate how important the following aspects are for choosing this mode of transportation: Choosing the cheapest option.
- $w(Time)$ - Please rate how important the following aspects are for choosing this mode of transportation: Travelling fast.
- $w(Role)$ - Please rate the extent to which you agree with the following statements: My acquaintances expect that I behave in an environmentally friendly manner.
- $w(Norm)$ - Regarding energy and saving energy, how strongly do you trust the information provided by the following people: Neighbors.
- $w(Self-concept)$ - Please rate how important each value is for you as a guiding principle in your life: Self-indulgent.
- $w(Emotion)$ - Please rate how important the following aspects are for choosing this mode of transportation: I enjoy this way of travelling.
- $w(Freq)$ - Please rate how important the following aspects are for choosing this mode of transportation: I am used to taking this means of transport.
- $w(Attitude)$ - Please rate how important each value is for you as a guiding principle in your life: Wealth.
- $w(Habit)$ - Please rate how important each value is for you as a guiding principle in your life: Social power.
- $w(Intention)$ = MAX_SCALE - $w(Habit)$.

After all the classes are defined, the framework can calculate the final U_{opt}(BehaviouralOuput) using the code and structure design provided in Section 4.3.

5.6 Calibration

The purpose of calibration is to improve the current population's compatibility with the target system (i.e. statistical data from MTMC). We focus on figuring out the most fitted ranking patterns for $U(Self-concept)$. We divided the agent population into four profiles, depending on their daily main transportations according to the SHEDS: 1) soft-mobility modes (walking/biking), 2) public vehicles (tram/bus/train), 3) private vehicles (car/motorbike) and 4) others. $U(Self-concept)$ for each of them can then be calibrated by permuting the ranking order of all the modal choices.

Our main objective is to minimise the error calculated from the Equation 5.1. It is measured from the total differences between the final sum of kilometres in each mobility mode at the end of a period (i.e. a year in this case) and historical data. The total kilometres result for one year of all mobility profiles can be obtained (i.e. walking/biking, bus/tram/train, car/motorbike, others) from MTMC [Mic]. Calibration involves using the permutation of these four sets of modes as configurations for the $U(Self - concept)$. We repeat this procedure for all agent profiles set at either deterministic (choose the best option) or stochastic (choose from a random function with probabilities provided by sampling distribution of final referenced values) to find the smallest error.

$$\underset{conf}{\text{minimise}} \quad err(conf) = \sum_{m \in \mathbb{M}} \mid census_m - sim_m(conf) \mid \tag{5.1}$$

where $\mathbb{M}$ is the set of walking/biking, bus/tram/train, car/motorbike and other. $census_m$ is census data for mode m (in kilometres). $sim_m(conf)$ is the simulation result for mode m (in kilometres). $conf$ is a ranking of four categories in $\mathbb{M}$ and can be captured in the form of permutations with repetition. Examples are:

- (1) walking/biking, (1) bus/tram/train, (1) car/motorbike, (1) others.
- (1) walking/biking, (1) bus/tram/train, (1) car/motorbike, (2) others.
- (1) walking/biking, (1) bus/tram/train, (2) car/motorbike, (1) others.
- (1) walking/biking, (2) bus/tram/train, (1) car/motorbike, (1) others.
- ...

We list the kilometres in census data and the top results of two types of agents in Section C.1. The best configuration is in the deterministic model (i.e. agents pick the option with the highest utilities), which has an error accounting for 6.5% of the total scheduled kilometres. The main differences are in the *public* (i.e. bus/tram/-train) transportation numbers. We also observe that the stochastic error is much more significant (above 51.8x10^9 kilometres) with an accuracy of 46%. This result is expected since agents in stochastic mode choose options based on a random function of probabilities derived from the utility values. Currently, no pattern is shown in the ranking function $U(Self - concept)$ of the results of *stochastic* mode, and hence additional runs with different distribution functions are needed to have a broader picture for this setting. Nevertheless, we decide to use the deterministic mode for the experiment below because it gives a more unambiguous indication of which results change with the modification in the agent's parameters.

5.7 Experiment with behavioural determinants in mobility modal choices

Our agent's decision-making platform offers a mechanism to measure the impact of different individual determinants on short-term transportation modal choices (i.e. car, bus, tram, trains, walking, biking). We demonstrate this capability through a series of setups to activate/deactivate the second and third level determinants of TIB in the agent's decision-making and compare the collective results after simulation. The current agent population contains a mapping of qualitative data in SHEDS to all

TIB's determinants, which is designed to reproduce the travelling patterns in MTMC. Hence, performing the experiment on this baseline can provide a practical insight into real-life situations where people often rely on a small set of factors and aspects to make their decision on modes of daily transport. These factors and aspects correspond to the ones mentioned in Section 1.1, including economic attitude, affective, social and habit. By singling out their effects, we also aim to identify the main drivers of different mobility mode choices in the agent population.

5.7.1 Design

In this experiment, we want to focus on observing the impact of second and third-level determinants in TIB, i.e. *Attitude, Social factors, Affect, Facilitating condition, Intention* and *habit*. It can be achieved by qualifying the mobility demand changes for different transport categories, including car, bus/tram, train, walking, biking, and other modes. In each scenario, we modify the corresponding weights in the agent's decision-making, i.e. $w(Attitude)$, $w(Social)$, $w(Affect)$, $w(Facilitating)$, $w(Intention)$, $w(Habit)$ (see Figure 5.7 and Table 5.1). This exercise is performed on the calibrated deterministic population described in Section 5.6; in which mode agents choose the best alternative for their trips. By keeping the weight(s) of the main determinant(s) as calibrated values and others to 0, the agent will only consider that key determinant(s) in decision-making and ignore the rest. In the first half of this setup, we focus on the second level of TIB, which connects to *Intention* in the third level. Hence, $w(Intention)$ is kept as in Section 5.6. This process is also applied similarly to $w(Attitude)$, $w(Social)$ and $w(Affect)$ in the second part to ensure $U(Intention)$ remains non-zero. All trips are scheduled within one year, so there is no difference in agents' accessibility to modes. Seasonal changes are planned for a future developing stage.

Table 5.2 shows a running example in the mobility domain which follows the TIB determinants mapping in Figure 4.6. An agent needs to make a working trip from Sion to Sierre and has access to three options: using a *car*, taking a *train* or a *bike*, which is assigned from data collected from [Mic]. In addition, U is a cost function, i.e. option with a smaller value is preferred. As explained above, the agent believes and evaluates the consequence/cost of this journey based on two criteria, namely *Price* and *Time*. It expects the *car* option to have the price of around 4 Swiss Franc, and so $U_{car}(Price) = 4$. Correspondingly, $U_{train}(Price) = 3$ and $U_{bike}(Price) = 0$. Their total value, $\sum U_(Price)$, is 7. The estimations for time are $U_{car}(Time) \approx 0.3$, $U_{train}(Time) \approx 0.2$ and $U_{bike}(Price) \approx 1$; the sum of which is 1.5. According to [Web+17], the agent has $w(Price)$ and $w(Time)$ are 2 and 4 respectively.

The utility values of car, train and bike at determinant *Attitude* are U_{car} (Attitude) ≈ 1.94, $U_{train}(Attitude) \approx 1.39$, $U_{bike}(Attitude) \approx 2.67$ respectively. They add up to 6 in total. Similarly, the sum of U for *Social factors* is 8 and sum of U for *Affect* is 1. With the *Attitude (At)* setup in Table 5.1, their weights ($w(Socialfactors)$ and $w(Affect)$) are 0. By applying Equation 4.7, the new expected value in the next level ($U(Intention)$) of *car* would be = 1.94/6*4 + 1.39/8*0 ≈ 1.3, *train* would be 1.39/6*4 + 2.67/8*0 + 0.33/1*0 ≈ 0.93, and *bike* would be 2.67/6*4 + 3.33/8*0 + 0.5/1*0 ≈ 1.78. In the third level, *Habit* and *Facilitating conditions* have their weights both equal to

TABLE 5.1: Experiment design

Main determinant(s)	$w(Attitude)$	$w(Social)$	$w(Affect)$	$w(Facilitating)$	$w(Intention)$	$w(Habit)$
Attitude (At)	as calibrated	0	0	0	as calibrated	0
Social Factors (SC)	0	as calibrated	0	0	as calibrated	0
Affect (Af)	0	0	as calibrated	0	as calibrated	0
At + SF	as calibrated	as calibrated	0	0	as calibrated	0
SC + Af	0	as calibrated	as calibrated	0	as calibrated	0
St + Af	as calibrated	0	as calibrated	0	as calibrated	0
Facilitating Conditions (FC)	as calibrated	as calibrated	as calibrated	as calibrated	0	0
Intention (I)	as calibrated	as calibrated	as calibrated	0	as calibrated	0
Habit (H)	as calibrated	as calibrated	as calibrated	0	0	as calibrated
FC + I	as calibrated	as calibrated	as calibrated	as calibrated	as calibrated	0
I + H	as calibrated	as calibrated	as calibrated	0	as calibrated	as calibrated
FC + H	as calibrated	as calibrated	as calibrated	as calibrated	0	as calibrated

0 and so, their utilities are not taken into account in the behaviour output. In other words, $U_{car} \approx 1.3$, $U_{train} \approx 0.93$, $U_{bike} \approx 1.78$. In contrast with the running example in Section 4.2.4, these utilities indicate that a *train* would now be the preferred option for this agent. By switching others' weights off (or assigning them to 0), this scenario highlights the connection of the agent's attitude toward using a train.

TABLE 5.2: Running example of an agent's decision-making adapted from Table 4.2 with *attitute (At)* setup.

Level	Determinant	w	EU
1st	Evaluation (Price - Swiss franc), Belief = 100%	2	$U_{car} = 4$ $U_{train} = 3$ $U_{bike} = 0$
	Evaluation (Duration - hours), Belief = 100%	4	$U_{car} \approx 0.3$ $U_{train} \approx 0.2$ $U_{bike} \approx 1$
	Norm (similarity with others)	3	$U_{train} = 1$ $U_{car} = 2$ $U_{bike} = 3$
	Role (environmental friendliness)	2	$U_{car} = 3$ $U_{train} = 2$ $U_{bike} = 1$
	Self-concept (personal preference)	3	$U_{car} = 1$ $U_{train} = 2$ $U_{bike} = 3$
	Emotion (enjoyment)	1	$U_{car} = 1$ $U_{train} = 2$ $U_{bike} = 3$
	Frequency (past similar trips - note that lower value means more usage)	3	$U_{car} = 0$ $U_{train} = 0$ $U_{bike} = 1$
2nd	Attitude (Evaluation + Belief)	4	$U_{car} \approx 1.94$ $U_{train} \approx 1.39$ $U_{bike} \approx 2.67$

Continued on next page

Table 5.2 – continued from previous page

Level	Determinant	w	EU
	Social factors (Norm + Role + Self-concept)	0	$U_{car} = 2$ $U_{train} \approx 2.67$ $U_{bike} \approx 3.33$
	Affects (Emotion)	0	$U_{car} \approx 0.17$ $U_{train} \approx 0.33$ $U_{bike} = 0.5$
3rd	Intention (Attitude + Social factors + Affect)	4	$U_{car} = 1.94/6{*}4 + 2/8{*}0 + 0.17/1{*}0 \approx 1.30$ $U_{train} = 1.39/6{*}4 + 2.67/8{*}0 + 0.33/1{*}0 \approx 0.93$ $U_{bike} = 2.67/6{*}4 + 3.33/8{*}0 + 0.5/1{*}0 \approx 1.78$
	Habit (Frequency)	0	$U_{car} = 0/1{*}0 = 0$ $U_{train} = 0/1{*}0 = 0$ $U_{bike} = 1/1{*}0 = 0$
	Facilitating conditions	0	$U_{car} = 0$ $U_{train} = 0$ $U_{bike} = 0$
	Behaviour output		$U_{car} = 1.3/4.{*}4 \approx 1.30$ $U_{train} = 0.93/4{*}4 \approx 0.93$ $U_{bike} = 1.78/4{*}4 \approx 1.78$

5.7.2 Results

After the simulation, all mobility modes' total kilometre results can be obtained (i.e. walking, biking, bus/tram, train, car, others). Comparing reference results in Section 5.6 against the outcomes of each setup above would give us an idea about the impact of the main determinants. The mapping of TIB's determinants and percentage composition of the modes can then be used to interpret the meaning of the difference in each test.

- **Attitudinal, Affective and Social determinants:** Table 5.3 and Figure 5.9 show the results of running BedDeM with the reference population and with one or two determinants of the second level turned on.

  In the *Attitude(At)* test case, a large number of car users switched to other options, i.e. bus/tram, train and walking. From Figure 5.7, this determinant consists of two elements - monetary price and time. Public modes and soft mobility (e.g. walking, biking) offer competitive prices in the current market. Their current speeds are, however, worst. Having a closer look into the agents who switch, we observe that the main reason is the higher values in their weights for cost, i.e. $w(Price) > w(Time)$. As the utility of *Attitude* is $U(Attitude) = U(Price) * w(Price) + U(Time) * w(Time)$, the weights, in this case, create a disadvantage for cars.

TABLE 5.3: Result of comparing the second level of TIB's determinants (All units are in 10^9 kilometres)

Main determinant	**Car**	**Bus / Tram**	**Train**	**Walk-ing**	**Bik-ing**	**Oth-ers**
Reference population	73.09	4.07	23.2	2.67	4.91	4.42
Attitude (At)	45.77	16.0	33.86	6.22	5.9	4.58
Social Factors (SF)	40.57	17.34	45.1	2.45	1.85	5.03
Affect (Af)	82.32	1.51	15.55	2.32	6.37	4.29
At + SF	37.97	16.9	47.22	2.88	2.22	5.16
SF + Af	69.44	3.38	27.19	2.81	5.12	4.42
At + Af	77.84	3.45	17.95	2.96	5.87	4.29

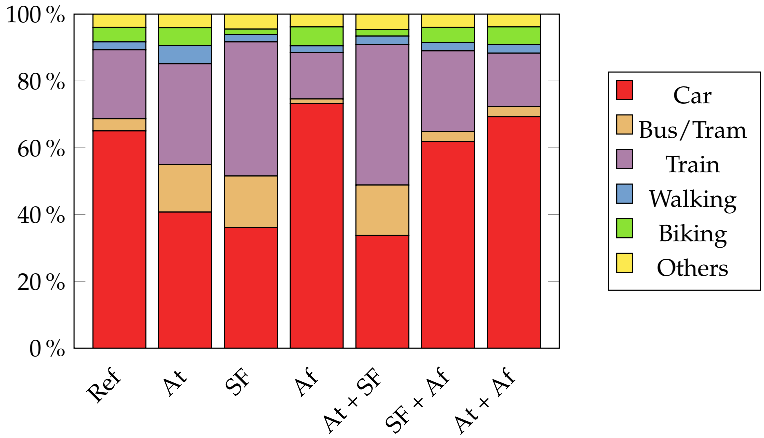


FIGURE 5.9: Percent composition of modes in the tests of second level of TIB's determinants

Similar shifts can also be seen in the *Social factors (SF)* test case with a more than 40% decrease in car usage. As they provide a place for socialisation ($U(Role)$) and are acceptable environmental friendly options, public transports have high utilities of *Norm* ($U(Norm)$) and *Role* ($U(Role)$). Hence, they see the most increase in number, whilst soft mobility usage sees a small decrease. When *Social factors* combine with other determinants (i.e. At + SF and SF + Af), we can observe a minor decrease in car usage.

With the main focus on *Affect (Af)* determinant, more agents pick cars than the other modes due to their convenience, comfortability and privacy, i.e. $U(Emotion)$ is high. It also explains the figures when two determinants are combined. When *Affect* is not considered (i.e. At + SF), the car usage goes down. When it is put together with others, the number increases significantly (up to 40%). Therefore, we can conclude that *Affect* is the main driver for the car option, while *Social factors* can encourage people to use more public transport, especially for environmental reasons.

- **Intentional, Habitual and Facilitating condition determinants:** The results of the third-level determinants test can be seen in Table 5.4 and Figure 5.10.

TABLE 5.4: Result of comparing the third level of TIB's determinants (All units are in 10^9 kilometres)

Main determinant	Car	Bus / Tram	Train	Walk-ing	Bik-ing	Oth-ers
Reference population	73.09	4.07	23.2	2.67	4.91	4.42
Facilitating Conditions (FC)	46.18	16.03	33.44	6.2	5.94	4.56
Intention (I)	67.72	4.12	28.12	2.77	5.21	4.42
Habit (H)	50.92	13.96	32.34	5.97	4.33	4.83
FC + I	67.82	4.16	28.0	2.76	5.2	4.42
I + H	69.23	3.45	27.73	2.63	4.9	4.42
FC + H	51.05	14.09	31.75	6.1	4.48	4.88

Although we put the "inconvenience of public connections" as the criteria for *Facilitating Condition (FC)* (see Section 5.5), there is still a large number of households favour walking and public transport over the car. The represented agents of these households do not have a large $w(Facilitating)$. Therefore, this particular condition does not contribute to the final decision significantly.

The *Habit* test case also has a lower percentage of private vehicles than the reference. The simulation time is short (1 year), so the agents have not accumulated enough "experience". However, the majority of agents have habit weight smaller than intention weight, i.e. $w(Habit) < w(Intention)$.

In contrast, *Intention* emerges as an important factor for car usage since the final figure for this mode is 10% larger than the figure for either *Habit* or *Facilitating*

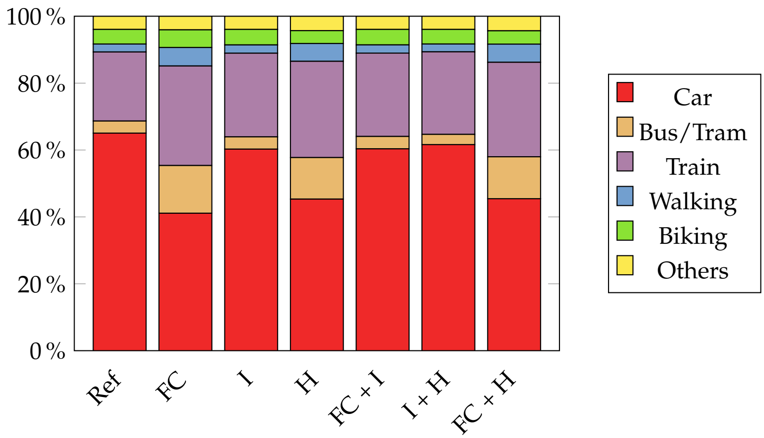


FIGURE 5.10: Percent composition of modes in the tests of third level of TIB's determinants

conditions. It can be confirmed in the combination cases where *Intention* is present, i.e. FC + I or I + H. Both of them have a higher number of car trips than other scenarios with only *Habit* or *Facilitating condition*. In TIB, *Intention* refers to the deliberation process of human decision-making, as opposed to *Habit*, which causes people to act on impulse. The simulation result at this level is an intriguing effect since social studies have pointed out that choosing to travel by private car is often considered habitual behaviour in a European country like Switzerland (e.g. [Kau00]).

The model allows us to isolate and link a determinant to its effect at the macro level. The current preliminary results observe the figure of car increase when the agents invoke *Affective* factors in the second level of TIB. The same pattern can be found where the agents put their *Intention* first by performing the deliberation process rather than acting based on past behaviours. On the contrary, *Social factors* and *Habit* appear to be the reason why the majority of Swiss households choose public or soft transport.

5.8 Advantages and limitations of our framework

Our framework allows users to incorporate the leading determinants of socio-psychological parameters: attitudes, social factors, affect and habit. We showed that BedDeM is capable of comparing the above determinants structurally by changing the associated weights. The modular architecture (see Figure 5.7) allows a programmatic approach to calculate the utility for each of these determinants. In turn, it connects the micro and macro levels so we can analyse and reason about the changes between the mode choices of individuals or groups. The performed simulation shows that our framework can be used as a tool to explore the implication of a set of factors/determinants on the whole system.

The experiment above could also not be repeated by utilising other state-of-the-art architectures in Chapter 3. The BDI architecture, its derivatives and other normative models only consider a smaller set of determinants compared to our setup in the experiment above. The cognitive architectures focus on understanding how people organise knowledge and produce intelligent behaviour based on knowledge derived from psychology experiments and employing quantitative measures. They do not cover concepts utilised in the experiment, such as social factors and affect. The MoHuB framework descriptive norm centres around the agent's components design and hence, is also not capable of specifying the levels of importance for different determinants. Similarly, Consumat framework only categorises different heuristics levels, not how determinant influences final output. Consequently, they have a disadvantage in this experimental causation design as we cannot create a link between micro (determinants) and macro (behaviour) levels.

There is still some small margin error from the calibration process (around 6.5% of the total scheduled kilometres). The agent's stochastic mode also does not reflect the macro patterns in MTMC effectively. Hence, we propose to address the issue of irrational behaviours by focusing on agents' learning in uncertain environments (e.g. reacting to changes in traffic rate/break-downs and others' opinions) in the upcoming developing stage. Currently, agents are only keeping track of the number of times they used a mode on trips with the same purpose, which accounts for *habit* in decision-making (see Section 5.5). The influence of past experience on the ranking functions (i.e. feedback loops) can be further extended by modifying the agent's belief about the consequences or changing the weights of determinants to prioritise better alternatives. Studies in Reinforcement Learning techniques (e.g. [Mni+15]) or Generalized Expected Utility Theory (e.g. [Qui12]) could be utilised for this purpose.

Chapter 6

Case study: The model of trust and reputation of rail service

The mobility model in previous chapter is extended in the this case study to reflect the change in demands to deal with the uncertainty of train schedules. The environment in this case is *stochastic*, *fully observable*, *static* and *unknown*. We utilise the idea of subjective logic to represent the trust of individuals and train's reputation in a network of agents. It also shows the effect of feedback loops in our framework. This chapter also contains the modified version of some elements from our publication [NS20a], including the related work, BedDeM's design, calibration and the experiment.

6.1 Introduction and description of case study

One of the main objectives of the social simulation research community is to create agent-based models capable of exhibiting human-like behaviours under the uncertainty of a complex environment [Dug13; Kam19]. Utilising trust and reputation can provide a way to express the confidence one can have in the quality of goods, services and even potential partnerships [Gra+15]. However, there is a lack of effort to develop models that focus on trust and reputation's impact and relationships with other determinants in human decision-making (see surveys such as [CCA15] and [Gra+15]).

The case study also aims to assess the ability to extend our utility function of the determinants to reflect more complex phenomena, i.e. trust and reputation. To represent them in an uncertain environment, we incorporate subjective logic to measure personal beliefs about available modal choices presented to the agent. Using feedback(s) from the environment, an agent can evolve its decision-making process over time depending on personal experiences and opinions from neighbours in its network. This case study also focuses on the effect of an environmental (external) factor on individuals' decision-making. In fact, in this case, the environment is considered to be *stochastic*, *fully observable*, *static* and *unknown*.

In this case study, we first consider some related projects that model trust and reputation at the individual level (Section 6.2). The following section describes the updated interaction mechanism between the environment and agents. Next, the specification of our decision-making process is provided, which contains the functions to calculate subjective probability and utility for each available option. A case study of the implementation platform - BedDeM - is then described in Section 6.5. Its main purpose is to reproduce the collective ground truth of the reputation for the rail services at both regional and national levels and to evaluate its impacts on the number of kilometres travelled. We then conclude our work in Section 6.6.

Since we use the same data sources as the previous mobility demand model, the description of the parametrisation of agents and environment and calibration are omitted. In fact, the main difference between this case study and the previous one is the stochastic environment. We will specify how the agent's interaction with the environment and its decision-making are updated accordingly.

6.2 Related work

Previously, we discussed the state-of-the-art agent's mobility model choice (see Section 5.2). In the section, we focus on agent-based models that incorporate individual trust and reputation in their decision-making. They can be divided into three categories:

- *Learning models*: The modeller studies the agents' behaviour patterns over a number of encounters. The agents' interaction results can be used directly as a measure of trust.
- *Reputation models*: The agent asks its network(s) about their opinions of the target(s).
- *Socio-cognitive models*: The agent model focuses on forming beliefs about different characteristics of potential partner(s) and reasoning about these beliefs to decide how much trust to put in them.

6.2.1 Learning models

Axelrod's tournaments of Prisoner's Dilemma [AD88] is the most cited example that illustrates the evolution of trust and cooperation over a sequence of interactions. Wu and Sun demonstrated that trust could emerge between agents using a cooperation strategy to adapt and evolve their relationships [WS01]. In their model, Sen and Dutta [SD02] show that reciprocity can emerge when agents are able to take into account future benefits of cooperation. Instead of building trust through a number of interactions, the model of Mukhejee et al. [MBS01] utilises mutually learning. Another learning method to simulate trust in a non-competitive environment is suggested in the work of Birk [Bir01]. All of these models assume that the environment is fully observable for the learning algorithms to work correctly and require strict assumptions to produce results.

Other models of this category implement trust metrics in agents so that they can evaluate how the target's action affects its goals using the data generated from their interactions. Witkowski et al. [FST01] developed a model of trading networks whose agent's trust in a partner is evaluated based on their past interactions and their type.

In situations that are new or unfamiliar, past observations are of little help in assessing trust. Chhogyal et al. propose a simple approach to trust assessment between agents based on shared values [Chh+19]. Recently, many investigations have been undertaken on trust mechanisms using various methods to evaluate the trustworthiness of an agent, such as fuzzy logic theory (e.g. [RRF18; Sch+07]), Bayesian network (e.g. [NB18; MB17]), game-theory (e.g. [Yin+19; Yah12; CP05b]). However, these mathematical-based models can be problematic if the perceptual input is wrong or incomplete.

6.2.2 Reputation models

The models in this subsection derive reputation as an aggregation of opinions from members of a community about one individual. For example, the concept of *referrals*, i.e. pointers to other sources of information, is used in the work of Yu and Singh [YS02]. The author developed an agent that can explore a network using referrals from its neighbours and gradually build up a model of its network. Zuo and Liu present a model for mobile agents to select the most reputable information host to search and retrieve information [ZL17]. They use opinion-based belief structure to represent, aggregate and calculate the reputation of an information host. Other high-level concepts, such as neighbours and groups, are employed to assess the reliability of the witnesses in the model of Sabater and Sierra [SS01]. However, all the mentioned models assume that agents that act as witnesses do not have personal agendas and share information truthfully.

Aggregating ratings is an alternative method popular within online communities. It is applied in the peer-to-peer rating system [Son+05] on the *eBay* website. However, this system has some limitations regarding the unresponsiveness of users or sellers' fraud with fake ratings. To deal with the absence of data, Yu et al. created a model that utilises Dempster Shafer's Theory of evidence [YFK94] and referrals, which handles the lack of belief in the agent as a state of uncertainty and combines various sources to derive reputation. The work of Schillo et al. furthers this technique to allow agents to handle lying witnesses using learning instead of subjective probability [SFR00]. Recent technologies, such as blockchain, have been utilised to provide the architecture and implementation of a system that allows the agents to interact with each other and enables tracking of how their reputation changes after every interaction [For+19; Cal+18].

6.2.3 Socio-cognitive models

The models above consider mainly the outcomes of interactions between agents. The subsection considers the models that utilise an individual's subjective perception of indirect interaction to enable a broader analysis of the nature of the potential partner [Gam+00]. Typical examples can be seen in the work of Castelfranchi and Falcone [CFP03; FC01], which is inspired by human behaviours and use the BDI agent architecture. Brainov and Sandholm [BS02] developed a strategy for an agent to model an opponent's trust with a rational approach. Bendiab et al. proposed a model that relies on fuzzy cognitive maps for modelling and evaluating trust relationships between the involved entities [Ben+19].

These socio-cognitive models are often being criticised for lacking empirical grounding in their reasoning mechanisms that other model types can provide [Bra+18; Gra+17; RHJ04]. A potential solution is using the same assessment method proposed above, i.e. taking into account trust, reputation and motivation of potential partner(s) over a number of interactions. However, it can be computationally intensive to have agents consider all aspects that influence their trust in others.

6.3 Interaction between the environment and agents

In this case study, the train timing provided by the environment is uncertain during the decision-making of the agent. It only knows whether a trip is on time or not after the action is performed. This information is returned to the agent in the form of *feedback*. Compared to the previous case where it is simply omitted, the feedback has a binary value (on time or late) which is randomly generated based on a success rate (in percentage). For example, the Valais region of Switzerland has a success rate of 90%, which means 90% of the total trips made by train are on time, and 10% are late. In the current model, an entity that extends the *Environment* class represents a region in Switzerland. Hence, the success of a train trip can be specific to that particular region.

In terms of the interfaces provided in Section 4.3, we implement the **Feedback** interface to provide a boolean value *isLate* to indicate the lateness information of the trip made by the chosen mode of transport (see Figure 6.1).

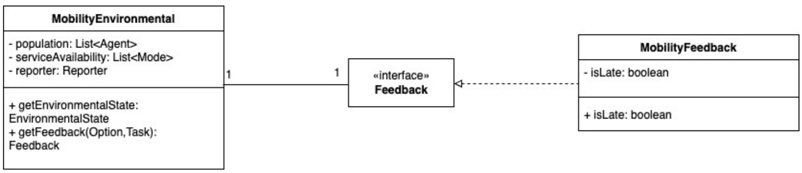


FIGURE 6.1: The implementation of the Feedback interface for the trust and reputation case study

6.4 Agent's decision-making process

A fundamental aspect of the human condition is that nobody can determine with absolute certainty whether a proposition about the world is true or false. Thus, whenever a statement is assessed, it should be done under the view of an individual and not be represented in a general and objective belief [Jøs97]. Reviews from [BK86] and [HH88] provide good examples of how standard logic and probabilistic logic are designed for an idealised world, where these important aspects are often not considered, and conclusions have to be drawn from insufficient evidence.

6.4.1 Subjective logic

In this case study, we follow the framework proposed by Jøsang [Jøs16] on subjective logic. It presents a specific calculus that uses a metric called *Opinion* to express our subjective beliefs about the world. An opinion, denoted as $\omega_x^A = (b, d, u, a)$, indicates party A's belief in statement x. In this case, b, d and u represent *belief, disbelief* and *uncertainty* respectively, where $b + d + u = 1$ and $b, d, u \in [0, 1]$. All sets of possible opinions can be mapped into a point inside an equal-sided triangle in Figure 6.2.

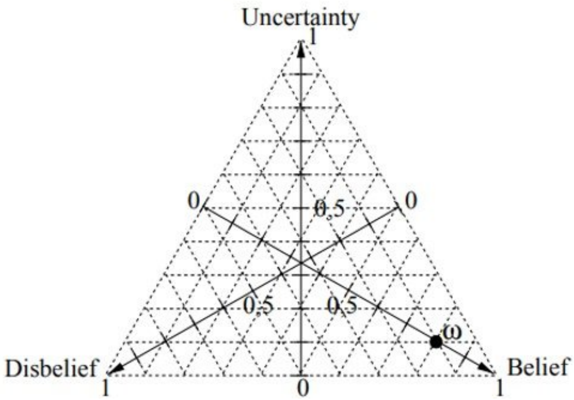


FIGURE 6.2: The opinion-space trust model / opinion triangle, taken from [Jøs16, p. 15]

The base rate parameter, $a \in [0, 1]$, determines how uncertainty shall contribute to the probability expectation value (see [Jøs16, p. 14]):

$$E(\omega_x^A) = b + au \tag{6.1}$$

In the binary event space (i.e. where there are only two possible outcomes - success or failure), subjective logic allows us to build an opinion from a set of evidence about x using the following equation (see [Jøs16, p. 16]):

$$\omega_x = \begin{cases} b_x = \frac{r}{r+s+W} \\ d_x = \frac{s}{r+s+W} \\ u_x = \frac{W}{r+s+W} \\ a_x = \text{base rate } x \end{cases} \tag{6.2}$$

where r is the number of positive evidence about x, s is the number of negative evidence about x, W is the non-informative prior weight, also called a unit of evidence, normally set to 2 and the default value of base rate a is usually set at $1/2$. In the case of no prior experience with the target, agent A's opinion of x is set as $\omega_x^A = (0, 0, 1, 1/2)$. Therefore, its probability expectation value is $E(\omega_x^A) = 1/2$.

Jøsang also proposed the consensus rule of independent opinions A and B [Jøs02]:

$$\omega_x^{AB} = \omega_x^A \oplus \omega_x^B = \begin{cases} b_x^{AB} &= \frac{b_x^A u_x^B + b_x^B u_x^A}{k} \\ d_x^{AB} &= \frac{d_x^A u_x^B + d_x^B u_x^A}{k} \\ u_x^{AB} &= \frac{u_x^A u_x^B}{k} \\ a_x^{AB} &= \frac{a_x^A u_x^B + a_x^B u_x^A - (a_x^A + a_x^B) u_x^A u_x^B}{u_x^A + u_x^B - W u_x^A u_x^B} \end{cases} \tag{6.3}$$

where $k = u_x^A + u_x^B - u_x^A u_x^B$. If k = 0, an alternative equation is applied instead:

$$\omega_x^{AB} = \omega_x^A \oplus \omega_x^B = \begin{cases} b_x^{AB} &= \frac{\gamma b_x^A + b_x^B}{\gamma + 1} \\ d_x^{AB} &= \frac{\gamma d_x^A + d_x^B}{\gamma + 1} \\ u_x^{AB} &= 0 \\ a_x^{AB} &= \frac{\gamma a_x^A + a_x^B}{\gamma + 1} \end{cases} \tag{6.4}$$

where $\gamma = u_x^A / u_x^B$, is the relative uncertainty between ω_x^A and ω_x^B. The consensus operator is commutative and associative, and thus, it allows the combination of more opinions. We model reputation in this case study using this consensus notion.

6.4.2 Updated utility function

Triandis suggests that one of the main factors to determine the intention of a behaviour is the value of perceived consequences, C, depending on the sum of the products of the subjective probability that a particular consequence will follow a behaviour (P_c) and the value of (or affect attached to) that consequence (V_c) (see page 16 [Tri77]). Thus, the equation for the utility of C is as follows:

$$U_c = \sum_{i=1}^{n} (P_{c_i} V_{c_i}) \tag{6.5}$$

where n is the number of consequences that a subject perceives as likely to follow a particular behaviour. The P_{c_i} value can be derived from Equation 6.1, i.e. $P_{c_i} = E(\omega_{c_i})$.

We update the full decision-making cycle with TIB's determinants, as illustrated in Figure 6.3. In the first level, the *Time* determinant is divided into two possibilities, on time and late. Each is also associated with a success rate in percentage.

In the model, the utility function of determinant *Evaluation* is an exception which follows Equation 6.5. In addition, the expected value of determinant *Norm* can be derived from the probability expectation value of the collective opinion formed by the consensus rule (see Equation 6.1 and Equation 6.3). In Section 6.5, we focus on a simplified binary event space, i.e. an action has two outcomes - success or failure. More complex scenarios could be considered in the future by extending the result space to multiple dimensions (e.g. time, cost, satisfaction).

We modify the running example in Table 4.2 to test the agent's belief in the *Time* of a trip. In this case, U is a cost function. In other words, when comparing two

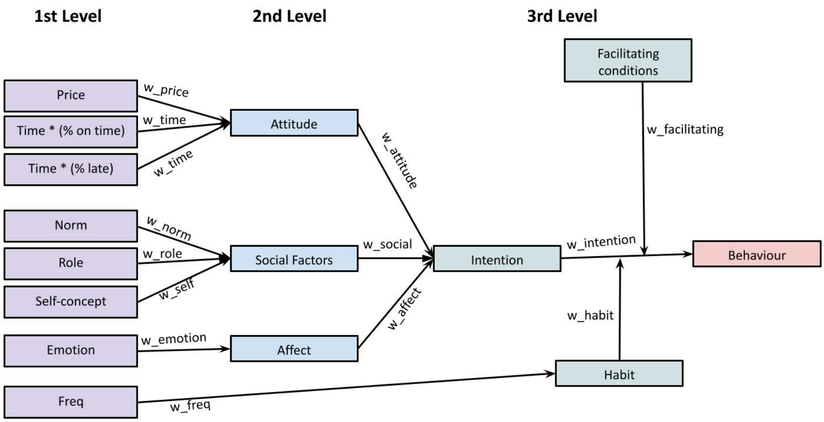


FIGURE 6.3: TIB's determinants mapping for the trust and reputation case study

options, the agent prefers the one that has smaller value. We assume that an agent has access to three options: *biking*, using a *car* or taking a *train*. It expects that a *car* journey would take around 0.3 hours for good traffic, which is believed to be a 20% chance. A late *car* drive would take up to 1.3 hours. Using subjective logic Equation 6.5 only for *Time*, we have $U_{car}(Time)$ = 0.3 * 20% +1.3 * 80% $\approx$ 1.1. In contrast, if 90% of trains are believed to only take 0.2 hour and the rest take 1.2 hour, $U_{train}(Time)$ will be 0.2 * 90% + 1.2 * 10% $\approx$ 0.3. If the agent has measured the exact biking time as 1 hour, $U_{bike}(Time)$ would simply be 1 * 100% = 1. Their total value, $\sum U(Time)$, is 2.4. If $w(Price)$ and $w(Time)$ are 2 and 4 respectively, the new expected value in the next level ($U(Attitude)$) of the *car* would be 4/7*2 + 1.1/2.4*4 $\approx$ 2.98, the *train* would be 3/7*2 + 0.3/2.4*4 $\approx$ 1.36 and the *bike* would be 0/7*2 + 1/2.4*4 = 1.67. By continuing to follow Equation 4.7 to the *Behaviour output* level, we have $U_{car} \approx$ 0.87, $U_{train} \approx$ 0.69 and $U_{bike} \approx$ 6.94. Instead of using the *car* as shown in Section 4.2.4, the best option for the agent is now *train* as it has the lowest utilities. This scenario shows that the agent's belief in the duration of a trip can change the final output of decision-making.

TABLE 6.1: Running example of an agent's decision-making adapted from Table 4.2 with updated belief in *Duration*

Level	Determinant	w	EU
1st	Evaluation (Price - Swiss franc), Belief = 100%	2	U_{car} = 4 U_{train} = 3 U_{bike} = 0

Continued on next page

Table 6.1 – continued from previous page

Level	Determinant	w	EU
	Evaluation (Time - hours)	4	U_{car}(0.3 - 20% / 1.3 - 80%) = 0.3 * 20% +1.3 * 80% $\approx$ 1.1 U_{train}(0.2 80% / 1.2 20%) = 0.2 * 90% + 1.2 * 10% $\approx$ 0.3 U_{bike} (1 100%) = 1 * 100% = 1
	Norm (similarity with others)	3	U_{train} = 1 U_{car}= 2 U_{bike} = 3
	Role (environmental friendliness)	2	U_{car} = 3 U_{train} = 2 U_{bike} = 1
	Self-concept (personal preference)	3	U_{car} = 1 U_{train}= 2 U_{bike} = 3
	Emotion (enjoyment)	1	U_{car} = 1 U_{train}= 2 U_{bike} = 3
	Frequency (past similar trips - note that lower value means more usage)	3	U_{car} = 0 U_{train}= 0 U_{bike} = 1
2nd	Attitude (Evaluation + Belief)	4	U_{car} = 4/7*2 + 1.1/2.4*4 $\approx$ 2.98 U_{train} = 3/7*2 + 0.3/2.4*4 $\approx$ 1.36 U_{bike} = 0/7*2 + 1/2.4*4 = 1.67
	Social factors (Norm + Role + Self-concept)	2	U_{car} = 1/6*3 + 3/6*2 + 1/6*3 = 2 U_{train} = 2/6*3 + 2/6*2 + 2/6*3 $\approx$ 2.67 U_{bike} = 3/6*3 + 1/6*2 + 3/6*3 $\approx$ 3.33

Continued on next page

Table 6.1 – continued from previous page

Level	Determinant	w	EU
	Affects (Emotion)	2	$U_{car} = 1/6*1 \approx 0.17$ $U_{train} = 2/6*1 \approx 0.33$ $U_{bike} = 3/6*1 = 0.5$
3rd	Intention (Attitude + Social factors + Affect)	4	$U_{car} = 2.98/6*4 + 2/8*2 + 0.17/1*2 \approx 2.83$ $U_{train} = 1.36/6*4 + 2.67/8*2 + 0.33/1*2 \approx 2.23$ $U_{bike} = 1.6/6*4 + 3.33/8*2 + 3/1*2 \approx 7.90$
	Habit (Frequency)	3	$U_{car} = 0/1*3 = 0$ $U_{train} = 0/1*3 = 0$ $U_{bike} = 1/1*3 = 3$
	Facilitating conditions (lower mean easier to access)	2	$U_{car} = 0$ $U_{train} = 0$ $U_{bike} = 0$
	Behaviour output		$U_{car} = 2.83/12.96*4 + 0/2*3 + 0/0*2 \approx 0.87$ $U_{train} = 2.23/12.96*4 + 0/2*3 + 0/0*2 \approx 0.69$ $U_{bike} = 7.9/12.96*4 + 3/2*3 + 0/0*2 \approx 6.94$

6.5 Experiment with trust and reputation of train

This section focus on the usage of BedDeM to perform some experimentation regarding the effects of trust and reputation in mobility modal choices. It is worth noticing that BedDeM is not a routing model, and hence, agents only take feedback from the environment as an indication of whether a trip is either a success or a failure.

The experiment also assesses whether the updated agent's decision-making can capture the ground truth of real-world data. In addition, we consider some test scenarios that highlight the effects of trust and reputation on mobility modal choices.

6.5.1 Design

The purpose of this experiment is to investigate the effect of public transportation's reputation on its demand, which is measured in yearly total kilometres travelled. The reputation, in this case, is represented by a constant percentage (ground truth value), i.e. punctuality. When the agent chooses to perform its trip by public transportation, we randomly generate a number. If this value is under the ground truth, the region will output a successful signal to the agent and vice versa. These consequences are connected with the duration of a trip for the agent. Successful feedback means the agent has performed the trip within its time estimation. In contrast, a failure message from the region means the trip is late and running time is doubled as a penalty. These information are taken into account in the next agent's decision-making as explained in Section 6.4. Following the decisions of all agents, the region computes the new reputation of the service by joining its residents' opinions using the consensus rule (see Equation 6.3), while ignoring all empty opinions. Currently, the agents do not have the ability to lie about their opinions, and thus, the reputation reflects collective perception about the punctuality of service. The national-level reputation of service is simply the combination of all regional reputation by also utilizing the consensus rule.

In this study, we focus on changing the successful rate of the rail service operated mainly by Swiss Federal Railways. On the one hand, punctuality is an important determinant of quality of service [Dur+17] and, in turn, affects yearly demand. On the other hand, other determinants (e.g. speed of trains, average fare per km and safety) can also contribute to the customer's perception of the performance index [Dur+17]. Hence, the rail service provides an excellent testing ground to observe the contribution and effects of trust, reputation along with other socio-psychological determinants in the agent's decision-making process.

A summary of the setup for the experiment can also be found in Table 6.2. We can divide them into five main categories, whose details are as follow:

TABLE 6.2: Collective ground truth / Punctuality (in percentage) of rail service in individual region for different testing scenarios

Scenario	Region 1	Region 2	...	Region 26
Perfect world	100	100	...	100
Real-world	87.8	87.8	...	87.8
High expectation	87.8	87.8	...	87.8
Low expectation	87.8	87.8	...	87.8
Disrupt in Region 1	***25***	87.8	...	87.8
Disrupt in Region 2	87.8	***25***	...	87.8
...	...	...	...	...
Disrupt Region 26	87.8	87.8	...	***25***

- *Perfect world scenario:* The success rate of all trips is kept at 100%. This setup is similar to what has been done previously in [NS19a]. Therefore, we expect almost no uncertainty in the agent's trust. It can act as a base to compare the effects of trust and reputation when they are implemented in later settings.

- *Real-world scenario:* According to [Rai21], around 87.8% of total trips were punctual in 2015, which is the year we calibrated our agents [NS19a]. We assign this number to all regional rail trips' success rates.

- *Agents starting with low or high expectation scenario:* The initial value of the probability expectation for an option x, $E(\omega_x)$, will be set at either 0 or 1 instead of 0.5, as suggested in Section 6.4. The punctuality of rail service is still kept at 87.8%. This setup changes the agents' parameters instead of the regional setting.

- *Regional disruption scenarios:* The punctuality of rail service in a single region is set at 25%. Otherwise, it is similar to the real-world case, i.e. 87.8%. These structures can be used to test the effect of disruption in one single region on the figures of other regions and the national level.

6.5.2 Results

As the model contains an element of randomness, we perform 100 simulations for each scenario and compute the averages. The results below show the reputation in terms of percentage and the demand for the rail service at the national level and three representative regions 1, 24 and 18. They represent the highest, midpoint and lowest in terms of kilometres demand in the perfect world scenario. Other disruptive scenarios follow the same pattern as these representatives.

TABLE 6.3: Result of reputation (measured in percentage) of rail service at the national level and three representative regions

Scenario	Reputation (percentage)			
	National	Region 1	Region 24	Region 18
Perfect world	***99.6***	99.9	99.4	95.7
Real world	***86.3***	89.0	87.8	83.0
High expectation	***87.7***	87.6	82.5	84.8
Low expectation	***85.2***	87.3	86.5	74.5
Disrupt Region 1	85.2	***25.7***	88.2	86.5
Disrupt Region 24	87.4	88.5	***22.0***	95.7
Disrupt Region 18	86.5	86.7	88.3	***12.8***

- **Reputation**: Table 6.3 shows the percentage of trust results from different scenarios. In essence, reputation measures at the national level and regions with significant train usage (e.g. Region 1) are nearly the same as the ground truth (see Table 5.1). In contrast, the lower the overall kilometres travelled by agents is, the less accurate the cumulative reputation of the whole region. This effect can be seen clearly in the figures of Region 18. It is mainly because reputation is a cumulative trust in neighbouring agents (see Function 6.3), which is also cumulative "experience" (see Function 6.2). In other words, the more agents in a network use the service, the more accurate they are in reflecting the ground truth.

When a disruption happens in a single region, the rail's reputation varies but stays around or below ground truth (25%). For example, the reputation of Region 1 is around 25% in the case of disruption at Region 1. The reputation of Region 24 is 22% when there is a problem at Region 24. Conversely, we observe no notable change in the national figures, i.e. around 87%. Except for the disruptive regions, other regional figures also show no significant difference. Although reputation is the cumulation of all agents' "experience", this result indicates that the inter-connections between agents from separate regions are weak or limited; therefore, interference cannot spread.

- **Total rail-kilometres**: Table 6.4 provides results in terms of total kilometres demand of train service in different scenarios. Having agents starting with high expectations affects the overall kilometres travelled positively, with the national figures almost doubled and high increases in all regions. Compared to others, this scenario shows the most different from the ground truth, especially in the case of Region 18. In the experiment design, the reputation of the train connects directly to the utility of *time* in the agent. This determinant is considered in our framework design to calculate the final utility value of the train option (see Figure 5.7). Hence, the higher the probability of trains arriving on time and having no penalties, the higher train usage. The agents were able to update their beliefs to be closer to the ground truth value (see Table 6.3). However, the updating process was significantly slower, which created this opportunity for the exceptional growth of total kilometres demand of train service in this scenario.

TABLE 6.4: Result of total demand (measured in kilometres travelled) of rail service at the national level and three representative regions

Scenario	Total demand (10^6km)			
	National	Region 1	Region 24	Region 18
Perfect world	***9081.588***	3114.568	104.632	0.173
Real world	***8958.592***	3112.998	104.632	0.173
High expectation	***16000.392***	5002.029	142.432	97.759
Low expectation	***8878.993***	3105.089	104.632	0.173
Disrupt Region 1	8865.848	***3105.089***	98.867	0.173
Disrupt Region 24	8937.483	3112.998	***102.267***	0.173
Disrupt Region 18	8947.010	3112.178	103.876	***0.173***

Conversely, by implementing a lower percentage of trust and reputation at the start of the simulation, we observe a significant decrease in usage at the national level and major regions. For instance, agents starting with low expectations cause a slight drop in national and Region 1's total, around 10 million kilometres. However, the demand in Regions 24 and 18 do not exhibit any sizeable changes in this scenario. As mentioned previously, the agents in these regions do not have enough "experience" to reflect the ground-truth value of the train's reputation. It leads to less accurate information in the agent's perception, so the reputation is still relatively high. In addition, as trust and reputation are only a subset of determinants in TIB, these observations also signify that agents have considered not only the probability expectation values but also other environmental differences.

There is no extensive impact on the national figures in the disruptive cases compared to the real-world scenario. The reduction is noticeable when the disruption happens in regions with significant demand where punctuality is the key issue (e.g. Zurich, Geneva). In contrast, the interference does not substantially change in the regions that initially have a small consumption number (i.e. Region 18 and 24). These observations are similar to the pattern in reputation columns as the total demand. Therefore, we can conclude that the usage of our model can create a direct connection between the agent's perception of the train's punctuality and the decision to use this mode of transport.

6.6 Advantages and limitations of our framework

In this case study, our main contribution in this case study is the development and implementation of a novel concept using subjective logic to represent trust and reputation elements in our decision-making framework that utilises TIB. It also shows that our utility function can be extended to cover more complex social phenomena. The experiment above provides a way to test this mechanism to see whether it can replicate the punctuality of rail service and observe the effect of trust and reputation on the number of kilometres travelled.

Compared to previous case study, in the agent's decision-making process, we adopt another psychology theory to our framework to better reflect the uncertainty of the environment: subjective logic [Jøs16]. It requires modifying the original utility formula to update the belief determinant. Preliminary results show that the reputation of the rail system in the agents' opinions can reflect the ground truth in the setup. The current model demonstrates a difference in reputations at two different levels - regional and national. We also detect the link between the train's reputation, time (individual determinant) and mobility demand (final decision output). This connection is a result of using the layered structure in our framework (see Figure 6.3). In addition, the belief in train reputation is one of the many determinants in the agent's decision-making (e.g. cost, ranking in environmental friendliness, and comfortability); thus, there are differences when comparing the change in total kilometres demands at the regional level in the high/low expectation and disruptive cases. By changing their weights in the expected utility function (similar to the previous case study), we also have the opportunity to explore further and compare their contribution to the agent's decision-making, which can affect the overall mobility demand, as we observed in the experiment.

Apart from what was discussed in the assessment of the case study (see Section 5.8), there are other limitations in this study. Agents have access to different modes of transport. Assuming that most of them can be perfectly predicted by the agent is unrealistic. A potential solution is to capture the dynamic traffic information in the real world to better capture whether the trip is on time or delayed at different hours of the day and in different regions. We also acknowledge that the mechanism to incorporate trust and reputation in this case study is simple and based on a theory (i.e. subjective logic [Jøs16]). Using qualitative data to inform agents' behaviours is also an ongoing research area within ABM community [An+21; RDG21; Edm15].

Chapter 7

Case study: The model of purchasing vehicle

The third study is about the decision of purchasing a model of vehicle. It also explored aspects of *bounded rationality*, mainly by utilising the Perception and Decision components of the agent, in a *deterministic*, *partially observable*, *static* and *known* environment. This chapter contains the modified version of some elements from our accepted paper [NPS22], including state-of-the-art, BedDeM design, calibration and the experiment.

7.1 Introduction and description of case study

The number of ABMs used to represent human decision making are increasing. Agent designs with the notion of perfectly rational maximise expected utility but crucially ignore the resource costs incurred. Investigations in Bounded Rationality (BR) offer an alternative to how to model behaviours in an uncertain environment with limited available cognitive resources. However, the ABMs utilised in these researches often focus only on simulating one particular type of BR and do not include other alternatives (see surveys such as [KAR18; Cas+20] and Section 7.2). This study assesses the ability of our framework to be extended to reflect the impact of multiple BRs on decision-making.

This case study follows the definition provided by Carley and Gasser [CG99] regarding two types of bounds in agents - limits to capabilities (i.e. the agent's physical, cognitive and computational architecture) and limits to knowledge (i.e. the ability to learn and construct intellectual history). We focus on using the *Perception* component to actively limit the agent's capability to observe relevant information. In this sense, BR is modelled as an extension of the model of the perfectly informed, optimised individuals to account for limited knowledge and resources, i.e. a form of *bounded optimality* [RN10, p. 1050]. Combining with the notions of bounded rationality in the work of Simon [Sim78] and the biases and heuristics advanced by Kahneman and Tversky [KT13], the following phenomena can be targeted for simulation:

- *Sequential decision-making* refers to algorithms that consider the dynamics of the world, thus delaying parts of the problem until they must be solved [FR14, p. 337].

- *Emotional decisions* happen when the people's emotional state influences the depth of information processing related to decision-making [Sim87].
- *Habit formation* is the process by which a behaviour becomes automatic when it is repeated with a routine [Sim87].
- *Multiple criteria* other than cost can be considered, depending on the decision-making context [RN10, p. 622-628].
- *Confirmation bias* is the tendency of people to select the information that supports their views, ignore contrary information, or when interpret ambiguous evidence as supporting their existing beliefs or values [Nic98].
- *Bandwagon effect* is a psychological phenomenon in which an idea or belief is being followed because everyone seems to be doing so [KS14].

We acknowledge that this list is limited and only covers the general ideas of each BR. However, it represents topics that are often mentioned in ABM research (see surveys such as [KAR18; Cas+20]) and provides a starting point for what can be considered in our study.

The TIB model developed in previous case studies has the ability to consider the impacts of *sequential*, *emotional*, *habitual* and *multiple criteria* decision-making. In this study, we modify the *Perception* component to consider the *confirmation bias* and the *bandwagon effect*. The environment is deterministic though partially observable due to the modified *Perception* component.

This case study will focus on Swiss personal mobility, an area where BR and information imperfections are particularly pervasive, as decisions are made on the level of heterogeneous individuals and households. It partially explains why the mobility sector remains one of the most challenging sectors (generating about one-third of the total CO_2 emissions) for the transition to net zero emission goals [Bou+21]. Purchasing new vehicles is an essential field for the energy transition strategy, especially when it provides an understanding of the need of individual consumers and requirements for future infrastructure. Due to the significant number of individual decision-makers involved and alternatives offered, ABM is often utilised for the assessment of BR's effects in the lab, as well as in the field (e.g. [HMS09; Kim+11]. Therefore, this area provides the most suitable test bed for assessing the effects of BR on energy-related decision-making processes.

This study contains the modified version of some elements from our accepted paper [NPS22], including state-of-the-art, BedDeM design, calibration and the experiment.

After considering some of the related ABM architectures in Section 7.2, we describe the process of parametrisation of agents and the environment and how they interact with each other in the next two sections. Next, the decision-making structure of our agent-based model is specified in Section 7.5. A case study is then provided to evaluate the result of applying this bounded Perception in Section 7.8. Finally, we conclude and suggest further development in Section 7.9.

7.2 Related work

In this section, we summarise agent-based architectures and frameworks that addressed the BRs mentioned, including sequential, emotional, habitual, multi-criteria decision-making, confirmation bias, and the bandwagon effect. Their full description can be seen in Chapter 3.

Sequential decision-making can be implemented as multiple steps/stages in decision-making in an agent to derive the action output. A typical example is Belief-Desires-Intentions (BDI) model [Geo+98]. Its extension (e.g. BOID [Bro+02], eBDI [Per+05], BRIDGE [DDJ08]), normative and and cognitive architectures that consist of a perception-deliberation-action cycle also belong to this category.

There are several BDI models that consider emotions agent decision-making, such as [SLC19; SDM+10]. Nevertheless, only a number of agent architectures cover emotions explicitly in their components, including eBDI [Per+05], PECS [Urb00], and BRIDGE [DDJ08]. eBDI added a dimension of emotion as an extension of the BDI architecture. As the name suggests, PECS provides a component-oriented agent architecture with integrative modelling of physical, emotional, cognitive and social influences. BRIDGE can use the *Ego* component to determine emotional responses to a number of different stimuli. However, they have limited practical application and have been mostly used as reference model [BG14].

In ABM literature, habits are often incorporated with hybrid approaches that have both heuristics and deliberative decision-making. Examples include the BRIDGE [DDJ08]and Consumat [JJ02] agent architectures. BRIDGE agents can utilise the idea of the basic needs of Maslow [Mas43] to overrule any deliberate decision-making process to make sure they can react when needed. The Consumat framework allows to model habit as one of the five heuristics to be utilised in place of the deliberation process in uncertain environments when the agent has low cognitive effort.

The *multi-attribute utility theory* can be used to represent the preferences of an agent over a number of alternatives under conditions of uncertainty [RN10, p. 622-628]. Thiriot and Kant created a multi-objective multi-agent system (MOMAS) to take into account the possible trade-offs between conflicting objective functions [TK06]. In this case, the users often have to define the criteria based on statistics or previous empirical studies.

Confirmation bias is used as a way to filter various sources of information in the perception phase. A typical example is the BDI agent filters information from all perceptions and other sensor stimuli using semantic association rules derived from its internal beliefs. The BRIDGE architecture includes an *Ego* component with several filters and ordering preferences to interpret the input stream of information to form the beliefs in the agent. The effect of confirmation bias can also be found in the opinion dynamics modelling frameworks. A quasi-Bayesian belief updating framework was proposed by Sobkowicz, in which incoming information is filtered by the cognitive biases or predispositions of the agent [Sob18]. In the work of Rollwage and Fleming, meta-cognition (accuracy of belief formation) of agents use confirmation bias to down weight contradictory information [RF20].

As the bandwagon effect is associated with the ability to consider social learning, it can often be found in normative models. Several architectures of this category can be found in literature, including BRIDGE, EMIL-A [And+07b], NoA [KN03] and Consumat. The BRIDGE design takes into account many social concepts, such as social interactions and culture. The EMIL-A model incorporates norms into decision-making through the processes of learning about norms in a society, the internalisation of norms and the use of these norms to derive the most appropriate action. The NoA architecture extends the notions of norms to include organisational concepts and ideas from legal systems (i.e. states of affairs are either obligatory, permitted or forbidden). The Concumat agent is able to reason by comparing the success of their actions to the success resulting from the actions of their neighbours. If its actions are less efficient, the agent simply the action of others.

Although many of the architectures above account for multiple aspects of behaviour, the agent architectures and implementations surveyed above do not comprehensively cover all BRs effects mentioned in Section 7.1. Therefore, in this case study, we aim to create an agent model capable of considering all these BR effects in its decision-making scheme.

7.3 Parametrisation of the environment and agents

7.3.1 Dataset

In this study, we utilised the same agent's population derived from the previous case study with data from MTMC and SHEDS (see Section 5.3.3). Further data from SHEDS related to mobility ownership, such as type and efficiency of the current vehicle, year of purchase, the price at purchase and kilometres made by the vehicle, is also extracted. In addition, we utilise a Swiss car catalogue [QE20] to capture the car models available in the market. It includes data about engine type, energy label, market price, brand and years of availability.

7.3.2 Environment parametrisation

The environment in this case study is considered to have the following properties: *deterministic*, *partially observable*, *static* and *known*. It includes two main entities: *Market* and *Opinion Platform*.

Figure 7.1 demonstrates the implementation of the **Environment** class from our framework (see Section 4.3). Using data captured from the Swiss car catalogue [QE20] and estimated by our economist specialist, the *Market* includes details of the currently available vehicle **Model**:

- Engine: the type of the engine, e.g. electric, gasoline, diesel, hydrogen or hybrid.
- Energy labels: A+/A, B, C, D, E, F.
- Price: the current purchasing price.

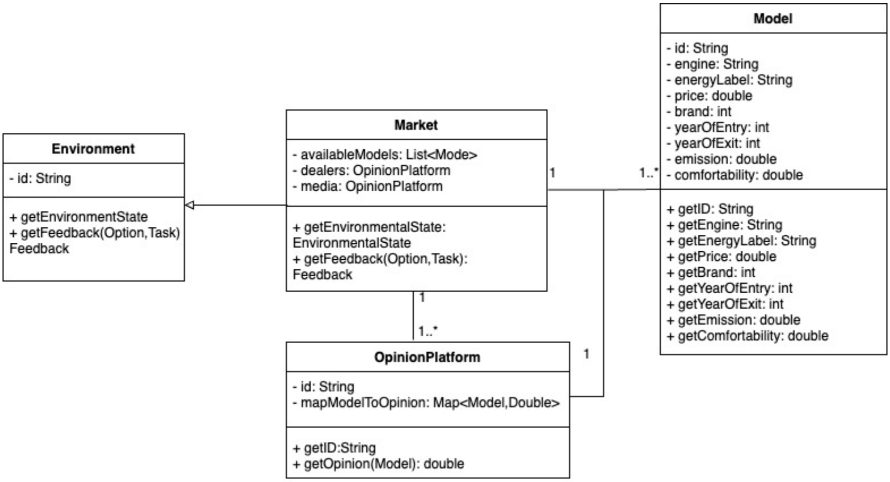


FIGURE 7.1: The implementation of the Environment class for the vehicle purchasing case study

- Brand: eight different groups of brands based on the place of production and associated image.
- Year of entry: the year in which the vehicle appears on the market.
- Year of exit: the year in which the vehicle no longer exists on the market.
- Comfortability: a number assigned by our economic collaborator to compare the luxury class of the vehicle, e.g. Sports car (50) versus Family car (30).
- Emission in terms of CO_2 per kilometre.

The *Opinion Platform* provides reviews (value from 0-1) from the neighbourhood, dealers and media. The neighbourhood network is set up similar to the previous case (see Section 5.3). An agent can subscribe to different media or dealers, which are biased toward specific models of different brands and engines. Media opinions can affect agents at the national level, while dealers target agents at the cantonal (i.e. regional) level. The subscriptions of agents are currently randomised in the model.

7.3.3 Agent parametrisation

We can define the following components in our agent framework: *InternalState*, *Task* and *Option*. Figure 5.4 is an UML diagram illustrating the implementation with the classes and interface provided in Section 4.3.

- **Agent's internal state**:

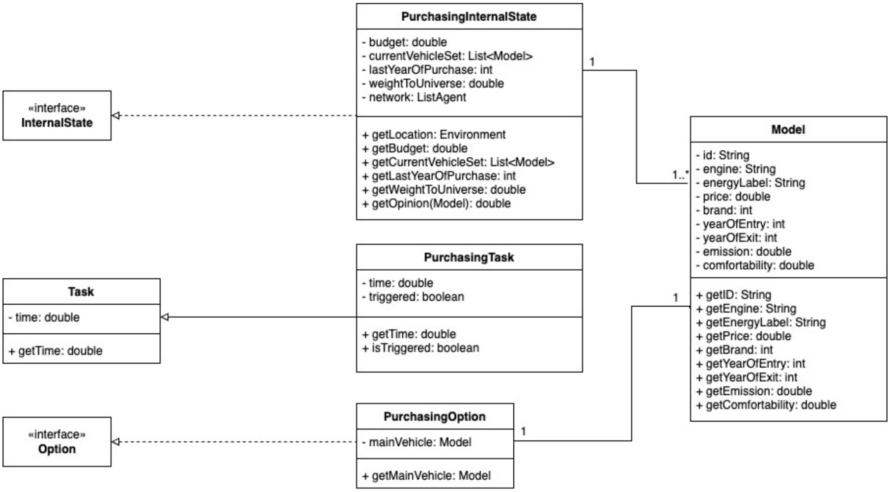


FIGURE 7.2: The implementation of the Environment class for the vehicle purchasing case study

- Budget: It represents the maximum price an agent can pay for a vehicle. To compute these values, we started with the last spending to purchase a vehicle for each profile in SHEDS data. Next, the GDP dynamics is divided by the population dynamics [10] to get a per-capita dynamics, which we can use as a global change of the budget.
- Current vehicle set: The current vehicle that the agent owns, as listed in SHEDS.
- Last year of purchase: The year that the vehicle was purchased.
- Weight to universe: The proportion of the population that the agent represents.
- Network: The agent keeps track of the neighbour network and all of its subscripted media and dealers. This neighbour network is created based on the location of the agent profile. The strength of the connection relies on its trust in different roles in society, which is extracted from SHEDS.

• **Task**: Two times a year, agents consider the available models on the market. If the condition for the trigger is satisfied, the agent will start the deliberation to decide which vehicle to buy. The probability of whether the agent purchases a model or not is calculated based on the following criteria:
  - Current vehicle kilometres made by the vehicle.
  - Family changes based on a yearly accumulation.

[10] https://data.sccer-jasm.ch/macroeconomic-drivers/2020-08-01/

  - Number of new models that are currently available.
  - Existence of a significantly better model than the one owned currently.
- **Option**: A model that is available on the market. There is also an option of not performing the scheduled activity due to the constraints from the agent's states (e.g. exhaustion of budget).

7.4 Interaction between agents and the environment

The *Market* is assumed to be open to all agents. In other words, agents have access to all the models that are available for that year. The *Opinion Platform* provides aggregated feedback to each agent depending on its subscription. Regarding our interfaces provided in Section 4.3, the **EnvironmentState** can be implemented to provide information on a certain model that is available in the market (see Figure 7.3) and its review (value from 0 to 1). At this stage of development, there is no response that the environment need to give back to the agent. Hence, the **Feedback** interface is not implemented in this model.

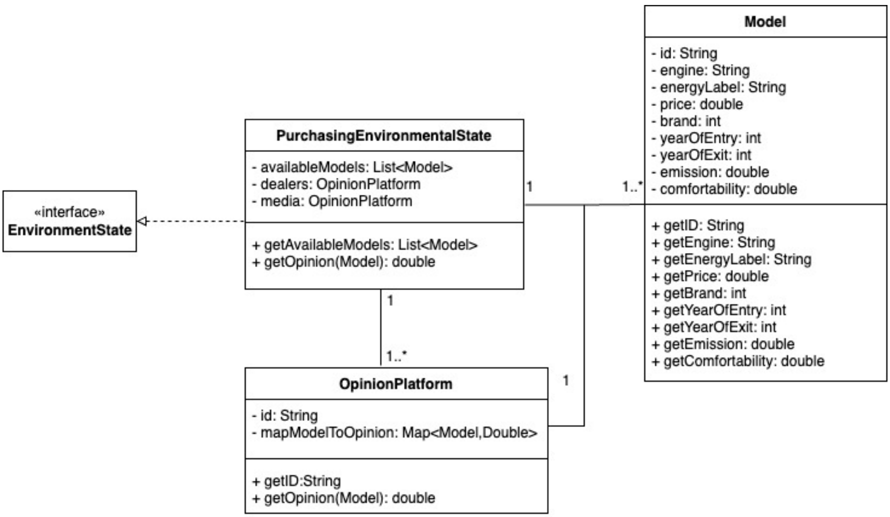


FIGURE 7.3: The implementation of the Environment class for the vehicle purchasing case study

7.5 Agent decision-making process

In this study, we create a new model to simulate purchasing a vehicle, utilising the Triandis model as a framework, as previously done in Chapter 5. In addition, we further extend the *Perception* component to cover several additional bounded rational aspects. Hence, BedDeM now addresses both the purchases of mobility resources and their usage in two intertwined decision-making schemes, each based on an implementation of the TIB model. The cycle starts with executing the mobility model to calculate the mobility demand for half a year. Then, the result (in terms of kilometres) is passed to the purchasing model. If an agent is triggered, it performs the purchasing decision-making and updates its available resources. Next, the mobility model runs again with an updated resource database for half of the year, at which point the purchasing model is carried out one more time. The cycle is repeated until the simulation timeline finishes.

The following section describes an overview of the model, including environments, purchasing agents and other existing entities, i.e. market and opinion platform. We then discuss how each BR type has been captured in the purchasing agent's architecture.

7.5.1 Model overview

The main components of our purchasing model are illustrated in Fig. 7.4). The environment contains not only the purchasing agents but also a *Market* and *Opinion Platform*. The *Market* stores all information about the available models, such as type, price and past sale figures. The *Opinion Platform* provides the recommendation level from the agent's neighbours, media and dealers. In this context, neighbours of a purchasing agent are other agents that connect with it in its social network. Media refers to public sources of information, such as television or magazine. Dealers are similar to media but only influence local levels. The recommendation of Media and Dealers and the agent's subscriptions are the interpretation of the data from [Web+17] by our economic specialist. The recommendation of an individual agent for a specific model starts at 0.5 and changes depending on whether the agent uses more or less the vehicle in the following year. All recommendations are taken into account in the *Perception* and the *Decision-making* components. Details are provided in the following subsections.

In each simulating decision cycle, an agent will perform the following sequence: First, it uses the *Perception* component to observe available models in the *Market*. Then, combined with information from the *Opinion Platform* of Neighbors, Media and Dealers and history of ownership, it filters, sorts and creates a shortlist of options. This list is passed to a *Trigger* component, which is implemented to reflect the irregularity of the purchasing activities. If specific criteria are achieved (e.g. number of new models, changes in the household, mileages of current car), the *Decision* component gets triggered. It follows the procedure of the TIB framework to evaluate the list of options in terms of a utility value. Finally, an option is selected based on the provided utility, either by choosing the best (deterministic agent) or using a probability (probabilistic agent). The *Communication* component then outputs this action to the current *Market* and updates the *Memory* state of the agent. It also informs the *Opinion Platform* for future reference by neighbours.

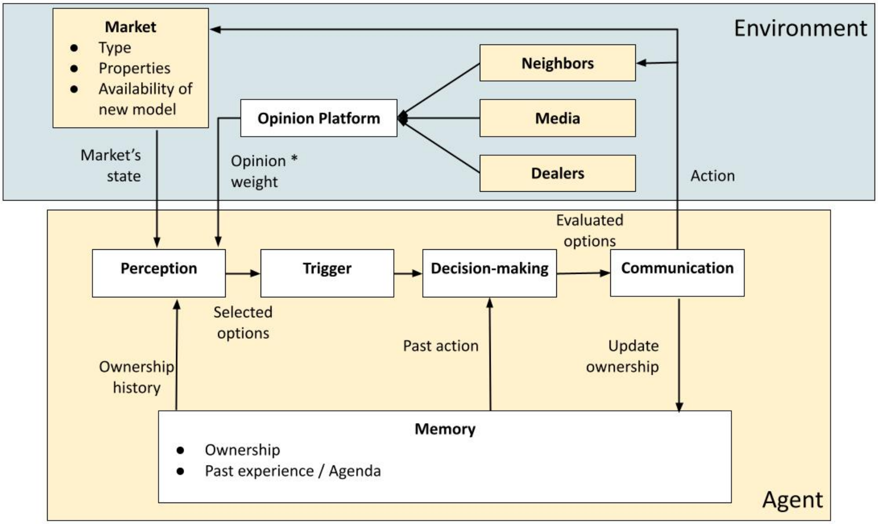


FIGURE 7.4: Agent's basic components for vehicle purchasing architecture

Compared to previous case studies, there is a *Trigger* component inside the agent due to the irregularity of the purchasing activities. It provides a probability of whether the agent purchases a model or not. In the latter case, the rest of the decision-making process is skipped. We calculate this probability based on the following criteria:

- Current vehicle kilometres made by the vehicle.
- Family changes based on a yearly accumulation.
- Number of new models that are currently available.
- Existence of a significantly better model than the one owned currently.

From the architecture above, the mentioned BRs can be captured in two agent components: *Perception* and *Decision*.

7.5.2 Perception component

The *Perception* component (see Fig. 7.5) first gathers information about the available options from the environment, including its neighbour's opinion. It then divides options into several lists, each satisfying specific criteria. These lists are then sorted, multiplied with specific weights and merged to form a list of selected options for decision-making. The criteria and their weights are based on the agent's personal

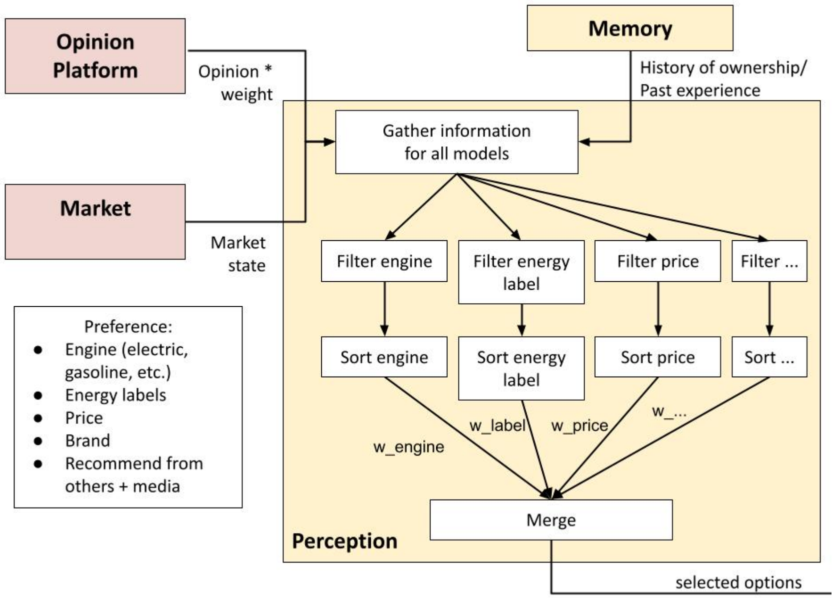


FIGURE 7.5: Perception component for vehicle purchasing architecture

preferences about the option's properties, which can be calibrated with empirical data.

The mechanism can be explained more clearer in the context of car purchasing. A consumer often starts by filtering out models that have a certain type of engine, price, energy labels and neighbour review. As human mental accounting mechanisms are limited [Hah+20], he/she has to sort the options to get the best one for each category and combine them to make a final list of available models for the final decision-making step.

Using this structure, the confirmation bias can be represented with the filtering process, with only relevant options being considered. The bandwagon effect is highlighted with the inclusion of the neighbour's opinion as one of the criteria. Using associated weights, modellers can decide on the influence of this effect on its final list of selected options for the agent's decision.

To capture the discussed BRs, the following references are included:

- Engine types (electric, gasoline, diesel, hydrogen, hybrid)
- Energy labels (A+/A, B, C, D, E or worst)
- Reserved price (a certain threshold depending on yearly income)

- Brands (group of producers divided by country of production and type of cars)
- Recommendation from Opinion Platform

With this setup, we can give an example. An agent wants to buy a car from the *Market*. It has the following references: electric engine (w = 0.5), below 30'000 (w = 0.25), and review larger than 0.5 (w = 0.25). Therefore, it has three sublists: 1) engine list, 2) price list and 3) review list. If the agent only wants to consider four models, we can cut two models from the engine list, one from the price list and one from the review list.

The current *Market* has a list of models with the following properties:

1. Engine: Electric, Label: A, Price: 30'000, Brand: 1, Review: 0.5
2. Engine: Electric, Label: A, Price: 25'000, Brand: 2, Review: 0.6
3. Engine: Hybrid, Label: B, Price: 25'000, Brand: 3, Review: 0.5
4. Engine: Gasoline, Label C, Price 20'000, Brand: 2, Review: 0.5
5. Engine: Diesel, Label C, Price 45'000, Brand: 1, Review: 0.4

The agent can then filter and sort these models into several sublists based on its references:

- Engine list: 1, 2
- Price list: 4, 2, 3
- Review list: 2, 1, 3, 4

After applying the cut for each of the lists, the final list of options for the *Decision* component includes 1, 2, and 4.

7.5.3 Decision component

Compared to the previous case study, the first level of the *Decision* component (Figure 7.6) is modified. *Price* and *Energy label* contribute to the result of determinant *Attitude*. *Social factors* consist of references from opinion channels, agent's status and brand bias. Comfortability is kept as the main contribution to *Affect*. The accessibility to refuelling points is considered as *Facilitating condition*. Finally, *Habit* takes into account the agent's history of similar models of the same engine type.

To implement these determinants using the framework pseudo code (see Section 4.3), the **LeafDeterminant** class is extended to different classes represent the determinants of the first level. Figure 7.7 demonstrates this process.

The *evaluateOptions* function is used to calculate the utility value U of the list of provided options. The $U(Price)$, $U(Label)$ and $U(Facilitating)$ can be extracted from a Swiss car catalogue [QE20]. The $U(Socialstatus)$ and $U(Emotion)$ are provided in

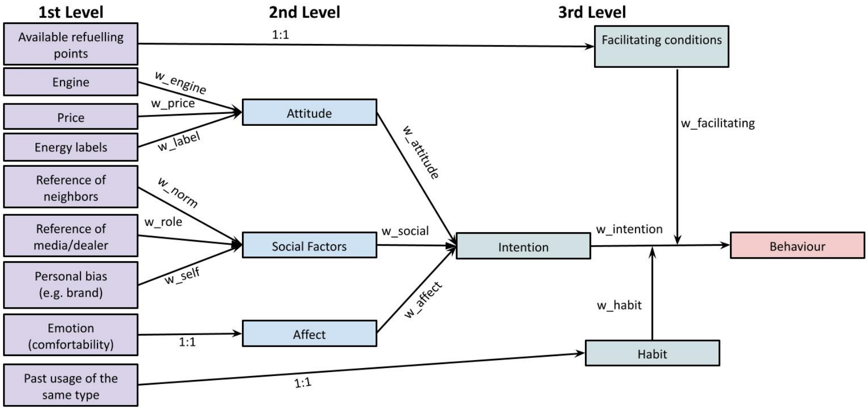


FIGURE 7.6: TIB's determinants mapping for the vehicle purchasing case study

[Web+17]. Weights of each determinant are calibrated, whose process can be seen in Section 7.7.

The expected value of each option $EU_d(opt)$ can be calculated or given a ranking value related to other option. They are as follow:

- $U(Engine)$ = 1 if the model's engine type is the same as the reference of agent; 0 otherwise.
- $U(Price)$ = MAX_PRICE - the price of the model.
- $U(Label)$ = 1) the model with label G, 2) the model with label F, ..., 6) the model with label B, 7) the model with label A.
- U(Review of neighbours) = value from 0(worst) to 1(best).
- U(Review of media/dealers) = value from 0(worst) to 1(best).
- $U(Brand)$ = 1 if the model's brand is the same as the reference of agent; 0 otherwise.
- $U(Comfort)$ = the vehicle's comfortability.
- U(Past usage) = 1 if the model has been used by the agent in the past; 0 otherwise.

7.6 Summary of the simulated bounded rationality effects

With the two *Perception* and *Decision* components, we can summarise how the BRs can be represented in our model:

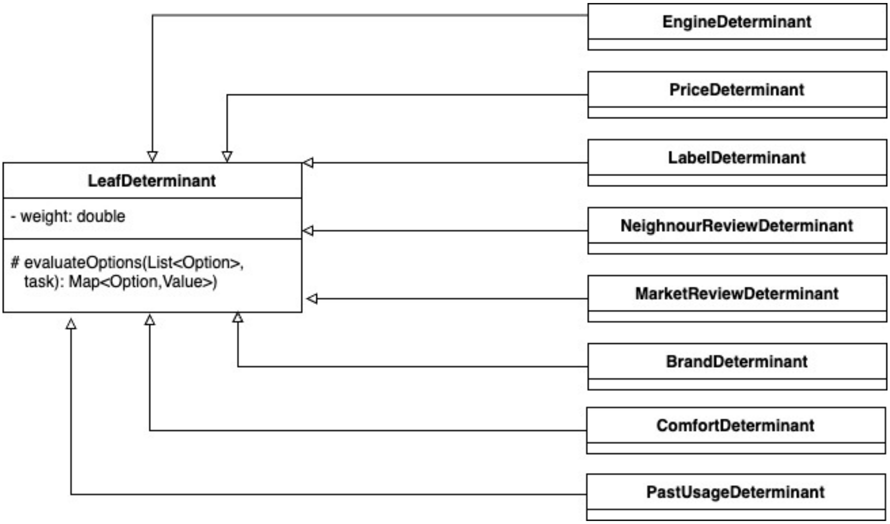


FIGURE 7.7: The implementation of agent decision determinant classes for purchasing vehicle study case

- **Sequential decision-making**: A decision-making cycle includes several steps, one after another. This procedure starts with the agent gathering information about the alternatives. Then, using its references, it filters, sorts, and cuts this list to a selected few options. If triggered, these selected options are evaluated in the decision-making component. Finally, the highest/lowest evaluated alternative is selected and communicated to the environment. Using a procedural approach, this process follows the description of sequential decision-making in Section 7.1, i.e. the current step waits for the result of the previous step.

- **Emotional decision-making**: It is captured in the determinant *Affect* in the second level of the *Decision* component (see Fig. 7.6). Its evaluation is dependent on the context of decision-making. For example, our purchasing agent can rank how much comfort/pleasure it can have from a model compared to others. The *Affect* determinant is associated with a weight ($w(Affect)$). By increasing this weight and lowering the weights of other related determinants, we can highlight the contribution of emotion to the overall behavioural output.

- **Habits**: Similar to emotion, the agent also accounts for past behaviour in its third level of the TIB framework (see Fig. 7.6). Its weight can be adjusted to mark its influence on the final choice.

- **Multiple criteria**: The TIB framework in the *Decision* component allows users to capture different factors in decision-making, i.e. *attitude (e.g. cost, time), norms, role, self-concept, emotion, habit, and past behaviour*. A mapping with empirical data can be provided better to interpret these factors in a decision-making context. Function 4.7 can be used to combine them in the form of a utility value. Using associated weights, the agent can also decide which has a more significant/lower impact on the final choice. This concept also allows the agent to express its preferences on specific criteria of decision-making.

- **Confirmation bias**: In the *Perception* component, an agent filters the information received from the environment to form different short lists of options. This process represents the idea that the agent selects the information that supports its preference. The associated weights of each criterion mark the contribution of this bias to the final list. For example, in the car purchasing context, an agent who only wants to receive information about electric cars has the filter to only allow electric engine cars and zero values for all weights, except for the engine's weight.

- **Bandwagon effect**: In its perception phase, the agent starts with observing its environment, including the patterns of its neighbours. It also accumulates the neighbours' opinions. This information is then used as a filter for in the Perception component (Figure 7.5) and fed into the Social factors determinant in the Decision-making component (Figure 7.6). Each of them is associated with a weight to provide a way to compare its effects to other factors in the decision-making.

7.7 Calibration

To calibrate this purchasing model, two different sets of parameters corresponding to different components - *Perception* and *Decision* - are selected.

In the *Perception* component, there are two main categories that are mentioned in Table 7.1: thresholds and their weights. The thresholds include: 1) preferred engine (Gasoline, Diesel, Electric, Hybrid, other), 2) energy label (A+/A, B, C, D, E and below), 3) price, 4) brand (1-8), recommendation level (value 0-1). In addition, each is associated with a weight, which also needs to be calibrated.

Regarding the *Decision* component, we calibrate the following determinants' weights: price, energy label, recommendation, social status, brand, emotion, habit, attitude, social factor, intention and facilitating condition (charging infrastructure). At this stage of development, all weights will take a value in the initial set (0, 0.25, 0.5, 0.75, 1).

The number of parameters is significantly large, increasing the combined number of test runs exponentially. Therefore, we choose to perform a sensitive test for all parameters. First, all parameters are set at the first item in Table 7.1, and the simulation is performed for reference. Each of them is then set at the last item value. By comparing the result of the simulation at this point with the reference, we can assess the impact of each parameter on the simulation outputs. The result of this sensitivity test can be seen in Table C.2. The parameters that make more impacts are then calibrated first with all the steps in Table 7.1. The less critical parameters are calibrated

TABLE 7.1: Calibration for agent's Perception component.

Parameter	**Logic**	**Item**	**N Step**
Filter - Engine Type	A preferred engine for each agent group	Gasoline - Diesel - Electric - Hydrogen - Hybrid	5
Filter - Energy label	A preferred energy label limit for each agent group	A+ - A - B - C - D/G	5
Filter - Price	Only take cars under budget	N/A	0
Filter - Brands	Clusters of brands for each agent group	1 - 2 - 3 - ... - 8	3
Filter - Recommendation	A real number represents the limit for accepted reference from others	0 - 0.25 - 0.5 - 0.75 - 1	4
Filter - New models	Select the new models only. No need for calibration	N/A	0
Percepted weight - Energy Label	Weight for energy label	0 - 0.25 - 0.5 - 0.75 - 1	4
Percepted weight - Price	Weight for price	0 - 0.25 - 0.5 - 0.75 - 1	4
Percepted weight - Brands	Weight for brands	0 - 0.25 - 0.5 - 0.75 - 1	4
Percepted weight - Recommendation	Weight for recommendation	0 - 0.25 - 0.5 - 0.75 - 1	4
Percepted weight - New models	Weight for new models	0 - 0.25 - 0.5 - 0.75 - 1	4

later and assigned only two steps (0-1) in data calibration. Next, the agents are divided into eight groups based on socio-economical analysis using data from [Web+17]. Each group has a set of parameters calibrated to have different values from the other group.

Our main objective is to minimise the error calculated by the total differences between the final number of vehicles purchased and real sales. If the difference is within an acceptable range, the error is multiplied by 0. Otherwise, it is multiplied by a weight (representing the adjusted importance):

1. If the error of the total unit sales is 8% , w = 0. Otherwise, w = 10.

2. If the error of the sum of sales of gasoline and diesel modes is 8%, w = 0. Otherwise, w = 2.

3. If the error of the negative of the sum sales of electric models is 5%, w = 0. Otherwise, w = 8.

4. If the error of the sum of sales of models with alimentation electric and hybrid is 5%, w = 0. Otherwise, w = 2.

5. If the error of the total sales of different clusters of models of different brands is 8%, w = 0. Otherwise, w = 2)

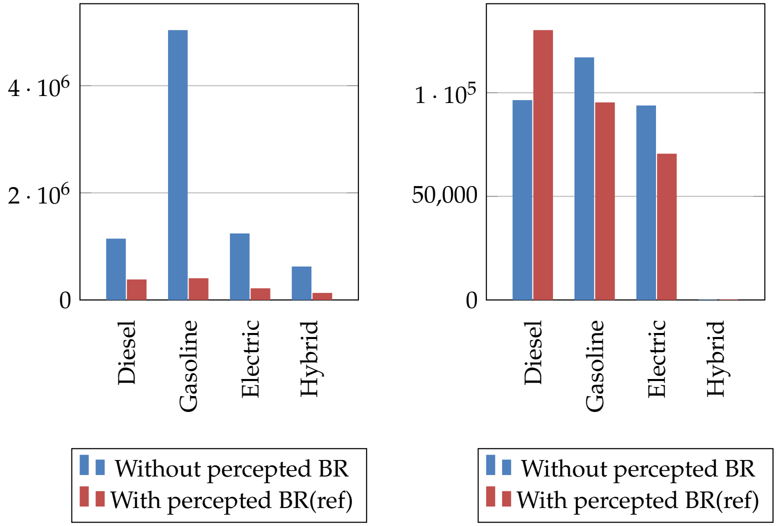
We calibrate with the data from 2015 to 2019. The more recent years, 2020 - 2021, are separated due to the effect of the COVID-19 pandemic. Therefore, its car stock is adapted directly from correspondences in SHEDS panel data. We repeat this procedure for all agent's profiles set at deterministic (i.e. choosing the best option) to find the smallest error. After a period of two weeks, the best setting satisfies the first, second and third criteria with the yearly average errors after multiplied with weights equal to 305'485. This setting can be seen in Appendix C.2.

7.8 Experiment

As the sequential, emotional, habitual, multi-criteria decision-making mechanisms are mainly implemented in the *Decision* component, their effects on behaviour can be observed by changing associated weights, similar to what was done in the previous case study (see Section 5.7). In this section, we focus on testing the functionality of the bounded *Perception* component in our agents. The number of vehicles (considered and purchased) calibrated for the final year (2019) is used as a reference. We experiment by turning the filtering, shorting and cutting functions off and evaluating the results against the reference. In this way, all models on the market (n=228) can be passed directly to the *Decision-making* component instead of only a limited number (max n=20) in the reference case.

Figure 7.8a shows the results as the number of models considered among the agent's population after the perception process. Figure 7.8b presents the final sales after the decision-making process. The figures are categorised by different engines, including diesel, gasoline, electric and hybrid vehicles.

The number of the models considered is much higher in the case without BR, especially for gasoline models (considered nearly 12 times). This bias is introduced with the models provided in the sample car catalogue [QE20]. In the reference case (with BR), the distribution between different engine types is more balanced though it is proportioned to the case. In the total sales of ground truth, the highest number is gasoline with 116 thousand vehicles. Even though electric vehicles are being considered more regularly, they have fewer sales, i.e. 93'591 cars. With the bounded perception applied, there are significant increases in the number of diesel cars sold. The gasoline and electric figures drop to 95'071 and 70'327 respectively. It is mainly due to betters models of diesel and fewer models from gasoline/electric type being selected after the perception phase. Overall, we can observe that the difference in the number of models being considered (individual perception level) can lead to the difference in the percentage of car types sold (macro level).



(A) Total models being considered

(B) Total sales

FIGURE 7.8: Simulation results in term of total number of vehicles per engine type

7.9 Advantages and limitations of our framework

This case study highlights the effect of determinants in long-term, strategic decision-making. The environment in this case study is considered to have the following properties: deterministic, partially observable, static and known. We adapt our simulation platform - BedDeM - to simulate the impacts of different types of BRs in the context of purchasing vehicles according to the definitions provided in Section 7.1. We acknowledge that these definitions are oversimplified and derived from the points of view of computer science concepts. However, the developed model demonstrated the flexibility of the framework components, which can be modified to further reflect other concepts from social studies. In particular, using the same design of *Decision-making* component as the previous case study, sequential, emotional, habitual and multiple criteria decision-making can be considered in the agent's architecture. The *Perception* component is further extended to cover the confirmation bias and bandwagon effect. The experiment shows that this extension has a certain impact on the results of the simulation at the macro level.

The increasing amount of agent parameters leads to a significant rise in complexity and time for calibration. In this study, we test the sensitivity of each parameter to the outcomes. If there is no noticeable change, the number of steps for calibration is reduced. Presetting parameters using literature/other theory data can also be a good substitution for calibration. Finally, depending on the context, it should be noted that our modular framework allows switching or turning off some determinants (by making their weight equal to 0) to reflect other behaviour theories.

Similar to previous case studies, the mapping for the first level of TIB is rather simple and intuitive. Although we have acquired the help of an economist specialising in environmental sustainability, it is also necessary to receive inputs from other fields to derive alternative mappings of empirical data (e.g. energy system engineering, consumer economics). In addition, the model's variability is not high. In particular, the associated weights of determinants are static over the simulation time, which means that the agents' preferences do not change over time. Further clustering techniques on SHEDS panel data of the past few years would also indicate how these weights have changed, especially when there is an innovation in technology or policy-making.

Chapter 8

Case study: The model of COVID-19 transmission

In the final study, we illustrate a working example of the public health domain - modelling the spreading of COVID-19 in a migrant centre. It focuses on rational behaviours in a *stochastic*, *fully observable*, *unknown* and *dynamic* environment. Due to limited socio-psychological data in this case study, we incorporate the elements of Maslow's hierarchy of needs inside the agent's decision-making mechanism.

8.1 Introduction and description of case study

The management of the COVID-19 pandemic in asylum centres is a critical public health issue, both because of the high risk of outbreak clusters and the socio-economic health preconditions of its populations. Indeed, high population density, belonging to a minority ethnic group or social deprivation are risk factors for contracting COVID-19 infection [de +20; SCM20; Ren+20]. To our knowledge, there is currently no published agent-based model analysing the transmission of the COVID-19 virus and the associated risk factors among asylum seekers living in asylum centres. Understanding these risk factors is crucial to determining targeted public health policies protecting these populations fairly and efficiently. In response, this case study was designed to explore the pandemic's spread into asylum centres (half-closed spaces) during the first wave of the pandemic in Switzerland. Specifically, it aims to identify which factors in migrant decision-making of daily tasks can increase the risk of COVID-19 infection after the first semi-confinement period (16 March to 27 April 2020).

Our agent decision-making framework is a suitable candidate due to its ability to isolate, highlight and link these factors to their effects at the macro level. However, in the given database, the amount of socio-psychological information is limited and insufficient for mapping all determinants in the TIB. To provide a way to derive the utility of an option, we decide to utilise Maslow's hierarchy of needs, which includes five layers: 1) physiological needs, 2) safety needs, 3) belongingness and love needs, 4) esteem needs, and 5) self-actualisation needs. This case study aims to address the question of the limitation of the model by assessing whether the modified version can be used in the same manner as previous studies.

In our model, each need in Maslow's theory is represented as a tank. Their volumes correspond to the levels in the hierarchy. If an agent satisfies a need, this need does not need to be recharged for some time. The utility of an option depends on the number of needs satisfied (details to follow in Section 8.5). This design allows us to test the adaptivity of our framework to the requirements of a new dataset (i.e. minimising the first level) and the exchangeability of a TIB factor, particularly *Attitude*. We also test our framework in an environment which is *stochastic*, *fully observable*, *unknown* and *dynamic*.

The next section considers the state-of-the-art that aims to model a certain population's reaction to a pandemic. Next, we provide a description of the database and how it is incorporated into the agents' and environment's parameters. The interaction between agents and the environment is then defined, followed by the description of the decision-making cycle. Finally, we summarise the advantages and issues of our framework in this case study.

8.2 Related work

In public health, Agent-Based Model (ABM) has historically been used almost exclusively to model infectious disease transmission and control in populations. Many ABMs of infectious disease transmission rely on the susceptible-infected-recovered (SIR) framework proposed by Kermack and McKendrick in the 1920s [SR13], in which the flows between susceptible, infected, and recovered states are governed by differential equations [EA96]. ABM extensions of SIR models have been used to introduce individual heterogeneity and more complex network interactions into these traditionally aggregate, compartmental models, providing further insight into infectious processes in real-world settings [Cho+16; EA96]. Notable ABMs in infectious disease epidemiology include comparisons of vaccination strategies to address a deliberate bioterrorist introduction of smallpox [Hal+02], tuberculosis control strategies [Mur02], use of targeted antiviral prophylaxis and social distancing measures to prevent an H5N1 influenza A (bird flu) pandemic [Fer+05], evacuation plans in the event of airborne contamination [EPH11], interventions to reduce human immunodeficiency virus (HIV) incidence [Esc+16; Mar+14], and vaccination strategies against influenza pandemics, including their impact on health care staff [Coo+10; Lee+10; Mar+12]. They have, thus, advanced to include increasingly sophisticated parameterisation of social networks and environmental influences to best inform public health policy and planning. Furthermore, many modelling capabilities developed, extended, and refined through infection-related ABM programs like MIDAS and the Framework for Reconstructing Epidemic Dynamics (FRED) [Gre+13], can be applied to public health problems beyond infectious disease.

In terms of COVID-19 simulation models, several agent-based influenza pandemic models have been repurposed to simulate the spread of COVID-19 transmission and the impact of social distancing measures in the United Kingdom [Fer+20], Australia [Roc+20], Singapore [Koo+20], and the United States [Cha+20]. Additionally, new agent-based models have been developed to evaluate the impact of social distancing and contact tracing [Ale+20; Kre+20; Kuc+20; Bla+20; Hoe+20] and super-spreading [Lau+20]. Features of these models include accounting for the number of households and non-household contacts [Cha+20; Kre+20; Kuc+20]; the age and clustering of contacts within households [Cha+20; Ale+20; Kuc+20]; and the microstructure in schools and workplace settings informed by census and time-use data [Ale+20]. Branching process models have also been used to investigate the impact of non-pharmaceutical intervention strategies [Pea+17; Hel+20] and the proportion of unobserved infections [Per+20]. Notably, Dignum et al. proposed a tool that uses Maslow's hierarchy of needs to analyse the health, social, and economic impacts of the COVID-19 pandemic when the government implements several interventions, such as closing schools, requiring that employees work in the assigned room, and providing subsidy for the population [Dig+20].

All these models are proper to explain the global behaviour of an epidemic on larger scales, considering general variables, e.g. economic factors, mobility and hygiene practices. However, there are many scenarios in which it is essential to analyse the socio-psychological factors in smaller populations or in facilities where the infection process can be identified by the interactions among their members [Qi+18]. The migrant population in the centre belongs to this category, where different determinants can play an important role in the transmission dynamics. A model developed using our framework allows researchers to explore the implication of a set of hypotheses regarding the spreading of COVID19 in this population.

8.3 Parametrisation of the environment and agents

8.3.1 Dataset

The data utilised in this study are part of SérocoVID, a seroepidemiologic study of COVID-19 infection conducted in the canton of Vaud, Switzerland [Mor+21]. Migrants living in an asylum centre, which is known to have had an epidemic outbreak, were invited to participate in this study. COVID-19 tests targeting the spike viral protein were measured in all participants. Each participant also completed a questionnaire measuring socio-demographic characteristics, medical history, health literacy, public health recommendations (wearing a mask, washing hands), behaviours and exposures (daily life activities, number of contacts weekly). The association of these independent variables with the serologic test result were estimated using a multivariable logistic regression model. The independent variables were obtained from the answers to the questionnaires divided into five main categories.

- Socio-demographic characteristics and health literacy, such as age, gender and education level.
- Health conditions, clinical risk factors and symptoms.
- Living conditions (e.g. access to single or family room, kitchen) and public health recommendations (e.g. wearing a mask in public, hygiene routines).

- Behaviour and exposure, such as place, context and number of meetings with other people.

8.3.2 Environment parametrisation

The migrants mainly stay within a centre, which has multiple facilities. For our framework, we consider the *Facility* as an extended class of the *Environment* where agents perform actions. Each has a capacity (i.e. the maximum number of agents allowed) and transmissible rate, which increases if an infected agent is present and decreases if agents apply hygiene rules. At the time of the investigation, there was no vaccine available and no mutated variants of the virus. In addition, there are two levels of individual infection risks of the agent if one is infectious, which is calculated using the model of [Lel+20]:

1. Masks and other preventive methods
2. No masks or other preventive methods

Whether or not an agent applies preventive methods is recorded in our dataset as an input variable. From Table 8.1, it can be seen that indoor actions can expose agents to more risks. One intriguing observation is that *staying in the assigned room* has the highest transmissible rate, but it is a top action due to their fear of the unknown outside environment (i.e. *Emotion* determinant).

TABLE 8.1: Action and associated location with two levels of transmissible rate

Action	**Location**	**With preventive methods (%)**	**Without preventive methods (%)**
Go to work	Workplace	2.5	6.2
Go shopping	Market	0.043	0.11
Stay in the assigned room	Small room	43	75
Stay in the assigned room	Family room	24	50
Eat in kitchen	Kitchen	0.22	0.55
Smoke outside	Garden	≈ 0	0.11
Socialise	Common room	5.9	14
Walk / Exercise	Outdoor	≈ 0	0.066
Talk with visitors / friends	Common room	5.9	14
Go to toilet	Toilet	5.1	12

After the agent chooses an action, a random percentage is drawn. If it is lower than the infection rate, the environment can signal the agent that it is infected. In this case, the environment is considered to be *stochastic* and *unknown*. In addition, the agent knows the environmental capacity and the number of people in it. Hence, it is also *fully observable*.

Using the classes and interface provided in Section 4.3, the implementation of the **Environment** interface (i.e. **Location**) and **Action** can be seen in Figure 8.1. Each **Location** has the following properties:

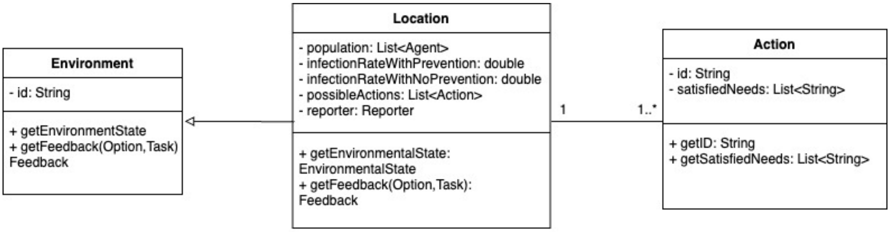


FIGURE 8.1: The implementation of the Location and Action classes

- Population: the current migrants presenting at a particular time of the day.
- Infection rate with prevention: the infection rate when the migrant applies preventive methods.
- Infection rate without prevention: the infection rate when the migrant does not apply preventive methods.
- Possible actions: the list of possible actions that can be performed in this location.
- Reporter: the report class records the results (e.g. infection in the location over time).

The **Action** class contains a list of possible needs that can be satisfied if the agent performes this action. Details of the need's implementation can be seen in Section 8.5.

8.3.3 Agent parametrisation

Since the number of individual records in the database is low (N=107), all profiles can be represented by agents. We can define the following components in our agent framework: *Internal state*, *Task*, *Option*. Figure 8.2 is an UML diagram illustrating the implementation with the classes and interface provided in Section 4.3.

- **Agent's internal state**:
  - Age
  - Gender

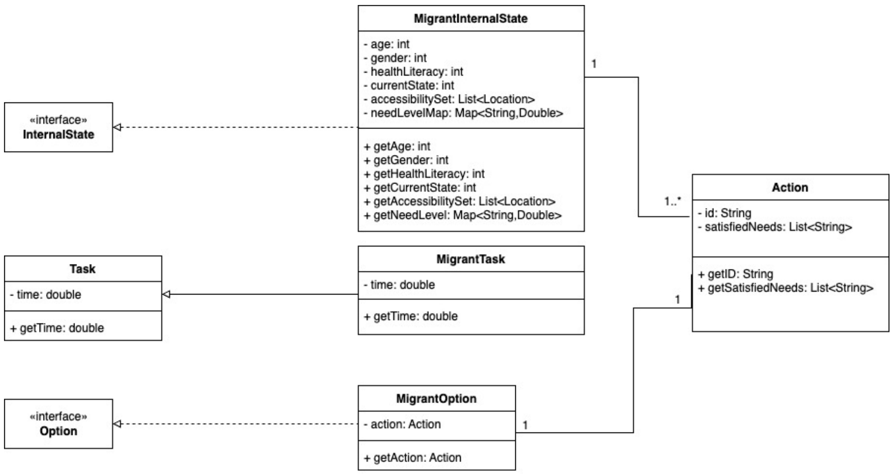


FIGURE 8.2: The extension of the Environmental and Mode class

  - Health literacy
  - Current state: 1) healthy, 2) asymptomatic, and 3)symptomatic. The agent transitions from one state to another in 14 day period.
  - Accessibility set: List of rooms or facilities that the agent has access to.

- **Task**: The agent has to pick out the top three actions to perform in a day, representing three day periods: morning, afternoon and evening.

- **Option**: The agent can choose the following actions: 1) Go to work, 2) Go shopping, 3) Stay in the assigned room, 4) Eat in the kitchen, 5) Smoke outside, 6) Socialise, 7) Walk/Exercise, 8) Talk with visitors/friends.

8.4 Interaction between the environment and agents

The environment entities, i.e. facilities, transmit the signal of estimated occupancies. If the activity can only be performed outdoors (e.g. eating, exercising, walking), then the capacity is usually unlimited. After an agent makes a decision, the environment updates its capacity and passes this information to other agents.

In terms of our interfaces provided in Section 4.3, the **EnvironmentState** can be implemented to provide information on the list of available actions at different locations accessible by the agent (see Figure 8.3).

We implement the **Feedback** interface to provide a boolean value *infected* to indicate whether the agent is infected after performing an action(see Figure 8.4).

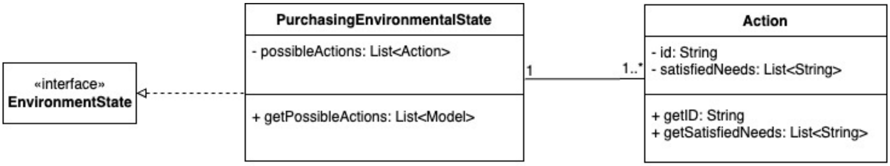


FIGURE 8.3: The implementation of EnvironmentalState interface for COVID19 case study

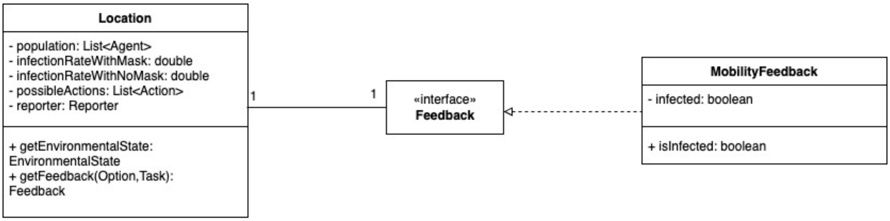


FIGURE 8.4: The implementation of Feedback interface for COVID19 case study

8.5 Agent decision-making process

In 1943, Abraham Maslow proposed the most widely known hierarchy of psychological needs [Mas43]. As seen in Figure 8.5, it includes five different levels: physiological, safety, love/belonging, esteem and self-actualisation. Furthermore, it states that people tend to spend more of their available time and resources on satisfying these basic needs before the higher order [TG13]. His approach generalised human motivations in various contexts but must be adapted for specific cultures or among different parts of the same society.

Each need in Maslow's hierarchy has a particular importance, which was utilised as a specific tank volume that required to be satisfied in a week (see Table 8.2). *Physiology* has a volume of 7 to represent seven days a week, which is also the time for all tanks to reset. It has the most significant amount, so more actions are required to fulfil. On the opposite side, *Self-actualisation* has a volume of 3 and requires the smallest number of actions to satisfy.

An agent will pick out the top three actions to perform in a day, representing three day periods: morning, afternoon and evening. There are some necessary actions (e.g. going to the toilet and going back to its assigned room), we assume that they are given for each period, and agents will be exposed to the risk in these locations. In addition, it can decide to do some of the actions in Table 8.3; each of which is associated with several fulfilled Maslow's needs. It should be noted that the action *Smoke outside* is only available for the smoker profile.

We also implemented a health cycle for COVID-19 patients. An agent will be in the state of *asymptomatic* after five days since first infected. It then becomes *symptomatic*,

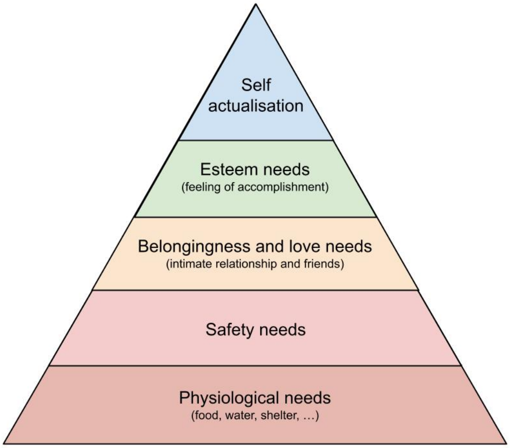


FIGURE 8.5: Maslow's hierarchy of needs (according to [DETL16])

TABLE 8.2: Mapping Maslow's need to tank volume.

Maslow's need	**Tank volume**
Physiology needs	7
Safety needs	6
Belongingness and love needs	5
Esteem needs	4
Self-actualisation needs	3

which means it mostly stays in a private room. Then, after 14 days, it returns to a healthy state, and this cycle starts again.

To incorporate Maslow's hierarchy of needs, the utility value of Evaluation follows Equation 8.1:

$$U(Evaluation) = \sum_{m=1}^{M} \text{Reminder of a tank(m)} - 1. \tag{8.1}$$

where M is the set of fulfilled Maslow's needs. If the remainder of a tank is 0, its need is no longer required to be satisfied and is excluded from the equation.

We perform data mapping according to Figure 8.6. In the first level, the *Evaluation* determinant uses the utility function above 8.1. The *Social Factors* determinant prioritises the actions that have direct contact with others, such as socialising with others in the common room, talking with visitors, eating in the kitchen and smoking. Contrarily, migrants often have fear as an *Emotion*. It leads to actions that avoid direct contact with others, such as staying in the assigned room or going for a walk. The

TABLE 8.3: Action and the fulfilled Maslow's needs.

Action	**Fulfilled Maslow's need**
Go to work	Self-actualisation Esteem
Go shopping / Pick-up essentials	Physiology
Stay in the assigned room	Safety
Eat in the kitchen	Physiology Belongingness and love
Socialise in common room	Belongingness and love
Walk / Exercise	Physiology Esteem
Talk with visitors / friends	Belongingness and love
Smoke outside (for smoker only)	Physiology Belongingness and love

Frequency of past behaviours is calculated by accumulating the past actions of the agent. In the second level, all utilities are kept the same as they are one-to-one connections to the first level. After that, the weights of *Attitude*, *Social Factors* and *Affect* are combined with these utilities using function 4.7 to produce *Intention*'s expected utility. The *Facilitating Condition* determinant depends on the previous action. If the agent is already outside the facility, they can continue with other related actions, e.g. going to work or exercising. The value of *Habit* is taken directly from *Frequency*. It is then coupled together with other third-level determinants to form *Decision output*.

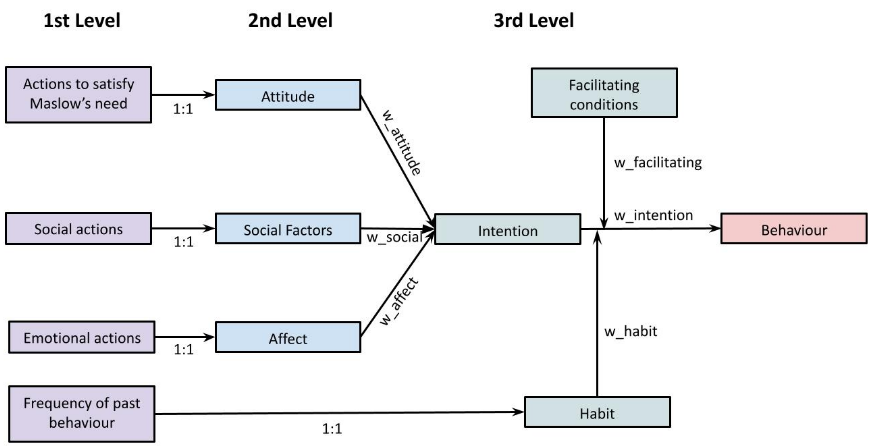


FIGURE 8.6: TIB's mapping for the COVID-19 case study

In terms of our framework implementation, we extend the **LeafDeterminant** to represent the determinants in the first/second level of TIB (see Figure 5.8). The *evaluateOptions* function can then be implemented with the expected utility of each option $U_{opt}(d)$, which can be calculated or be given a ranking value related to other options:

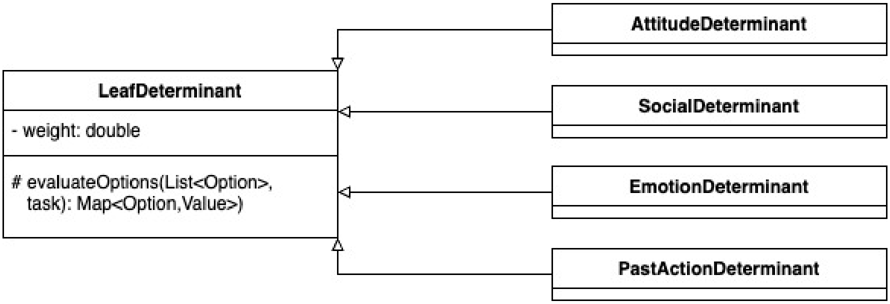


FIGURE 8.7: The implementation of Feedback interface for COVID19 case study

- $U(Attitude)$ = how much the action can balance Maslow's needs.
- $U(Social)$ = ranking based on how much interaction an agent can have during the performance of an action. We give a higher score for more social actions: 6) Socialise in the common room, 5) Talk with a visitor, 4) Eat in the kitchen, 3) Smoke with others, 2) Go to work, and 1) Other actions.
- $U(Affect)$ = ranking based on which can reduce the feeling of fear. We give a higher score for the action that is most employed in the survey: 6) Stay in the assigned room, 5) Walk/exercise, 4) Go shopping/pick up essentials, 3) Smoke , 2) Go to work, and 1) Other actions.
- $U(Facilitating)$ = value 0/1, representing whether the previous action happened at the same location.
- $U(Frequency)$ = the number of times the agent used the same mode the previous year.

We acknowledge that the above mapping is oversimplified, partially due to the limited dataset. However, the mapping is also designed to be easy to understand and intuitive, so further surveys can be performed to explore relevant psychological aspects of migrant actions. For our experiment, it is sufficient to be able to consider all the leading aspects: *Attitude*, *Social Factors*, *Affect*, *Facilitating Conditions*, *Habit* and *Intention*.

Table 8.4 shows a running example in the COVID-19 case study, which follows the TIB determinants mapping in Figure 8.6. An agent has three options for an afternoon:

eating in the kitchen, *staying in the assigned room* or going for a *walk*. We assume that the agent just started the week, and so its tank volumes for Maslow's needs are equal to what is shown in Table 8.2. According to Table 8.3, *eating in the kitchen* can fulfil the belongingness and physiological needs; *staying in the assigned room* fulfils the safety and physiological needs; *walking* satisfies physiological and esteem needs.

Using Equation 8.1, the utilities of *Evaluation* are: $U_{eating}(Evaluation)$ = (5-1) +(7-1) = 10, $U_{stayingintheassignedroom}(Evaluation)$ = (6-1) +(7-1) = 11 and $U_{walking}$ (Evaluation) = (4-1) +(7-1) = 9. Their total is 30. It should be noted that U is a maximising function, i.e. option that has a larger value is preferred. Hence, the agent would choose to stay in the assigned room. In terms of *Social actions*, we apply the rankings for $U(Socialactions)$, i.e. *eating in the kitchen* is 4, *walking* and *staying* is 1. Similarly, $U_{eating}(Emotion) = 1$, $U_{staying}(Emotion) = 6$ and $U_{walking}(Emotion) = 5$.

The second level's determinants have the same utility as the first level due to the one-to-one mapping (see Figure 8.6). We assume the weight of *Attitude*, *Social factors* and *Affect* are 2, 1, 2 respectively. Applying Equation 4.7, the expected value of *Intention* of *eating in the kitchen* would be 10/30*2 + 4/6*1 + 1/12*2 = 1.5. Similarly, $U_{staying}(Intention)$ is 1.9 and $U_{walking}(Intention) = 1.6$. As the agent is already outside, its utilities for *Facilitating Conditions* of *eating in the kitchen*, *staying*, *walking* are 1, 0 and 1. At behaviour output, expected values are U_{eating} is 2.45, $U_{staying}$ is 2.52 and $U_{walking} \approx 3.03$. These utilities indicate that *walking* would be the best option for this agent. We choose this example to highlight the importance of *Facilitating Conditions* in decision-making because the best choice would have been *staying in the assigned room* if the agent only made an evaluation based on its *Intention*.

8.6 Calibration

There were 107 migrants who participated in the survey. From the previous preview, we divide the total population into different groups following characteristics as proposed in [Mor+21]:

- Age group (12-20 and >20)
- Gender (Male/Female)
- Health literacy (Good/Medium/Bad)

Using these characteristics, the agent population can be grouped into 12 profiles. The weights of each agent profile will be calibrated so that the number of final infections is the same in real-life data. In terms of the *Decision* component, we calibrate the following determinants' weights: $w_{social}, w_{affect}, w_{habit}, w_{condition}$. At this stage of development, all weights will take a value in the set (0, 0.25, 0.5, 0.75, 1). Our main objective is to minimise the error calculated by the total differences between the final number of infections in the simulation and real data.

There are 20'736 configurations to be tested. Since the results involve some randomness in the environment, we perform 100 simulations for each configuration and calculate the average. Overall, the total number of runs is 2'073'600.

TABLE 8.4: Running example of an agent's decision-making in COVID-19 case study

Level	Determinant	w	EU
1st	Evaluation	1	U_{eating} = (5-1) +(7-1) = 10 $U_{staying}$ = (6-1) +(7-1) = 11 $U_{walking}$ = (4-1) +(7-1) = 9
	Social actions	1	U_{eating}= 4 $U_{staying}$ = 1 $U_{walking}$ = 1
	Emotion	1	U_{eating}= 1 $U_{staying}$ = 6 $U_{walking}$= 5
	Frequency	3	U_{eating}= 2 $U_{staying}$ = 8 $U_{walking}$= 6
2nd	Attitude	2	U_{eating}= 10 $U_{staying}$ = 11 $U_{walking}$= 9
	Social factors	1	U_{eating}= 4 $U_{staying}$ = 1 $U_{walking}$= 1
	Affects	2	U_{eating}= 1 $U_{staying}$ = 6 $U_{walking}$= 5
3rd	Intention	4	$_{U}eating$= 10/30*2 + 4/6*1 + 1/12*2 = 1.5 $U_{staying}$= 11/30*2 + 1/6*1 + 6/12*2 = 1.9 $U_{walking}$= 9/30*2 + 1/6*1 + 5/12*2 = 1.6
	Habit (Frequency)	2	U_{eating}= 2 $U_{staying}$ = 8 $U_{walking}$= 6
	Facilitating conditions	2	U_{eating}= 1 $U_{staying}$ = 0 $U_{walking}$= 1
	Behaviour output		U_{eating}= 1.5/5*4 + 2/16*2 + 1/2*2 = 2.45 $U_{staying}$ = 1.9/5*4 + 8/16*2 + 0/2*2 = 2.52 $U_{walking}$= 1.6/5*4 + 6/16*2 + 1/2*2 = 3.03

The best configuration from calibration is described in Appendix C.3. The total error (difference of positive tests between simulation and real data) is 18, i.e. 82% accuracy compared to the final COVID-19 tests. We acknowledge that the interval for calibration (0.25, 0.5, 0.75, 1.0) is significantly large. Future sensitive analysis can help identify the correct weight parameters that most affect decision-making to perform the calibration in smaller intervals.

8.7 Experiment with behavioural determinants

TABLE 8.5: Experiment design

Main determinant(s)	w_attitude	w_social	w_affect	w_facilitating	w_intention	w_habit
Attitude (At)	as calibrated	0	0	0	as calibrated	0
Social Factors (SC)	0	as calibrated	0	0	as calibrated	0
Affect (Af)	0	0	as calibrated	0	as calibrated	0
At + SF	as calibrated	as calibrated	0	0	as calibrated	0
SC + Af	0	as calibrated	as calibrated	0	as calibrated	0
St + Af	as calibrated	0	as calibrated	0	as calibrated	0
Facilitating Conditions (FC)	as calibrated	as calibrated	as calibrated	as calibrated	0	0
Intention (I)	as calibrated	as calibrated	as calibrated	0	as calibrated	0
Habit (H)	as calibrated	as calibrated	as calibrated	0	0	as calibrated
FC + I	as calibrated	as calibrated	as calibrated	as calibrated	as calibrated	0
I + H	as calibrated	as calibrated	as calibrated	0	as calibrated	as calibrated
FC + H	as calibrated	as calibrated	as calibrated	as calibrated	0	as calibrated

The experiment in this case study is set up similarly to the first case study (see Section 5.7). It aims to assess whether we can draw the same connection between the decision-making determinants and their effects at the macro level. In turn, it evaluates the impact of replacing empirical data with another theory - Maslow's hierarchy of needs.

8.7.1 Design

This experiment focuses on observing the impact of core determinants in TIB, i.e. attitude, social factors, affect, facilitating condition, intention, and habit. This can be achieved by adjusting the corresponding weights in the models, i.e. $w(Attitude)$, $w(Social)$, $w(Affect)$, $w(Facilitating)$, $w(Intention)$, $w(Habit)$ (see Table 8.5). This exercise is similar to the experiment in the mobility demand case study (see Section 5.7). In this case, the experiment is performed on the calibrated deterministic population described in Section 8.6; in which mode, agents choose the best alternative for the action to be done during one day. By keeping the weight(s) of the main determinant(s) as calibrated values and others to 0, the agent will only take into account that key determinant(s) in decision-making and ignore the rest. In the first half of this setup, we focus on the second level of TIB, which connects to *Intention* to the third level. Hence, $U(Intention)$ is kept as calibrated in Section 8.6. This process is applied similarly to $w(Attitude)$, $w(Social)$, $w(Affect)$ in the second part to ensure $U(Intention)$ remains non-zero.

8.7.2 Result

TABLE 8.6: Result of comparing the second level of TIB's determinants

Main determinant	**Work**	**Shopping**	**Staying in assigned room**	**Eat in kitchen**	**Smoking**	**Socialise**	**Walking**	**Talking with visitors**
Reference population	754	4201	10443	2020	556	2806	3153	4957
Attitude (At)	363	6072	11636	2353	540	2256	3636	4034
Social Factors (SF)	554	4031	8598	3242	823	3032	3153	5457
Affect (Af)	854	5124	13523	604	456	1019	6053	1257
At + SF	687	4721	8926	2532	421	3516	2873	5214
SF + Af	556	4124	11356	2406	424	2519	4053	2257
At + Af	563	5752	10716	1053	652	1463	5623	3068

After the simulation, the number of actions (e.g. going to work, shopping, staying in the assigned room, walking) during the three-month period can be obtained.

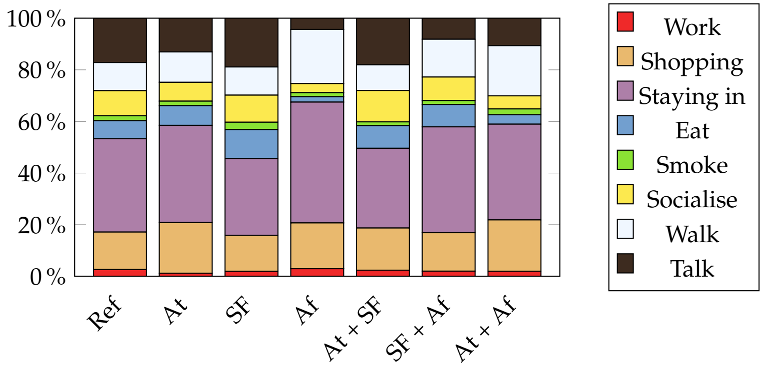


FIGURE 8.8: Percent composition of modes in the tests of second level of TIB's determinants

Comparing reference results in Section 8.6 with the outcomes of each setup in Table 8.5 show whether a determinant can have a significant impact on the agent population's behaviours. The mapping of the TIB's determinant in Figure 8.6 and the percentage composition of the activities in the cumulative figures can then be used to interpret the meaning of the difference in each test.

- **Attitudinal, Affective and Social determinants:** Table 8.6 and Figure 8.8 show the results of running BedDeM with the reference population and with one or two determinants of the second level turned on.

  In the *Attitude(At)* test case, the majority of actions follow the priorities of Maslow's needs. On the one hand, a large number of people *go shopping* and *walking* (which satisfies *Physiology*), or *staying in the assigned room* (which satisfies *Safety*. On the other hand, there are decreases in actions that fulfil *Self-actualization* (e.g. *go to work*) and *Belongingness and love* (e.g. *eating out*, *talking with friends and visitors*) compared to the figures of the reference case. They are caused by the differences in tank volumes in the *Attitude* determinant (see Table 8.2).

  The number of social actions does not reflect the order in Table 8.3. Although there is a decline in the number of migrants *staying in the assigned room* (from 10'443 actions to 8'598 actions), *eating in the kitchen* or *taking with visitors* are not among the top options. Due to the percentage of transmissions of the associated location being significantly higher, more agents got sick quickly and stayed in the assigned room more in a later stage. It can boost the total number of socialising actions when combined with other factors (i.e. At + SF or SF + Af).

  With the main focus on *Affect(Af)* determinant, more agents choose to *stay in the assigned room* due to the fear of being infected. It also discourages other actions that involve meeting with others. For example, *talking with visitors and friends* reduce from 2'806 to 1'019. *Eating in the kitchen* is chosen 1'416 times less than the reference case. In addition, it explains the figures when two determinants are combined. When *Affect* is not considered (i.e. At + SF), *socialising* goes down by 48%. Therefore, we can conclude that *Affect* is the main driver for *staying in the assigned room* while *Social Factors* can strengthen the opposite effects.

Main determinant	**Work**	**Shopping**	**Staying in assigned room**	**Eat in kitchen Smoking**	**Socialise**	**Walking**	**Talking with visitors**	
Reference population	754	4201	10443	2020	556	2806	3153	4957
Facilitating Conditions (FC)	682	4521	9762	2141	631	2942	3354	4857
Intention (I)	689	4452	11112	2122	569	2578	2695	4673
Habit (H)	1053	5024	8427	1832	1140	3052	3146	5216
FC + I	711	4219	11532	1941	479	2717	3077	4214
I + H	899	4901	9432	2242	724	2814	3021	4857
FC + H	763	4872	9016	1942	1002	3067	3203	5025

TABLE 8.7: Result of comparing the third level of TIB's determinants

- **Intentional, Habitual and Facilitating condition determinants:** The results of the third-level determinants' tests can be seen in Table 8.7 and Figure 8.9.

  Although there is a small variation, the percentages of different actions in the *Facilitating Conditions(FC)* test case are similar compared to the reference case. The current setup has *Facilitating Conditions* as agents prefer to continue to stay outside if they already perform an outdoor activity. Otherwise, they would prefer an indoor action. Upon having a closer observation of the calibration result in Appendix Table C.4, the represented agents of these households do not have a significant weight $w(Facilitating)$. It would mean that this particular condition does not significantly contribute to the final decision.

  We expected that migrants prefer to *staying in the assigned room* as a *Habit*. However, its test case also has a lower percentage of this option than the reference. It is due to the lack of habit information questions in the dataset. In addition, the simulation time is short (3 months), so agents could not create a new habit. Moreover, most migrants had just arrived at the centre and were not familiar with the situation.

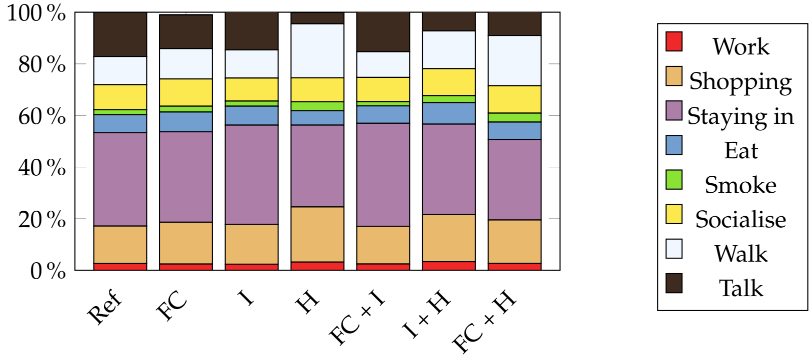


FIGURE 8.9: Percent composition of modes in the tests of third level of TIB's determinants

In contrast, *Intention* emerges as an important factor for *staying in the assigned room* since the final figure of this mode is at least 20% larger than the one of either *Habit* or *Facilitating Conditions*. It can be confirmed in the combination cases where *Intention* is present, i.e. FC + I or I + H. Both have a higher number of people choosing to stay in the assigned room than other scenarios with only *Habit* or *Facilitating Conditions*. In TIB, *Intention* refers to the deliberation process of human decision-making, as opposed to *Habit*, which causes people to act on impulse.

From the experiment above, we can see the effect of the leading determinant(s) on the agent's total weekly actions. However, as their weights are calibrated and different assumptions have to be made for the application of our framework in this case study, it is not possible to draw a conclusion on the actual impacts of these factors in practice.

8.8 Advantages and limitations of our framework

This case study provides an application in a different domain - healthcare for migrants, where various socio-psychological factors can play an important role in preventing the spreading of a pandemic such as COVID19. Compared to the models mentioned in Section 8.2, the tree-like and layered structure of TIB has inspired us to develop a new agent model that can combine many different determinants in human decision-making. In addition, the case study assesses the usage of our framework in a scenario in which socio-psychological is limited.

The amount of statistical data is low compared to the previous three studies due to the nature of the migrant population. To resolve this problem, we adopt a psychological theory - Maslow's hierarchy of needs - into our framework. The *Attitude* determinant in the Theory of Interpersonal Behaviours (TIB) needed to be more clearly defined to apply our utility function. It requires a design of tanks for the different needs and modifications to the original utility. In addition, the mapping of the first level of the TIB model is simplified (see Figure 8.6), and additional assumptions are made to derive a suitable utility value for each action of the agent. This design only aids the process of theoretical exploration and provides rather limited practical insight into a real-world phenomenon.

An experiment with the second and third-level determinants is then proposed to test their effects on different actions in the agent population. Preliminary results show that social factors do not necessarily lead to more activities outside the centre. The emotion of *fear* plays a significant role in making migrants stay in the assigned room. However, as mentioned, the experiment only allows us to draw theoretical explanation on the observations in the previous statistical study [Mor+21]. Further investigations and qualitative approaches are required to understand more finely how living conditions, risks and behaviours such as tobacco consumption and the adoption of protective measures impact COVID-19 infection.

The limitation of the dataset can restrict the practical application of the model. In order to better utilise our framework in future research, it would require careful design for specific focus groups or survey questionnaires protocol to accommodate the number of different socio-psychological factors. Additional important aspects of data collection should also focus on the validity of the model and boosting its reliability and replicability. For further reading on this topic, we recommend different books on data collection, standards, and best practices, such as [Bra+06; Sch+21].

Chapter 9

Discussion

The last four chapters 5 to 8 show the implementation of our framework to develop agent-based models in four case studies in the domain of mobility, trust and reputation, vehicle purchasing and public health. The variety of application areas and contexts allow us to assess the ability of our framework to explain the correlation between socio-psychology factors and macro social patterns.

In this chapter, we first summarise all these case studies and the lessons learned during the application of our framework. This summarisation will be the basis for answering the research questions in the following section.

9.1 Summary of the case studies

Following the steps in Section 4.4, we demonstrated how to adapt the environment and agent's reasoning to different decision-making types and available datasets. In each case study, we first give an introduction of the reasons for model development and a description of background information (Step 1). Related work is then introduced, including the applications of the state-of-the-art in Chapter 3. From the description of the dataset, the agents and environment parametrisation are then performed (Step 2). We also specify how agents interact with each other and with the environment (Step 3). The decision-making cycle with the main components is then described (Step 4). After, the model is calibrated with empirical data (Step 5). Finally, an experiment is provided to evaluate the suitability of implementation or extension of the components of our framework for the case study (Step 6).

The first case study simulates the change in mobility demand, which is a challenge for interdisciplinary researchers in the transportation and energy sector. The number of papers devoted to applying agent-based technologies in the transportation engineering domain has grown enormously. However, there is still a need for modelling platforms capable of exploring the influence of different psychological factors on individual decision-making. By utilising our platform to create a model of Swiss households, we propose an experimental method to test and investigate the impact of core determinants in the TIB on the usage of different transportation modes. Comparing the results with a calibrated population of Swiss households data, we conclude that *Intention* and *Affect* have a positive effect on the usage of private vehicles, while *Habit* and *Social Factors* can encourage people to travel with public or soft transportation modes.

Trust and reputation are currently being researched broadly in multiple disciplines and are often considered the main drivers of human actions in many different scenarios. However, in the agent-based simulation community, there are still concerns about qualifying and modelling them with sufficient details and adequateness in decision-making. Besides, the diversity of application domains requires a method to combine trust and reputation with other determinants to provide a complete picture of the deliberating process in complex systems. The second case study presents a novel solution by applying subjective logic in conjunction with a modelling framework that utilises TIB to simulate the modal choices of households. It uses the concept of opinion as a metric to measure an agent's belief about the consequence(s) of action, which can be updated through feedback loops. In addition, its consensus rule allows us to combine relevant opinions of the neighbours to evaluate the target's reputation. By performing an experiment set up in the mobility domain, we demonstrate the model's ability to capture the ground truth of a service's reputation at different simulation scales and highlight the effects of these concepts on train demand.

The third case study investigates the possibility of simulating bounded rationality effects in an agent's decision-making scheme by limiting its capability of perceiving information and utilising a decision-making framework of TIB. Based on previous work on an agent-based platform, BedDeM, we propose how to capture the effects of sequential, emotional, habitual and multi-criteria decision-making. Considering confirmation bias and the bandwagon effect in terms of the simplified definitions provided in Section 7.1, the Perception component in the agent is further extended with new filtering functions to limit the number of available options. We demonstrate the functionality of this model in the context of purchasing vehicles in Switzerland's households. The model is first calibrated with empirical purchasing data, which becomes a reference. The filtering functions in the Perception are then turned off, allowing all market models to be available for the decision-making process. We show that this effect the final number of models purchased.

In the final study, a model is created to simulate the migrant's choice of activities. Previous studies on this topic were unable to specify a systematic way to compare the effects of determinants in decision-making. Using our framework, we aimed to provide a finer view of different socio-psychological aspects that can affect a minority population - migrants in an asylum centre of Switzerland - in a closed environment. However, due to the lack of relevant data, Maslow's hierarchy of needs and several additional assumptions have to be made in order to provide a coherent narrative for the developed model. Therefore, the experiment in this case study only provides a way to explore the implication of a set of hypotheses rather than practical insights into a real-world phenomenon.

Overall, the design of basic components in our framework enables adaptations for the different contexts of choice modelling. These include transportation modal choice, buying a vehicle or daily activities. Most of the work can be done by changing the first-level determinants in the TIB's model based on the phenomena simulated and the available data. It corresponds to the *LeafDeterminant* component of our framework. When the environment is uncertain or partially observable, one can modify the *Perception* and *Feedback* interface to reflect these observations over time better. Users can also implement the interfaces in our framework, e.g. *InternalState, Option, Task*, to enable the flow of data between the agent's components.

The variety of determinants in TIB's model encourages the modeller to consider more determinants and aspects in social studies. By implementing them in a modular framework, we allow specific modifications to the structure of these determinants. For example, users can turn off a dimension that is not related to a study by assigning some weights to zero. The TIB structure can be replaced by other theories (such as TPB) by customising the *Determinant*s' organisation in the *Decision* component.

There are also some limitations of our implementation in those case studies. First, there is a more extensive data requirement due to the number of determinants. It can also affect the overall time needed for calibration. One effective solution is clustering with the relevant statistical model to identify the common features that can also limit the number of profiles. In addition, presetting parameters using literature/other theory data can be a good substitution for calibration. Our modular framework allows switching or turning off some determinants (by making their weight equal to 0) to reflect other behaviour theories. The modeller can use this feature to test and find a more suitable set of determinants that fit the dataset and the decision-making context.

Our framework is based on a socio-psychological theory and requires collaboration between disciplines to translate and map correct knowledge. To demonstrate its functionality, the mapping in the case studies is simple and intuitive. Different interpretations or contexts should also be explored to allow variability in terms of social research. As mentioned, the framework can also change its structure to reflect other multi-layer theories, such as TIB or TPB, by switching some determinants' weight to zero. They must, however, ensure that the relationship or mechanisms highlighted in the agent-based model are plausible explanations of real-world phenomena, which often involve a separate analysis of empirical data. Similar experiments to assess the causal relationships between decision determinants and behaviour can then be performed.

There are also some promising research directions for our models. They can provide a good indication of the roles of determinants in future scenarios (such as new infrastructures or government policies). Coupling our models with other models from different sectors can also provide a complete picture of all the internal and external factors that can affect human decision-making. In addition, our framework can act as a guide for the data collection procedure. Especially in the design of interviews and questionnaires, the modeller can better represent domain knowledge in the form of the decision-makers' utilities, features, heuristics or decision-making characteristics that he/she can deploy in a model.

9.2 Conclusions of research questions

In this chapter, we summarise the experience gained while creating our agent framework and in the four case studies to answer the research questions.

- **Question 1**: What does an implementation of an agent decision-making framework with TIB offers? Does it help to close the conceptual gap between MAS and ABM?
- **Question 2**: When developing models for different case studies, what are the limitations of our agent decision-making framework with TIB?
- **Question 3**: For which research purposes are the models based on our agent decision making framework especially useful?

9.2.1 Question 1

Building an agent decision-making framework based on TIB has several advantages, including:

- Compared to other behavioural theories, TIB considers various determinants, which can inform a wide range of research designs and methodologies as it is a flexible tool that identifies a set of potentially relevant factors and their interactions (see Section 4.1).
- The TIB determinants are organised in a tree-like structure (see Figure 4.4). Hence, the decision-making based on TIB can be flexible enough to reflect other behavioural theories by exchanging determinants and assigning weights to mark their contribution to the agent's decision-making process.
- It provides the user with a more sophisticated mechanism to compute the probability of performing an action from a set of available alternatives using an *additive value function* 4.7.
- By changing the mapping with the first level determinants, we could adapt the framework for different decision-making contexts, such as mobility choice, vehicle purchasing and migrant's daily activity (see chapters 5 to 7).

In Section 1.1, we identified a conceptual gap between the two fields of ABM and MAS. On the one hand, the most popular and highly-cited method of ABM often employ ad-hoc, simple *condition-action* rules based on theoretical assumptions or derived over statistical distributions. On the other hand, the state-of-the-art in MAS does not cover the variety of behavioural dimensions and different decision-making strategies (see Chapter 3). By satisfying the criteria in Section 1.2, our framework provides a better approach to help to close the conceptual gap mentioned:

- It covers all dimensions that we listed in Section 1.2, including cognition, affective, social factors, norm and learning (details comparison is in Section 4.1). The developed model allows theoretical exploration of the macro effects when the agent's decision-making is led by one or more of these determinants. An example can be seen in the experiment of the first and last case study (Section 5.7 and 8.7).

- It provides a general agent architecture design which can be extended to allow the expression and simulation of many social phenomena. For example, in the trust and reputation case study, we adapted the utility function with subjective logic to represent these aspects in the agent's decision-making component (see Section 6.4). Another example can be seen in the third case study, in which the Perception component is extended to represent the confirmation bias and bandwagon effect (see Section 7.5). By adding or removing these implementations in the experiments, we can observe their effects on the macro simulation results.

- It provides a way to incorporate empirical data and use the variety of psychological ontologies in the simulation process to produce a more transparent explanation for social phenomena. The first three case studies utilise the real-world statistical data in the mapping of determinants and calibrate the models after (see Section 5.3 and 7.3). This process improves the reliability and transparency of the agent's decision-making. Subsequently, it can also facilitate the engagement of psychologists, sociologists, economists and the general public with multiagent modelling projects.

9.2.2 Question 2

In the four chapters 5 to 8, we identified the following three potential limitations from the technical perspective when applying our framework:

- It is challenging to adapt our framework correctly to capture human aspects in terms of qualitative data. As our framework includes qualitative determinants, such as *Emotion* or *Social Factors*, it is difficult to qualify exactly their utility values. It would further require a careful design for specific focus groups or survey studies to qualify the level of trust and reputation more accurately. It is also an ongoing topic in the ABM community [Sei14; Edm15; An+21; RDG21]. In this framework, we propose to rank the available option and use the rankings with their associated weights to create a difference in utilities. Other methods to use qualitative data to inform behaviour rules are also encouraged for future research [PSG10; GDS15].

- Due to the number of determinants in the TIB, there is a significant amount of micro data required for mapping and calibration. This amount can grow significantly with the diversity of agents' profiles that are represented in our agent population. An example of this can be seen in the complexity and required runtime of the calibration process of the third case study (Section 7.7). Reducing the number of agent profiles using statistical and clustering techniques is an immediate solution (e.g. [JJM13; Saa+18; BS19a; GAE20]). Another method that was applied in the third case study is performing sensitivity analysis (see Section 7.7), which involves varying a system's inputs to assess the individual impacts of each variable on the output and ultimately provide information regarding the different effects of each tested variable. The parameters that have more impact on the behaviour output can be further tested in smaller interval sets.

- In the scenario where socio-psychological data is limited, additional theories and assumptions can be made to calculate the utility value for each option. However, this design can only aid the process of theoretical exploration and provides somewhat limited practical insight into a real-world phenomenon. The fourth case study is a typical example, where we adopt a psychological theory - Maslow's hierarchy of needs to calculate the utility of *Attitude* determinant (see Section 8.5). The experiment only allows us to provide a theoretical explanation of the observations in the previous statistical study [Mor+21]. To better utilise our framework in future research, an additional data collection process with a careful designed protocol to accommodate the number of different socio-psychological factors is recommended.

9.2.3 Question 3

One model can have indeed different purposes. In his paper [Edm17a], Edmond looks at seven of them, including prediction, explanation, description, theoretical exposition, illustration, analogy and social learning. Their definitions can be seen in Section 1.3. We consider each of these purposes with the models developed in the four chapters 5 to 8:

- *Prediction*: In the first three case studies, we generate a reference baseline by calibrating with empirical data. By applying policies that target certain determinants of the whole population, we can explore their effects on alternative pathways in the future. An example can be seen in one of our work [Bek+18]. Hence, these models can be used for prediction purpose.

- *Theoretical exposition*: As our framework is based on the TIB, all the models we developed in the case studies serve this purpose since they can be used to explore the consequences of theoretical assumptions and properties through computer simulation.

- *Description*: The models in the first three case studies utilise plausible mechanisms to match outcome data in a well-defined manner, which fits the description purpose category.

- *Explanation*: Our models deliberately made many connections between aspects of the simulation and evidence of various kinds. Some of these connections might be in the form of comparing the outcomes of the simulation to data. In this case, it is tempting to suggest that the simulation supports an explanation of those outcomes. However, the simulation has not been established in a range of settings. We could only test simple mappings from data to the TIB's determinants. Hence, our models are not general enough to make a scientific explanation for the simulated phenomena.

- *Illustration*: A model can be designed to be an illustration or playful exploration as being sufficient for the purpose of a theoretical exposition. However, an illustration does not test the code intensively to check the behaviour and the assumptions. Our frameworks have been applied in several case studies in different domains. Hence, it does not have an *illustration* purpose.

- *Analogy*: We formalise TIB into an agent framework. Therefore, the models cannot be used as a way of thinking about something in an informal manner and so do not serve this purpose.

- *Social learning*: Models developed for this purpose aim to capture a shared understanding (or set of understandings) of a group of people. TIB focus on individual behaviour. Thus, our generated models do not serve this purpose.

Chapter 10

Conclusion

In this chapter, we first summarise the main findings of this work, including the following statements:

- There is a research gap for an agent decision-making framework that covers different concepts in social science research.
- A framework based on the Theory of Interpersonal Behaviours (TIB) with modular determinants has been designed to bridge this gap.
- The feasibility of this framework has been proven by applying it in four case studies.
- The advantages and disadvantages of the framework have been analysed and detailed.

Finally, we present future works and potential directions.

10.1 Findings in the study

Agent-Based Model (ABM) is useful for social research as it introduces the possibility of a new way of thinking about different social and engineering processes based on the idea of the emergence of complex behaviours from relatively simple activities [Sim96]. The advance in research in Multiagent System (MAS) could provide a robust and novel approach to understand societies for researchers in ABM. However, these benefits are limited in practice mainly due to the difference in methodologies and conceptual gaps between MAS and social research. Building agent architecture based on socio-psychological theories is a promising direction to minimise this gap. There are a significantly large number of aspects that can be considered in an agent's decision-making. Tina Balke and Nigel Gilbert have identified five high-level dimensions that should be considered in an agent decision-making architecture [BG14]: *cognition, affect, social factors, norm and learning*. Using them as an initial set, we aim to find an agent architecture and framework that can cover them comprehensively. The objectives of our research include:

- The framework shall have a sophisticated decision-making mechanism, moving away from ad-hoc, oversimplified behavioural rules.
- The framework shall allow the expression of assumptions, postulates and concepts explicitly drawn from social sciences. At the minimum, it should include the following five dimensions: cognition, affective, social factors, norm and learning.
- The framework shall have an extensible mechanism that allows reflection on various decision-making aspects.
- The framework shall offer a mechanism to incorporate empirical data.
- The framework can be applied in different decision-making contexts and domains, e.g. mobility mode choice, health care, and public policy.

First, the theoretical background of agents and their decision-making mechanism has been laid out. An agent architecture needs to include all basic components: a way to percept the current state of the environment, a way to derive the agent's internal state, and a way to communicate the action to the environment and a decision-making process. For a single decision-maker, utility functions are required to compare different alternative outcomes. The framework should be based on an abstract architecture that can organise the flow of information between the decision-making determinants. Finally, it should consider different properties of the environment, such as static/dynamic, know/unknown, fully/partly observable and deterministic/stochastic.

Using our objectives and criteria above, a large body of research has been documented in the literature. We provided a mix of categories of agent architectures and frameworks in Chapter 3, including BDI and its derivatives, normative models, cognitive models, and socio-psychology-inspired frameworks. They were, however, not suitable for our criteria due to the following reasons:

- Each architecture based on BDI only focuses on one particular dimension, such as emotion (eBDI) or norm (BRIDGE, BOID). There is no architecture in this category that can cover all five dimensions.
- Normative agent architectures have the same problem as the architectures of the BDI category as their primary purpose is on how norms are captured in human deliberation.
- Cognitive architectures lack the ability to reflect social realism and do not cover the affective and social dimensions.
- MoHuB only provides a simple way to adapt a standard agent architecture to reflect different theories. It aims to include a different set of behavioural theories into formal models. Comparing these theories with the criteria in Section 1.2, the implementation of them in MoHuB still does not meet all of them sufficiently. In addition, the agent components have to be redefined to implement a new theory, or additional processes are needed. This process

requires a certain level of expertise and effort from the users, which can limit the reusability of the framework.

- Consumat considers different heuristics in terms of contextual design. However, formalising a specific domain requires more effort and deliberate choices than a rational actor approach. If the modelled context becomes complex, it will require significant work to formalise the complete framework, the need-satisfying capacities and resource demands of many different opportunities. In addition, its concept design is specific, so users cannot consider a different setup or interpretation in another context.

Since there was not an architecture or framework that could satisfy our criteria, we decided to build our own agent decision-making framework. Building a framework from a theory of human decision-making that covers a broad set of decision-making determinants is one promising direction for our research objectives. It facilitates the reuse and comparison of models since a theory could serve as a standard reference [Bel+15; CCB08]. Modelling an agent's decision-making based on a theory can also help limit the enormous options of what aspect could be included in the model to only those deemed relevant by the theory [Edm17b]. In addition, this approach can provide a standard practice in interdisciplinary teams and facilitate communication between modellers and social scientists [DEB07].

From our survey of different socio-psychological theories, the Theory of Interpersonal Behaviours (TIB) was chosen due to its broad set of determinants and inclusion of an additive value function. However, it does not cover the *Learning* aspect of the agent's decision. We provided a solution by adding a feedback loop from the environment and by developing a full agent architecture that can combine many different determinants in human decision-making, each of which can also be enhanced by empirical data.

To implement a decision-making framework with TIB, a combination of the horizontal and vertical one-pass layered architectures is suitable as TIB shares a similar layout with these architectures. Determinants from the same group (e.g. norm, role, self-concept) can be put in layers of a horizontal layout, while the different levels of decision-making (i.e. three levels of TIB model) can follow the one-pass architecture. An agent's main components and functions are illustrated in Figure 10.1. A typical decision-making cycle is as follows: When a task is assigned, the *Perception* observes the current state of the *Environment* and combines them with the agent's internal state to produce a list of perceived options. Then, they are given to the *Decision* unit to be evaluated. Details of this process are described in Section 4.2.3. The *Communication* component then utilises this result to execute the chosen option(s) with *Environment* and other agents. The *Environment* can then provide feedback(s) based on the nature of the system associated with the action. The agent remembers these feedbacks in the *Memory*, which can then be used to modify the probability of expected values in future decision-making.

Figure 10.2 illustrates the decision-making steps in our framework. An agent is given a list of tasks that isolated decision-making task needs to be sequentially executed and a list of actions. To perform a task, the agent first filters the list of actions

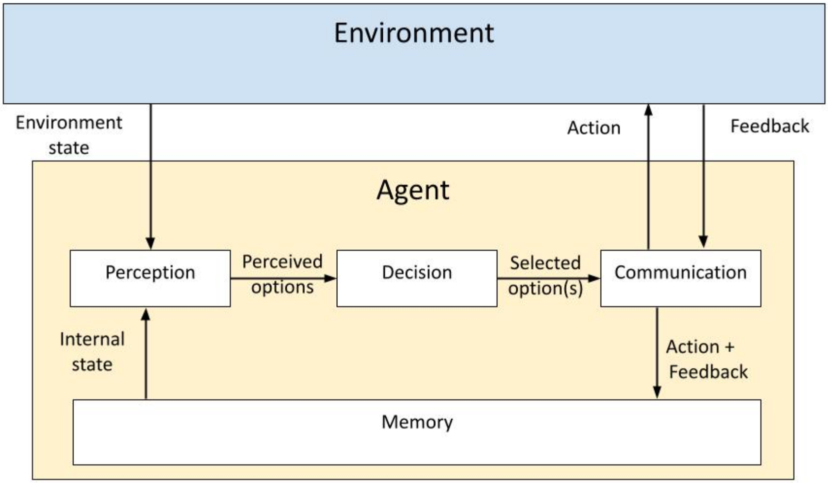


FIGURE 10.1: Overview of agent's basic components

with the information from its internal state and the external environmental state to generate a set of possible options.

For all determinants (d) in TIB, each option (o) is then given an utility value which comes from comparing its properties with other's ($U_o(d)$). In the first level, this value can be in the form of a cardinal utility measure (for determinants such as price or time) or ordinal utility ranks (for determinants such as emotion). Both of them can be calculated from empirical data (e.g. census, survey) or calibrated with experts' knowledge and stakeholders' assessment. The results for these determinants are then multiplied with an normalised weights (called $w(d)$). This process is captured in the following equation, which is adapted from the additive value function in Section 2.1.1:

$$U_o(d) = \sum_{c=1}^{C} U_o(c) * w(c) \tag{10.1}$$

where $U_o(d)$(opt) is the utility value of an option o at determinant d. C is the set of all children c of d, i.e. determinants connect with d in the previous level. Therefore, $c \in C$ where $connect(d, C)$. $w(c)$ is the normalised weight of child determinant c.

Next, an UML diagram was provided to show the framework classes, interfaces and relationship between them. A set of pseudo code for each agent component was then provided. A modeller can adopt this agent's framework by producing a glue code for each component and some important interfaces. The following steps provide a standard approach to use our framework:

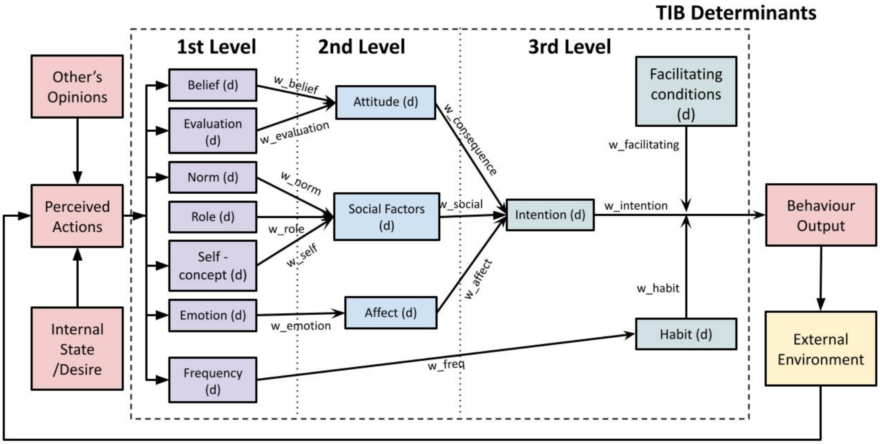


FIGURE 10.2: Agent's decision-making mechanism with TIB's determinant

1. Specify the targeted behaviours or the interesting features/phenomena that can be simulated through the given modelling context.

2. Parametrise agents and environment and their attributes using the available data. This way, all the fields and variables of the main *Environment* and *Agent* classes should be defined.

3. Design how the agents and environment interact, including *Agent-Self*, *Environment-Self*, *Agent-Agent*, *Environment-Environment*, *Agent-Environment* (see Section 2.3). In particular, the modeller can define the variable fields in the *EnvironmentState* and *Feedback* interfaces.

4. Define how an agent filter and evaluate an option. This process involves the implementation of the *InternalState* and *Option* interfaces. Users also need to extend the main components of the agents, i.e. *PerceptionComponent*, *MemoryComponent*, *DecisionComponent* and *CommunicationComponent*, and specify the functions in these classes. In addition, the function *evaluateOptions* in the *LeafDeterminant* should be defined. The default of the computation is set as TIB three-layered model, but a user can create their own decision-making structure that is more suitable for the context and available database.

5. Calibrate the agent parameters so that the outputs reflect empirical/historical data.

6. Set up an experiment to demonstrate or test the model's functionality base on the type of decision-making and the model's purpose. Its results can then be interpreted and discussed.

In the next four chapters 5 to 8, we demonstrated the framework's usage by creating fully-working models for four case studies from Switzerland's mobility, car purchasing and health care domains. In each of them, there was an adaptation of TIB's determinants to a specified context. It was followed by detailed data mapping and calibration processes. Several experiments were then proposed to observe and qualify the macro pattern change when a decision parameter was modified.

In the model of mobility demand study, agents made decisions on transportation mode for their routine, i.e. *short-term* decision-making. The environment was *deterministic*, *fully observable*, *static* and *known*. Trips data from Transport Microcensus (MTMC) and psychological data from Swiss Household Energy Demand Survey (SHEDS) were utilised in the mapping with our framework, as well as in the calibration. With the amount of data available, all the suggested decision-making strategies or aspects mentioned in Section 1.1 could be represented in the model. Our experiment was set up by turning on/off the second and third levels of TIB's determinants to test their impacts on the final behaviour output. We were able to provide insight into the agent decision process and reason about the changes of transportation when agents only on one or two determinants.

In the second study, the same agent's population was utilised. To represent the trust and reputation of train services, we adopted the theory of subjective logic in the reasoning of the agent's *Decision* component. In this case, the environment was *stochastic*, *fully observable*, *static* and *unknown*. We then tested the model functionality by modifying the punctuality of the train at the regional and national levels. The experiments showed that this new mechanism was able to reproduce the ground truth. We also explored the theoretical link between the train's reputation, time and final decision output.

A new model for vehicle purchasing was created in the third case study. Agents were derived from the same profiles as the previous cases, but their attributes were updated. In addition, we modified the *Perception* component to represent the two types of bounded rationality: confirmation bias and the bandwagon effect. The experiment section provided a simple test of the impact of this implementation on the number of vehicles considered and purchased. This case study focused on *long-term*, *strategic* decision-making in a *deterministic*, *partially observable*, *static* and *known* environment.

The fourth case study modelled the daily activities of migrants during a three-month period during the COVID-19 pandemic. The type of decision is similar to the first case, i.e. *short-term* decision-making in a *fully observable* environment. However, the environment was also *stochastic*, *unknown* and *dynamic*. The data set was unsuitable for mapping with all TIB. We utilised another psychology theory, i.e. Maslow's hierarchy of needs, to evaluate the utility of the determinant *Attitude*. Other assumptions had to be made for other core determinants. Due to this design, although we can perform the same experiment as the first case study, the developed model can only serve the theoretical exploration purpose and provides rather limited practical insight into a real-world phenomenon.

Overall, the TIB framework is able to provide a way to incorporate empirical data and make use of the variety of psychological ontologies in the simulation process to produce a more transparent explanation for the decision-making process in the agent. The framework also allows a systematic way to represent different social aspects and strategies in human behaviours. As determinants in TIB can be organised in a tree structure, we can formalise the utility function to capture the effect of changing the value of a small branch in decision-making. With this setup, it can provide a link between micro-macro levels. In turn, it can act as a test bed for policy designers to test the policies that focus on many aspects of the individual, i.e. not only economic incentives but also social and emotional factors. As a result, it can facilitate the engagement of psychologists, sociologists, economists and the general public with multiagent modelling projects.

The main limitation of using the TIB structure is the extensive data requirement due to the number of determinants. By applying a modular design with associated weights, the framework allows some flexibility in mapping from data sources to elements of TIB. Despite many determinants, we have the possibility to choose only the ones that are relevant to the case study using the associated weights. It is demonstrated in Figures 5.7, 6.3, 7.6 and 8.6. By assigning the weight to 0, the framework can be cut back to reflect other theories, such as the Theory of Reasoned Action (TRA) or Theory of Planned Behavior (TPB). Depending on the context, the user can adapt the Decision-making component design to minimise the complexity of data collection.

In addition, the mappings between datasets and the first level TIB's determinants are still simple and intuitive in our generated models. It can be improved by incorporating inputs from the domain experts to derive alternative mappings. Supplementary interviews and questionnaires can also be designed to understand better the association between the theoretical and real-world determinants of human decision-making. Another approach that requires interdisciplinary cooperation is creating hybrid prediction models, which combine methods from both the expert-driven and the data-driven paradigms using machine learning techniques (e.g. [Not+16; Plo+17]).

10.2 Future work

The proposed decision-making framework covers the research gap identified in this thesis. The results of this thesis can, therefore, be an initial prerequisite for many subsequent research projects. We outline some aspects that should and will be investigated in future research.

One limitation of our models is that many elements are static over simulation time, e.g. determinant weights and schedule. As a solution, panel data of the past few years (e.g. SHEDS) can be further investigated to identify the pattern of change over the course of the simulation. Another method is making agents change their weights to prioritise the selected option. In addition, studies in Reinforcement Learning techniques (e.g. [Mni+15]) or Generalized Expected Utility Theory (e.g. [Qui12]) could offer some insights on this topic.

As our models are built on a behavioural theory framework, we need to validate our agent-based models empirically. Bektas et al. proposed to use an unsupervised machine learning algorithm based on cluster analysis of real and artificial individuals to create meso-level behavioural patterns [BPS21]. The algorithm generates a validation score by comparing the balanced composition of real and artificial agents among these clusters. This method was applied in the mobility case studies [BPS21]. It can also be done similarly for the vehicle purchasing study. Due to the relatively smaller population in the COVID-19 case model, cross-validation can be utilised, which involves checking with informants and other domain experts that the behaviour of the agents conforms to that of the individuals or organisational units the agents represent [ME05].

The number of case studies applying the framework needs to be increased to allow for more reliable evaluation results. Therefore, it is necessary to use our framework and compare the modelling effort with other approaches in various domains, and, ideally, the process should be performed by different developers. Therefore, our approach has been designed to be generally applicable. However, in the case that a variety of case studies exists, it could become helpful to analyse which domains are particularly well suited for the application of our framework and if application areas require more specialised or enhanced versions of the process or additional guidance.

Another aspect that could be used in future case studies is the investigation of different theories as a baseline to organise behaviour determinants. As already pointed out, finding the correct structure of determinants and formulating utility functions across different domains and contexts can become a complex task. Therefore, it might be helpful to investigate modelling protocols that facilitate, on the one hand, a formally grounded description of the scenario and, on the other hand, offer additional value during the implementation and can foster an implementation phase.

In this thesis, we emphasised the importance of identifying contextual requirements. They can be used to specify characteristics of the decision-making process and formalise criteria that have to be considered so that clustering mechanisms can be applied. Moreover, they can offer additional value. For example, existing data can be classified according to a fixed set of calibration requirements and, therefore, form a catalogue of clustering mechanisms that can foster the identification of agents' profiles. Thus, there is a requirement to build a new scheme for classifying input data for different application domains.

Capturing qualitative data is one of the challenges of our framework implementation as it is used to evaluate some of our dimensions, such as emotion. It is an ongoing research area within ABM community [Sei14; Edm15; An+21; RDG21]. As a solution, we suggest using ordinal ranking in the utility functions. However, we acknowledge that extracting domain knowledge from experts in a way that can be easily used by an automated agent can be very complex[11]. Even if some protocols are widely accepted and standardised nowadays, there are rarely any techniques for qualifying qualitative data in behavioural rules. Overall, there is still a gap in research

[11] Readers are recommended to find more discussion on the topic in [Hof14]

for generating better general criteria, rules or mathematics formulation to capture these data in the agent's design.

There is also a need to mention the important aspects of data collection, which focus on maintaining the validity of the data and boosting its reliability and replicability. These aspects are not specific or unique to the data collection of human decision-making. One of the promising directions is the *think-aloud method* for data collection, which requires participants to say whatever comes into their minds as they make decisions [RK18]. All verbalisation is transcribed and then analysed. It could allow the researcher to evaluate better the adequacy of his/her experimental design or selected theoretical exposition.

Using our framework as a foundation, several Artificial Intelligence (AI) and MAS techniques can be applied to include authentic and realistic features of human deliberation. For example, one can extend the *Communication* component to include some interaction protocols mentioned in Section 2.3. The *Perception* component and *Feedback* mechanism can be enhanced to take into account other agents' actions. These improvements can help address interesting topics such as negation and cooperation in an environment with limited resources.

Regarding software and documentation, the core agent framework and its implementation - BedDeM - are developed in Java using an agent-based platform called RePast [Nor+13]. Although Repast's documentation is still limited, it is easy to understand and has reduced the learning curve for the development process. RePast is also actively updated for newer Java versions and functionalities. We are using the R language to handle and analyse empirical input data. We have published the core architecture of BedDeM[12] [Git] and plan to provide further complex examples for the decision-making mechanism. It will allow us to have feedback from multiple perspectives to improve the platform for research across different domains.

[12] Link: https://github.com/SiLab-group/beddem_simulator

Appendix A

Summary of Theories Search Results

TABLE A.1: Summary of search results

Theory	Main idea	References
Action Identification Theory	The theory specifies the principles by which people adopt a single act identity for their behaviour and outlines the conditions under which people maintain this act identity or adopt a new one.	[VW14]
Attachment Theory	The theory explains the evolution of that bond, its development, and its implications for human experience and relationships across the life course.	[HS94]
Attribution Theory	The family of attribution theories are concerned with the question of how ordinary people explain human behaviour, e.g. usage of folk psychology, observing regularities and differences in behaviours.	[CN92; Gil98; Mal04]
Balance Theory	The theory claims that unbalanced structures between individuals and objects are associated with an uncomfortable feeling of negative affect, and that this negative feeling leads people to strive for balanced structures.	[Ins84]
Broaden-and-Build Theory of Positive Emotions	The theory was mainly developed to explain why people experience positive emotions.	[Fre98]

Continued on next page

Table A.1 – continued from previous page

Theory	Main idea	References
Cognitive Dissonance Theory	The theory defines cognitive dissonance as the aversive state of arousal that occurs when a person holds two or more cognitions that are inconsistent with each other, which then can be used to explain a variety of ordinary and extraordinary events in our social lives.	[Fes57; AM59]
Correspondent Inference Theory	The theory outlines when it is appropriate to infer that a persons personality corresponds to his/her behaviour.	[Jon90]
Drive Theory	The theory claims that survival, culturally determined or learned drives can motivate people to reduce desires by choosing responses that will most effectively do so.	[Cot+68]
Dual Process Theories	This group of theories describes how people think about information when they make judgments or solve problems. They distinguish two basic ways: intuitive associations and systematic reasoning.	[CT99]
Dynamic Systems Theory	The theory studies the behaviour of systems that exhibit internal states that evolve over time (i.e., internal dynamics) and how these systems interact with exogenously applied input (often referred to as perturbations).	[Kel95; NV98]
Equity Theory	The theory states that people feel most comfortable when they are getting exactly what they deserve from their relationshipsno more and certainly no less.	[WBW76]
Error Management Theory	The theory proposes that the direction of a bias in social judgment is tied to how costly different kinds of errors are.	[HN06]
Escape Theory	The theory refers to the tendency for people to engage in behaviours to avoid an unpleasant psychological reaction.	[Bau90]
Excitation-Transfer Theory	The theory focuses on physiological manifestations of bodily arousal.	[Apt92; LeD98; Zil96]

Continued on next page

Table A.1 – continued from previous page

Theory	Main idea	References
Implicit Personality Theory	The theory refers to a persons notions about which personality characteristics tend to co-occur in people.	[Bor92]
Inoculation Theory	The theory acts as a strategy to protect attitudes from change, to confer resistance to counter attitudinal influences, whether such influences take the form of direct attacks or sustained pressures.	[CP05a]
Interdependence Theory	This theory describes the structural properties that characterize interactions and the implications of such structure for human psychology. Different from most psychological theories, it regards the relationships between people as important as the people themselves.	[RL12]
Learning Theory	The theory discusses two general types of learning: non-associative and associative learning.	[Sch89]
Logical Positivism	The theory states that a event was meaningful only if it could be verified or confirmed through experience.	[Pas43]
Opponent Process Theory	The theory prefers a stimulus that initially inspires displeasure will likely be followed by a pleasurable after-feeling and vice versa. In addition, the after-feeling can become the prevailing emotional experience associated with a particular stimulus event over time.	[Sol80]
Optimal Distinctiveness Theory	The theory is about social identityhow people come to define themselves in terms of their social group memberships.	[Bre91]
Prospect Theory	It assumes that people derive utilities from gain and loss, which are measured relative to some reference points, rather than from the resulting outcome of the decision.	[TK79]
Realistic Group Conflict Theory	The theory describes how perceived competition for limited resources can lead to hostility between groups.	[Jac93]

Continued on next page

Table A.1 – continued from previous page

Theory	Main idea	References
Reasoned Action Theory	The theory suggests that a person's behaviour is determined by their intention to perform the behaviour and that this intention is, in turn, a function of their attitude toward the behaviour and subjective norms.	[Ajz85]
Reductionism	The theory is the idea that one can completely explain the human psyche by breaking it down into several general principles.	[BC04]
Regulatory Focus Theory	The theory addresses the motivations that people have in goal pursuit, particularly as those motivations address achievement of desired states.	[CH97]
Relational Models Theory	The theory describes the four fundamental forms of social relationships: communal sharing, authority ranking, equality matching, and market pricing.	[Fis91; FH05]
Role Theory	The theory examines how roles (i.e. the collection of expectations that accompany a particular social position) influence a wide array of psychological outcomes, including behaviour, attitudes, cognitions, and social interaction.	[EWD00; EK02]
Self-Affirmation Theory	The theory states that people have a fundamental motivation to maintain self-integrity, a perception of themselves as good, virtuous, and able to predict and control important outcomes.	[Cre+05; SC06]
Self-Categorization Theory	The theory assumes that a person might act as a unique personality in one context, but display collective similarities as a group member in another.	[Tur+94; OT04]
Self-Determination Theory	The theory theorises that the type, rather than amount, of motivation is the more important predictor of outcomes. In addition, the type of motivation is determined by the degree of satisfaction of the basic needs.	[DKR99; RD00]

Continued on next page

Table A.1 – continued from previous page

Theory	Main idea	References
Self-Discrepancy Theory	The theory proposes that people's actual self is compared with their self-guides, the kind of person they want or desire to be. When there is a discrepancy between them, people suffer emotionally.	[Hig87; Str89]
Self-Expansion Theory	The theory states that people are motivated to enter relationships in order to enhance the self and increase self-efficacy.	[ANA98]
Self-Perception Theory	The theory proposes that people determine their attitudes and preferences by interpreting the meaning of their own behaviour.	[Bem72]
Self-Verification Theory	The theory asserts that people want others to see them as they see themselves and will take active steps to ensure that others perceive them in ways that confirm their stable self-views.	[SJDLRH94]
Social Exchange Theory	The theory includes a broad social psychological perspective that attempts to explain how human social relationships are formed, maintained, and terminated.	[DW13]
Social Identity Theory	The theory predicts certain intergroup behaviours on the basis of perceived group status differences, the perceived legitimacy and stability of those status differences, and the perceived ability to move from one group to another.	[Bur06]
Social Impact Theory	The theory depicts that the amount of influence a person experiences in group settings depends on power or social status, physical or psychological distance and the number of people in the group exerting the social influence.	[HL98]
Sociobiological Theory	The theory aims to use demographic parameters and the genetic structure of populations to predict patterns of social organization across species.	[Wil00]
Stress Appraisal Theory	The theory refers to the process by which individuals evaluate and cope with a stressful event.	[Laz63]

Continued on next page

Table A.1 – continued from previous page

Theory	Main idea	References
Symbolic Interactionism	The theory assumes that people respond to elements of their environments according to the subjective meanings they attach to those elements	[Hew76]
Temporal Construal Theory	The theory describes the effects of psychological distance object and events on thinking, decision making, and behaviour.	[TL03]
Terror Management Theory	The theory proposes that people strive to sustain the belief they are significant contributors to a meaningful universe to minimize the potential for terror engendered by their awareness of their own mortality.	[GSP97]
Theory of Mind	The theory refers to the capacity to understand other people by surmising what is happening in their mind.	[CS96]
Theory of Reasoned Action	The theory depicts that a person's behaviour is determined by their intention to perform the behaviour and that this intention is, in turn, a function of their attitude toward the behaviour and subjective norms	[FA75]
Theory of Planned Behaviour	The theory links beliefs to behaviour. It has three core components, i.e. attitude, subjective norms, and perceived behavioural control. Together, they shape an individual's behavioural intentions.	[Ajz85]
Theory of Interpersonal Behaviour	The theory proposes that , a function partly of the intention, partly of the habitual responses, and partly of the situational constraints and conditions.	[Tri77]
Threatened Egotism Theory	The theory states that violence is related to a highly favourable view of the self, combined with an ego threat.	[BSB96; BB98]

Appendix B

Java Code Listing of Agent's Core Components

B.1 Agent overview

Listing B.1: Source code of Perception component

```
package framework.agent.core;

import java.util.LinkedList;
import java.util.List;
import java.util.ListIterator;
import java.util.Map;
import java.util.Set;
import java.util.logging.Level;
import java.util.logging.Logger;

import framework.concept.AgentInfo;
import framework.concept.EnvironmentalState;
import framework.concept.Feedback;
import framework.concept.InternalState;
import framework.concept.Option;
import framework.concept.Pair;
import framework.concept.Task;
import framework.environment.Environment;

/**
 * Define all the fields and decision making process of an agent. Here the
      agent
 * perform the step() method which was scheduled by the ContextManager.
 *
 * @author khoa_nguyen
 *
 *
 */
public abstract class TaskExecutionAgent implements IAgent {

  private static Logger LOGGER = Logger.getLogger(TaskExecutionAgent.class.
       getName());

  protected String id;

  // Current agent's location, which contains environmental information.
  protected Environment loc;
```

```java

  // All the agent's scheduled tasks to be performed.
  private List<Task> schedule;

  // Internal components of the agent.
  protected PerceptionComponent perceptionComponent;
  protected MemoryComponent memoryComponent;
  protected DecisionComponent decisionComponent;
  protected CommunicationComponent communicationComponent;

  public TaskExecutionAgent(String id, Environment loc) {
    this.id = id;
    this.loc = loc;
    this.schedule = new LinkedList<Task>();
    this.decisionComponent = createDecisionComponent();
    this.memoryComponent = createMemoryComponent();
    this.communicationComponent = createCommunicationComponent();
    this.perceptionComponent = createPerceptionComponent();
  }

  /**
   * Implementation of perception component.
   *
   */
  protected abstract PerceptionComponent createPerceptionComponent();

  /**
   * Implementation of memory component.
   *
   */
  protected abstract MemoryComponent createMemoryComponent();

  /**
   * Implementation of decision-making component.
   *
   */
  protected abstract DecisionComponent createDecisionComponent();

  /**
   * Implementation of communication component.
   *
   */
  protected abstract CommunicationComponent createCommunicationComponent();

  @Override
  public void step() throws Exception {
    LOGGER.log(Level.FINE, "Agent " + this.id + " is stepping.");
    // Get the next event from schedule.
    Task task = schedule.remove(0);

    EnvironmentalState environmentalState = loc.getEnvironmentalState();
    InternalState internalState = this.memoryComponent.getInternalState();

    List<Option> options = this.perceptionComponent.generateOptions(task,
         environmentalState, internalState);

    Map<Option, List<Pair<Double, Double>>> evaluatedOptions = this.
         decisionComponent.evaluateOptions(options,
        task);

```

```
    Option pickedOption = this.communicationComponent.pickOption(
         evaluatedOptions, internalState);
    //if (options.size() >0 && pickedOption == null)
    //  System.out.println(this.id);
    Feedback feedback = this.loc.getFeedback(task, pickedOption,
         memoryComponent.generateAgentInfo());

    this.memoryComponent.updateInternalState(task, pickedOption, feedback);
  }

  /**
   * Add a task to agent's schedule. The schedule is sorted in the order of
   * executing time of tasks.
   *
   * @param task
   *            The task needed to be add to agent's schedule.
   */
  public void addToSchedule(Task task) {
    ListIterator<Task> scheduleIt = this.schedule.listIterator(0);
    if (!scheduleIt.hasNext()) {
      this.schedule.add(task);
    } else {
      while (scheduleIt.hasNext()) {
        Task taskInList = scheduleIt.next();
        if (taskInList.compareTo(task) > 0) {
          scheduleIt.previous();
          break;
        }
      }
      scheduleIt.add(task);
    }

  }

  /**
   * Get this agent current location information.
   *
   * @param task
   *            The task needed to be add to agent's schedule.
   */
  public Environment getLoc() {
    return this.loc;
  }

  public CommunicationComponent getCommunicationChannel() {
    AgentInfo newAgentInfo = this.memoryComponent.generateAgentInfo();
    this.communicationComponent.updateAgentInfo(newAgentInfo);
    return this.communicationComponent;
  }

  @Override
  public final boolean isThreadable() {
    return true;
  }

  @Override
  public String getID() {
    return this.id;
  }

  @Override
```

```
  public String toString() {
    return "Agent " + this.id;
  }

  @Override
  public boolean equals(Object obj) {
    if (obj == null) {
      return false;
    }
    if (!TaskExecutionAgent.class.isAssignableFrom(obj.getClass())) {
      return false;
    }
    final TaskExecutionAgent other = (TaskExecutionAgent) obj;
    if ((this.id == null) ? (other.id != null) : !this.id.equals(other.id)) {
      return false;
    }
    return true;
  }

  @Override
  public int hashCode() {
    return this.id.hashCode();
  }

}
```

B.2 Perception

Listing B.2: Source code of Perception component

```
package framework.agent.core;

import java.util.List;

import framework.concept.EnvironmentalState;
import framework.concept.InternalState;
import framework.concept.Option;
import framework.concept.Task;

public interface PerceptionComponent {

  /**
   * Function to be implemented that takes current environmental (external) state
   * and internal state and generate the list of options for the tasks. The
   * information/properties storing in EnvironmentalState and InternalState can be
   * included in the implementation of EnvironmentalState and InternalState
   * interfaces respectively.
   *
   * @param environmentalState Information of current state of the environment (to
   *                           be implemented).
   * @param internalState      Information of the internal state of the agent,
   *                           which is stored in memory (to be implemented).
   * @return List of available options to perform the task.
   *
   * @see EnvironmentalState
   * @see InternalState
   *
   */
  List<Option> generateOptions(Task task, EnvironmentalState environmentalState, InternalState internalState);

}
```

B.3 Memory

Listing B.3: Source code of Memory component

```java
package framework.agent.core;

import framework.concept.AgentInfo;
import framework.concept.Feedback;
import framework.concept.InternalState;
import framework.concept.Option;
import framework.concept.Task;

public interface MemoryComponent {
  /**
   * Provide the current internal state of the agent. The information/properties
   * storing in InternalState can be included in the implementation of
   * InternalState Interface.
   *
   * @return The current internal state of the agent.
   */
  InternalState getInternalState();

  /**
   * Update the internal state of the agent using information provided from the
   * feedback.
   *
   * @param task
   *            The task to be performed.
   * @param option
   *            The option chosen by the agent.
   * @param feedback
   *            The information/properties needed to update the internal state.
   *            included in the implementation of Feedback Interface.
   */
  void updateInternalState(Task task, Option option, Feedback feedback);

  AgentInfo generateAgentInfo();

}
```

B.4 Decision-making

Listing B.4: Source code of Decision component

```java
package framework.agent.core;

import java.util.List;
import java.util.Map;
import framework.concept.Option;
import framework.concept.Pair;
import framework.concept.Task;

public interface DecisionComponent {

  /**
   * Evaluate all options in the provided set.
   *
   * @param options
   *            The options provided for evaluation.
   * @param task
   *            The task to be performed.
   * @return @return The map for values to all the options scored that
       value.
   */
  Map<Option, List<Pair<Double, Double>>> evaluateOptions(List<Option>
      options, Task task);


}
```

Listing B.5: Source code of Decision component

```java
package framework.agent.reasoning;

import java.util.List;
import java.util.Map;

import framework.concept.Option;
import framework.concept.OutcomeProps;
import framework.concept.Task;

/**
 * A class represent a standard determinant (psychology) in decision making
     .
 * User needs to define how agent would evaluate a set of option based on
     its
 * internal state and the task at hand.
 *
 * @author khoa_nguyen
 *
 */
public abstract class Determinant {

  private String id;
  private double weight;

  public Determinant(String id, double weight) {
    this.id = id;
    this.weight = weight;
  }

  protected abstract Map<Option, Double> evalOptions(List<Option> options,
      List<OutcomeProps> outcomePropsList, Task task);

  public String getID() {
    return this.id;
  }

  public double getWeight() {
    return this.weight;
  }

  @Override
  public String toString() {
    return this.getID() + ", weight: " + this.weight;

  }
}
```

Listing B.6: Source code of Decision component

```
package framework.agent.reasoning;

import java.util.HashMap;
import java.util.List;
import java.util.Map;

import framework.concept.Option;
import framework.concept.OutcomeProps;
import framework.concept.Task;

/**
 * The base node in decision making model. In which user has to define how
     agent
 * values an option based on its internal state and the task at hand.
 *
 * @author khoa_nguyen
 * @see IAgent
 *
 */
public abstract class LeafDeterminant extends Determinant {

  public LeafDeterminant(String id, double weight) {
    super(id, weight);
  }

  @Override
  public Map<Option,Double> evalOptions(List<Option> options, List<
       OutcomeProps> outcomePropsList, Task task) {
    Map<Option,Double> results = new HashMap<Option,Double>();
    for (int i = 0; i < options.size(); i++) {
      results.put(options.get(i), evalOpt(options.get(i), outcomePropsList.
           get(i), task));
    }
    return results;
  }

  protected abstract double evalOpt(Option opt, OutcomeProps outcomeProps,
       Task task);

}
```

Listing B.7: Source code of Decision component

```
package framework.agent.reasoning;

import java.util.ArrayList;
import java.util.HashMap;
import java.util.List;
import java.util.Map;

import framework.concept.Option;
import framework.concept.OutcomeProps;
import framework.concept.Task;

/**
 * A class represent a parent determinant in decision making model. Its
     ranking
 * function depends on
 *
 * @author khoa_nguyen
 * @see IAgent
 */
public class ParentDeterminant extends Determinant {

  private List<Determinant> children;

  public ParentDeterminant(String id, double weight) {
    super(id, weight);
    this.children = new ArrayList<Determinant>();
  }

  public void addChildDeterminant(Determinant determinant) {
    this.children.add(determinant);
  }

  @Override
  public Map<Option, Double> evalOptions(List<Option> inputOptions, List<
       OutcomeProps> outcomePropsList, Task task) {
    Map<Option, Double> results = new HashMap<Option, Double>();
    for (Determinant child : this.children) {
      if (child == null) {
        for (Option opt : inputOptions) {
          if (!results.containsKey(opt)) {
            results.put(opt, 0.0);
          }
        }
      } else {
        double sumValue = 0;
        Map<Option, Double> childEvaluation = child.evalOptions(
             inputOptions, outcomePropsList, task);
        for (Option opt : childEvaluation.keySet()) {
          sumValue += childEvaluation.get(opt);
        }
        if (sumValue == 0) {
          sumValue = 1;
        }
        for (Option opt : inputOptions) {
          double childValue = childEvaluation.get(opt) * child.getWeight()
               / sumValue;
          if (results.containsKey(opt)) {
            results.put(opt, results.get(opt) + childValue);
          } else {
```

```
                results.put(opt, childValue);
              }
            }
          }
        }
        return results;
    }

}
```

B.5 Communication

Listing B.8: Source code of Communication component

```
package framework.agent.core;

import java.util.List;
import java.util.Map;

import framework.concept.AgentInfo;
import framework.concept.InternalState;
import framework.concept.Option;
import framework.concept.Pair;

public abstract class CommunicationComponent {

  private AgentInfo agentInfo;

  /**
   * Pick an option based on list evaluated outcomes and probability.
   *
   * @param evaluatedOptions
   *            A map between options and their list of utility values and
   *            probabilities.
   * @param internalState
   *            Internal state of the agent. Used to provide past feedback
        from
   *            the same option.
   * @return The selected option.
   */
  protected abstract Option pickOption(Map<Option, List<Pair<Double, Double
       >>> evaluatedOptions, InternalState internalState);

  protected void updateAgentInfo(AgentInfo newAgentInfo) {
    this.agentInfo = newAgentInfo;
  }

  public AgentInfo getAgentInfo() {
    return this.agentInfo;
  }

}
```

Appendix C

Calibration Results for Case Studies

C.1 Case study: Mobility demand

Type	**conf** [a]	**CM**[b]	**BTT**[b]	**WB**[b]	**O**[b]	**err**[b]
Census		72.7	27.5	8.6	3.7	n/a
Determi-nistic	R_{CM} = (1)CM, (2)BTT, (3)WB, (4)O R_{BTT} = (3)CM, (1)BTT, (4)WB, (2)O R_{WB} = (4)CM, (2)BTT, (1)WB, (3)O R_O = (2)CM, (4)BTT, (3)WB, (1)O	73.1	26.7	3.3	4.4	7.3
Stochas-tic	R_{CM} = (1)CM, (2)BTT, (4)WB, (3)O R_{BTT} = (3)CM, (1)BTT, (4)WB, (2)O R_{WB} = (4)CM, (3)BTT, (1)WB, (2)O R_O = (4)CM, (2)BTT, (3)WB, (1)O	46.7	6.0	5.0	4.6	51.9

TABLE C.1: Best configuration for the mobility demand case study

[a] Abbreviation - CM: Car/ Motobike, BTT: Bus/Tram/Train, WB: Walking/Biking, O:Others.
[b] All units are in 10^9 kilometres.

C.2 Case study: Car purchasing

Determinant	Diesel sale	Gasoline sale	EV / Hybrid sale	Total difference	Calibration order
Baseline	460′145	408′345	307′245		
Filter - Engine Type	120′215	104′330	1′506′870	1′843′570	1
Filter - Energy label	155′995	234′410	872′395	1′043′235	12
Filter - Brands	141′470	134′620	955′345	1′240′500	10
Filter - Recommendation	237′775	544′810	432′660	484′250	17
Percepted weight - Energy Label	148′995	150′350	1′126′710	1′388′610	6
Percepted weight - Price	94′600	645′435	934′865	1′230′255	9
Percepted weight - Brands	200′705	151′400	1′103′545	1′312′685	8
Percepted weight - Recommendation	102′880	119′880	1′290′430	1′628′915	3
Percepted weight - New models	1′52′970	97′545	1′163′175	1′473′905	5
$w_{Charging}$	352′705	417′975	772′900	582′725	16
w_{Price}	88′700	131′150	1′360′515	1′701′910	2
$w_{Energylabel}$	251′750	448′675	520′035	461′515	18
$w_{Recommendation}$	523′585	563′485	356′455	267′790	19
$w_{Socialstatus}$	608′215	354′465	951′920	846′625	13
w_{Brands}	375′190	220′980	1′203′075	1′168′150	11
$w_{Emotion}$	224′455	471′030	815′835	806′965	15
w_{Habit}	117′040	156′750	1′086′510	1′373′965	7
$w_{Attitude}$	172′535	111′580	1′295′550	1′572′680	4
$w_{Social factors}$	570′685	332′870	960′430	839′200	14

TABLE C.2: Sensitivity test for car purchasing case study

Determinant	**Profile 1**	**Profile 2**	**Profile 3**	**Profile 4**
Filter - Engine Type	Electric	Gasoline	Diesel	Hydrogen
Filter - Energy label	A	D	C	B
Filter - Brands	7	5	1	4
Filter - Recommendation	0.7	0.5	0.5	0.5
Percepted weight - Energy Label	0.75	0.5	0.5	0.75
Percepted weight - Price	0.25	0.75	0.75	0.5
Percepted weight - Brands	0.5	0.75	0.75	1
Percepted weight - Recommendation	0.25	0.75	0.75	0.5
Percepted weight - New models	0.25	0.5	0.5	0.5
$w_{Charging}$	1	0	0	0
w_{Price}	0.5	0.75	0.75	0.5
$w_{Energylabel}$	1	0.5	0.5	0.25
$w_{Recommendation}$	0.25	0.5	0.75	0.5
$w_{Socialstatus}$	1	0.5	0.5	0.75
w_{Brands}	0.75	0.5	0.5	0.25
$w_{Emotion}$	0.75	0.5	0.25	0.5
w_{Habit}	0.25	0.25	0.25	0.75
$w_{Attitude}$	0.75	0.75	0.75	0.5
$w_{Social factors}$	0.75	0.5	0.5	0.5

TABLE C.3: Best configuration for the vehicle purchasing case study

C.3 Case study: Covid-19

Profile	w_{social}	w_{affect}	w_{habit}	$w_{condition}$
1	1	0.5	0.25	0.25
2	1	0.5	0.5	0.5
3	0.1	1	0.5	0.25
4	0.1	0.5	0.25	0.5
5	1	0.5	0.25	0.25
6	0.5	0.5	0.25	0.5
7	1	0.5	0.25	0.25
8	0.5	0.5	0.5	0.5
9	0.1	0.5	0.25	0.5
10	1	0.5	1	0.25
11	0.5	0.25	0.25	0.75
12	1	1	0.25	0.25

TABLE C.4: Best configuration for the COVID-19 case study

Bibliography

[AD88] Robert Axelrod and Douglas Dion. "The further evolution of cooperation". In: *Science* 242.4884 (1988), pp. 1385–1390.

[AG13] Masih Akhbari and Neil S Grigg. "A framework for an agent-based model to manage water resources conflicts". In: *Water resources management* 27.11 (2013), pp. 4039–4052.

[AG16] Carole Adam and Benoit Gaudou. "BDI agents in social simulations: a survey". In: *The Knowledge Engineering Review* 31.3 (2016), pp. 207238.

[AH81] Robert Axelrod and William D Hamilton. "The evolution of cooperation". In: *Science* 211.4489 (1981), pp. 1390–1396.

[Aia+15] E. Aiassa, J.P.T. Higgins, G. K. Frampton, M. Greiner, A. Afonso, B. Amzal, J. Deeks, L. Dorne, J. Glanville, G. L. Lövei, K. Nienstedt, A. O'Connor, A.S. Pullin, A. Rajic, and D. Verloo. "Applicability and feasibility of systematic review for performing evidence-based risk assessment in food and feed safety". In: *Critical reviews in food science and nutrition* 55.7 (2015), pp. 1026–1034.

[Ajz05] Icek Ajzen. *Attitudes, Personality and Behaviour*. 2nd. McGraw-hill education (UK), 2005.

[Ajz85] Icek Ajzen. "From Intentions to Actions: A Theory of Planned Behavior". In: *Action Control: From Cognition to Behavior*. Ed. by Julius Kuhl and Jürgen Beckmann. Berlin, Heidelberg: Springer Berlin Heidelberg, 1985, pp. 11–39.

[AL98] John R Anderson and Christian J Lebiere. *The atomic components of thought*. 1st. Psychology Press, 1998.

[Ale+20] Alberto Aleta, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E. Dean, M. Elizabeth Halloran, Ira M. Longini Jr, Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro, and Yamir Moreno. "Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19". In: *Nature Human Behaviour* 4.9 (2020), pp. 964–971.

[AM59] Elliot Aronson and Judson Mills. "The effect of severity of initiation on liking for a group." In: *The Journal of Abnormal and Social Psychology* 59.2 (1959), pp. 177–181.

[Amp04] Liz Elizabeth Ampt. "Understanding voluntary travel behaviour change". In: *Transport Engineering in Australia* 9.2 (2004), pp. 53–66.

[An+21] Li An, Volker Grimm, Abigail Sullivan, BL Turner II, Nicolas Malleson, Alison Heppenstall, Christian Vincenot, Derek Robinson, Xinyue Ye, Jianguo Liu, et al. "Challenges, tasks, and opportunities in modeling agent-based complex systems". In: *Ecological Modelling* 457 (2021), p. 109685.

[ANA98] Arthur Aron, Christina C Norman, and Elaine N Aron. "The self-expansion model and motivation." In: *Representative Research in Social Psychology* (1998), pp. 1–13.

[And+04] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. "An integrated theory of the mind." In: *Psychological review* 111.4 (2004), pp. 1036–1060.

[And+07a] Giulia Andrighetto, Marco Campennì, Rosaria Conte, and Mario Paolucci. "On the immergence of norms: a normative agent architecture". In: *AAAI Fall Symposium: Emergent Agents and Socialities*. 2007.

[And+07b] Giulia Andrighetto, Rosaria Conte, Paolo Turrini, and Mario Paolucci. "Emergence In the Loop: Simulating the two way dynamics of norm innovation". In: *Normative Multi-agent Systems*. Ed. by Guido Boella, Leon van der Torre, and Harko Verhagen. Vol. 7122. Dagstuhl Seminar Proceedings (DagSemProc). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2007, pp. 1–30.

[Anw+14] AHMM Anwar, Kiet Tieu, Peter Gibson, Khin Than Win, and Matthew J Berryman. "Analysing the heterogeneity of traveller mode choice preference using a random parameter logit model from the perspective of principal-agent theory". In: *International Journal of Logistics Systems and Management* 17.4 (2014), pp. 447–471.

[Apt92] Michael J Apter. *The dangerous edge: The psychology of excitement*. 1st. Free Press, 1992.

[Axe97] Robert Axelrod. "The complexity of cooperation". In: *The Complexity of Cooperation*. Princeton university press, 1997.

[BA15] Elizabeth Bruch and Jon Atwell. "Agent-based models in empirical social research". In: *Sociological methods & research* 44.2 (2015), pp. 186–221.

[BA98] Michael D Byrne and John R Anderson. "Perception and action". In: *The atomic components of thought* 16 (1998), pp. 23–28.

[BAB99] Moshe Ben-Akiva and Michel Bierlaire. "Discrete choice methods and their applications to short term travel decisions". In: *Handbook of transportation science*. Springer, 1999, pp. 5–33.

[Bal+04] Michael Balmer, Nurhan Cetin, Kai Nagel, and Bryan Raney. "Towards truly agent-based traffic and mobility simulations". In: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*. IEEE Computer Society. 2004, pp. 60–67.

[Bau03] Roy Baumeister. "The psychology of irrationality: Why people make foolish, self-defeating choices". In: *The psychology of economic decisions* 1 (2003), pp. 3–16.

[Bau90] Roy F Baumeister. "Suicide as escape from self." In: *Psychological review* 97.1 (1990), pp. 90–113.

[BB98] Brad J Bushman and Roy F Baumeister. "Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence?" In: *Journal of personality and social psychology* 75.1 (1998), pp. 219–229.

[BC04] Gary G Berntson and John T Cacioppo. "Multilevel Analyses and Reductionism: Why Social Psychologists Should Care about". In: *Essays in social neuroscience* (2004), pp. 107–120.

[BDF16] Lars Böcker, Martin Dijst, and Jan Faber. "Weather, transport mode choices and emotional travel experiences". In: *Transportation Research Part A: Policy and Practice* 94 (2016), pp. 360–373.

[BDT05] Jan Broersen, Mehdi Dastani, and Leendert van der Torre. "Beliefs, obligations, intentions, and desires as components in an agent architecture". In: *International Journal of Intelligent Systems* 20.9 (2005), pp. 893–919.

[Bek+18] Alperen Bektas, Khoa Nguyen, Valentino Piana, and René Schumann. "People-centric policies for decarbonization: Testing psycho-socio-economic approaches by an agent-based model of heterogeneous mobility demand". In: *CEF Conference*. 2018.

[Bel+15] Andrew R Bell, Derek T Robinson, Ammar Malik, and Snigdha Dewal. "Modular ABM development for improved dissemination and training". In: *Environmental Modelling & Software* 73 (2015), pp. 189–200.

[Bem72] Daryl J Bem. "Self-perception theory". In: *Advances in experimental social psychology*. Vol. 6. Elsevier, 1972, pp. 1–62.

[Ben+19] Gueltoum Bendiab, Stavros Shiaeles, Samia Boucherkha, and Bogdan Ghita. "FCMDT: A novel fuzzy cognitive maps dynamic trust model for cloud federated identity management". In: *computers & security* 86 (2019), pp. 270–290.

[BFS10] Claudia R Binder, Giuseppe Feola, and Julia K Steinberger. "Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture". In: *Environmental impact assessment review* 30.2 (2010), pp. 71–81.

[BG14] Tina Balke and Nigel Gilbert. "How do agents make decisions? A survey". In: *Journal of Artificial Societies and Social Simulation* 17.4 (2014), p. 13.

[BIP88] Michael E Bratman, David J Israel, and Martha E Pollack. "Plans and resource-bounded practical reasoning". In: *Computational intelligence* 4.3 (1988), pp. 349–355.

[Bir01] Andreas Birk. "Learning to trust". In: *Trust in Cyber-societies*. Springer, 2001, pp. 133–144.

[BK86] Raj K Bhatnagar and Laveen N Kanal. "Handling uncertain information: a review of numeric and non-numeric methods". In: *Machine Intelligence and Pattern Recognition*. Ed. by Laveen N. Kanal and John F. Lemmer. Vol. 4. Elsevier, 1986, pp. 3–26.

[Bla+20] Tony Blakely, Jason Thompson, Natalie Carvalho, Laxman Bablani, Nick Wilson, and Mark Stevenson. "The probability of the 6-week lockdown in Victoria (commencing 9 July 2020) achieving elimination of community transmission of SARS-CoV-2". In: *Medical Journal of Australia* 213.8 (2020), pp. 349–351.

[Bor92] Peter Borkenau. "Implicit personality theory and the five-factor model". In: *Journal of personality* 60.2 (1992), pp. 295–327.

[Bou+21] Konstantinos Boulouchos, Christian Bach, Christian Bauer, Dominik Bucher, Davide Cerruti, Amin Dehdarian, Massimo Filippini, Maximilian Held, Stefan Hirschberg, Ramachandran Kannan, Tom Kober, Albert Mancera Sugrañes, Valerio De Martinis, Véronique Michaud, Kirsten Oswald, Martin Raubal, Kyle Seymour, and Andrea Vezzini. *Pathways to a net zero CO2 Swiss mobility system: SCCER Mobility Whitepaper*. Tech. rep. Zurich: ETH Zurich, 2021.

[BPS21] Alperen Bektas, Valentino Piana, and René Schumann. "A meso-level empirical validation approach for agent-based computational economic models drawing on micro-data: a use case with a mobility mode-choice model". In: *SN Business & Economics* 1.6 (2021), pp. 1–25.

[Bra+06] Giovanna Brancato, Stefania Macchia, Manuela Murgia, Marina Signore, Giorgia Simeoni, Karen Blanke, and Jürgen Hoffmeyer-Zlotnik. "Handbook of recommended practices for questionnaire development and testing in the European statistical system". In: *European Statistical System* (2006).

[Bra+18] Diego De Siqueira Braga, Marco Niemann, Bernd Hellingrath, and Fernando Buarque De Lima Neto. "Survey on computational trust and reputation models". In: *ACM Computing Surveys (CSUR)* 51.5 (2018), pp. 1–40.

[Bra87] Michael Bratman. *Intention, Plans, and Practical Reason*. Cambridge: Cambridge, MA: Harvard University Press, 1987.

[Bre91] Marilynn B Brewer. “The social self: On being the same and different at the same time”. In: *Personality and social psychology bulletin* 17.5 (1991), pp. 475–482.

[Bro+01] Jan Broersen, Mehdi Dastani, Joris Hulstijn, Zisheng Huang, and Leendert van der Torre. “The BOID architecture: conflicts between beliefs, obligations, intentions and desires”. In: *Proceedings of the fifth international conference on Autonomous agents*. ACM. 2001, pp. 9–16.

[Bro+02] J. Broersen, M. Dastani, J. Hulstijn, and L. van der Torre. “Goal generation in the BOID architecture”. In: *Cognitive Science Quarterly* 2.3-4 (2002), pp. 428–447.

[Bro+04] Daniel G Brown, Scott E Page, Rick Riolo, and William Rand. “Agent-based and analytical modeling to evaluate the effectiveness of greenbelts”. In: *Environmental Modelling & Software* 19.12 (2004), pp. 1097–1109.

[BS02] Sviatoslav Braynov and Tuomas Sandholm. “Contracting with uncertain level of trust”. In: *Computational Intelligence* 18.4 (2002), pp. 501–514.

[BS14] Rahmatollah Beheshti and Gita Sukthankar. “A Normative Agent-Based Model for Predicting Smoking Cessation Trends”. In: *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*. AAMAS ’14. International Foundation for Autonomous Agents and Multiagent Systems, 2014, pp. 557–564.

[BS15] Federico Bianchi and Flaminio Squazzoni. “Agent-Based Models in Sociology”. In: *WIREs Comput. Stat.* 7.4 (July 2015), pp. 284–306.

[BS19a] Alperen Bektas and René Schumann. “How to optimize gower distance weights for the k-medoids clustering algorithm to obtain mobility profiles of the swiss population”. In: *2019 6th Swiss Conference on Data Science (SDS)*. IEEE. 2019, pp. 51–56.

[BS19b] Alperen Bektas and René Schumann. “Using mobility profiles for synthetic population generation”. In: *Conference of the European Social Simulation Association*. Springer. 2019, pp. 207–219.

[BSB96] Roy F Baumeister, Laura Smart, and Joseph M Boden. “Relation of threatened egotism to violence and aggression: the dark side of high self-esteem.” In: *Psychological review* 103.1 (1996), p. 5.

[BT06] David A Bennett and Wenwu Tang. “Modelling adaptive, spatially aware, and mobile agents: Elk migration in Yellowstone”. In: *International Journal of Geographical Information Science* 20.9 (2006), pp. 1039–1066.

[BTV06] Guido Boella, Leon van der Torre, and Harko Verhagen. “Introduction to Normative Multiagent Systems”. In: *Computational & Mathematical Organization Theory* 12 (Jan. 2006), pp. 71–79.

[Bur06] Peter James Burke. *Contemporary social psychological theories*. 1st. Stanford University Press, 2006.

[BV03] Lisa Brouwers and Harko Verhagen. "Applying the Consumat model to flood management policies". In: *Proceedings of Agent-based Simulations*. Apr. 2003, pp. 29–34.

[BVDT07] Guido Boella and Leendert Van Der Torre. "A game-theoretic approach to normative multi-agent systems". In: *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2007.

[BWK99] Ana Lécia C Bazzan, Joachim Wahle, and Franziska Klügl. "Agents in traffic modelling - from reactive to social behaviour". In: *Annual Conference on Artificial Intelligence*. Springer. 1999, pp. 303–306.

[Byr00] Michael D Byrne. "The ACT-R/PM project". In: *Simulating Human Agents: Papers from the 2000 Fall Symposium*. 2000, pp. 1–3.

[BZW12] Walter Brenner, Rdiger Zarnekow, and Hartmut Wittig. *Intelligent Software Agents: Foundations and Applications*. 1st. Springer Science & Business Media, 2012.

[Cal+18] Davide Calvaresi, Valerio Mattioli, Alevtina Dubovitskaya, Aldo Franco Dragoni, and Michael Schumacher. "Reputation management in multi-agent systems using permissioned blockchain technology". In: *2018 IEEE/WIC/ACM international conference on web intelligence (WI)*. IEEE. 2018, pp. 719–725.

[Cal+21] Roberta Calegari, Giovanni Ciatto, Viviana Mascardi, and Andrea Omicini. "Logic-based technologies for multi-agent systems: a systematic literature review". In: *Autonomous Agents and Multi-Agent Systems* 35.1 (2021), pp. 1–67.

[Cas+20] Juana Castro, Stefan Drews, Filippos Exadaktylos, Joël Foramitti, Franziska Klein, Théo Konc, Ivan Savin, and Jeroen van den Bergh. "A review of agent-based modeling of climate-energy policy". In: *Wiley Interdisciplinary Reviews: Climate Change* 11.4 (2020), e647.

[Cas+99] Cristiano Castelfranchi, Frank Dignum, Catholijn M Jonker, and Jan Treur. "Deliberative normative agents: Principles and architecture". In: *International workshop on agent theories, architectures, and languages*. Springer. 1999, pp. 364–378.

[CBA08] Francesco Ciari, Michael Balmer, and Kay W Axhausen. "A new mode choice model for a multi-agent transport simulation". In: *Arbeitsberichte Verkehrs-und Raumplanung* 529 (2008).

[CC99] R Conte and C Castelfranchi. "From conventions to prescriptions. Towards a unified theory of norms". In: *AI and Law* 7 (1999), pp. 323–340.

[CCA15] Jin-Hee Cho, Kevin Chan, and Sibel Adali. "A survey on trust modeling". In: *ACM Computing Surveys (CSUR)* 48.2 (2015), pp. 1–40.

[CCB08] Andrew Crooks, Christian Castle, and Michael Batty. “Key challenges in agent-based modelling for geo-spatial simulation”. In: *Computers, Environment and Urban Systems* 32.6 (2008), pp. 417–430.

[CCD99] Rosaria Conte, Cristiano Castelfranchi, and Frank Dignum. “Autonomous Norm Acceptance”. In: *Intelligent Agents V: Agents Theories, Architectures, and Languages*. Ed. by Jörg P. Müller, Anand S. Rao, and Munindar P. Singh. Vol. 1555. Springer Berlin Heidelberg, 1999, pp. 99–112.

[CF21] Rafael C Cardoso and Angelo Ferrando. “A review of agent-based programming for multi-agent systems”. In: *Computers* 10.2 (2021), p. 16.

[CFP03] Cristiano Castelfranchi, Rino Falcone, and Giovanni Pezzulo. “Integrating trustfulness and decision using fuzzy cognitive maps”. In: *International Conference on Trust Management*. Springer. 2003, pp. 195–210.

[CG99] Kathleen M Carley and Les Gasser. “Computational organization theory”. In: *Multiagent systems: A modern approach to distributed artificial intelligence* (1999), pp. 299–330.

[CH12] Andrew T Crooks and Alison J Heppenstall. “Introduction to agent-based modelling”. In: *Agent-based models of geographical systems*. Springer, 2012, pp. 85–105.

[CH97] Ellen Crowe and E Tory Higgins. “Regulatory focus and strategic inclinations: Promotion and prevention in decision-making”. In: *Organizational behavior and human decision processes* 69.2 (1997), pp. 117–132.

[Cha+20] Dennis L Chao, Assaf P Oron, Devabhaktuni Srikrishna, and Michael Famulare. “Modeling layered non-pharmaceutical interventions against SARS-CoV-2 in the United States with Corvid”. In: *medRxiv* (2020).

[Chh+19] Kinzang Chhogyal, Abhaya C. Nayak, Aditya Ghose, and Hoa Khanh Dam. “A Value-based Trust Assessment Model for Multi-agent Systems”. In: *CoRR* abs/1905.13380 (2019).

[Cho+16] Gerardo Chowell, Lisa Sattenspiel, Shweta Bansal, and Cécile Viboud. “Mathematical models to characterize early epidemic growth: A review”. In: *Physics of life reviews* 18 (2016), pp. 66–97.

[Chu+09] Thanh-Quang Chu, Alexis Drogoul, Alain Boucher, and Jean-Daniel Zucker. “Interactive learning of independent experts’ criteria for rescue simulations.” In: *Journal of Universal Computer Science* 15.13 (2009), pp. 2701–2725.

[CKR91] Robert B Cialdini, Carl A Kallgren, and Raymond R Reno. “A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior”. In: *Advances in experimental social psychology*. Vol. 24. Elsevier, 1991, pp. 201–234.

[CL18] Dongkyu Choi and Pat Langley. "Evolution of the ICARUS cognitive architecture". In: *Cognitive Systems Research* 48 (2018), pp. 25–38.

[CN92] Patricia W Cheng and Laura R Novick. "Covariation in natural causal induction." In: *Psychological Review* 99.2 (1992), pp. 365–382.

[Con+12] R. Conte, N. Gilbert, G. Bonelli, C. Cioffi-Revilla, G. Deffuant, J. Kertesz, V. Loreto, S. Moat, J. P. Nadal, A. Sanchez, A. Nowak, A. Flache, M. San Miguel, and D. Helbing. "Manifesto of computational social science". In: *The European Physical Journal Special Topics* 214.1 (2012), pp. 325–346.

[Con+21] SM Constantino, Maja Schlüter, EU Weber, and N Wijermans. "Cognition and behavior in context: a framework and theories to explain natural resource use decisions in social-ecological systems". In: *Sustainability Science* 16.5 (2021), pp. 1651–1671.

[Coo+10] Philip Cooley, Bruce Y Lee, Shawn Brown, James Cajka, Bernad ette Chasteen, Laxminarayana Ganapathi, James H Stark, William D Wheaton, Diane K Wagener, and Donald S Burke. "Protecting health care workers: a pandemic simulation based on Allegheny County". In: *Influenza and other respiratory viruses* 4.2 (2010), pp. 61–72.

[Cot+68] Nickolas B Cottrell, Dennis L Wack, Gary J Sekerak, and Robert H Rittle. "Social facilitation of dominant responses by the presence of an audience and the mere presence of others." In: *Journal of personality and social psychology* 9.3 (1968), pp. 245–250.

[CP05a] Joshua A Compton and Michael Pfau. "Inoculation theory of resistance to influence at maturity: Recent progress in theory development and application and suggestions for future research". In: *Annals of the International Communication Association* 29.1 (2005), pp. 97–146.

[CP05b] Jacomo Corbo and David Parkes. "The price of selfish behavior in bilateral network formation". In: *Proceedings of the twenty-fourth annual ACM symposium on Principles of distributed computing*. 2005, pp. 99–107.

[CP14] Rosaria Conte and Mario Paolucci. "On agent-based modeling and computational social science". In: *Frontiers in psychology* 5 (2014), p. 668.

[CRA01] Michael D Cohen, Rick L Riolo, and Robert Axelrod. "The role of social structure in the maintenance of cooperative regimes". In: *Rationality and Society* 13.1 (2001), pp. 5–32.

[Cra88] Iain D Craig. "Blackboard systems". In: *Artificial Intelligence Review* 2.2 (1988), pp. 103–118.

[Cre+05] J David Creswell, William T Welch, Shelley E Taylor, David K Sherman, Tara L Gruenewald, and Traci Mann. "Affirmation of personal values buffers neuroendocrine and psychological stress responses". In: *Psychological science* 16.11 (2005), pp. 846–851.

[Cri13] Natalia Criado. "Using norms to control open multi-agent systems". In: *AI Communications* 26.3 (2013), pp. 317–318.

[CRK90] Robert B Cialdini, Raymond R Reno, and Carl A Kallgren. "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places." In: *Journal of personality and social psychology* 58.6 (1990), pp. 1015–1026.

[CS11] Amit K Chopra and Munindar P Singh. "Agent communication". In: *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence,* (2011), pp. 101–141.

[CS96] Peter Carruthers and Peter K Smith. *Theories of theories of mind*. Cambridge university press, 1996.

[CT99] Shelly Chaiken and Yaacov Trope. *Dual-process theories in social psychology*. Guilford Press, 1999.

[CTN07] Hui-Qing Chong, Ah-Hwee Tan, and Gee-Wah Ng. "Integrated cognitive architectures: a survey". In: *Artificial Intelligence Review* 28.2 (2007), pp. 103–130.

[CW04] Alpha C Chiang and Kevin Wainwright. *Fundamental methods of mathematical economics*. 4th. McGraw-Hill, 2004.

[Dan+07] Vincent Danos, Jérôme Feret, Walter Fontana, Russell Harmer, and Jean Krivine. "Rule-based modelling of cellular signalling". In: *International conference on concurrency theory*. Springer. 2007, pp. 17–41.

[Das14] Mehdi Dastani. "A survey of multi-agent programming languages and frameworks". In: *Agent-Oriented Software Engineering: Reflections on Architectures, Methodologies, Languages, and Frameworks*. Ed. by Onn Shehory and Arnon Sturm. Springer. 2014, pp. 213–233.

[Dav+15] Rachel Davis, Rona Campbell, Zoe Hildon, Lorna Hobbs, and Susan Michie. "Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review". In: *Health psychology review* 9.3 (2015), pp. 323–344.

[DD19] Donald L DeAngelis and Stephanie G Diaz. "Decision-making in agent-based modeling: A current review and future prospectus". In: *Frontiers in Ecology and Evolution* 6 (2019), p. 237.

[DDJ08] Frank Dignum, Virginia Dignum, and Catholijn M Jonker. "Towards agents for policy making". In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer. 2008, pp. 141–153.

[de +20] Simon de Lusignan, Jienchi Dorward, Ana Correa, Nicholas Jones, Oluwafunmi Akinyemi, Gayatri Amirthalingam, Nick Andrews, Rachel Byford, Gavin Dabrera, Alex Elliot, Joanna Ellis, Filipa Ferreira, Jamie Lopez Bernal, Cecilia Okusi, Mary Ramsay, Julian Sherlock, Gillian Smith, John Williams, Gary Howsam, Maria Zambon, Mark Joy, and F D Richard Hobbs. "Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study". In: *The Lancet Infectious Diseases* 20.9 (2020), pp. 1034–1042.

[DEB07] Jason P Davis, Kathleen M Eisenhardt, and Christopher B Bingham. "Developing theory through simulation methods". In: *Academy of Management Review* 32.2 (2007), pp. 480–499.

[DETL16] Emily Corrigan Doyle, M. Carolina Escobar-Tello, and Kathy Pui Ying Lo. "Alternative pathways to understanding and designing for happiness in the home". In: *Iterations* 4 (2016), pp. 16–23.

[DHH10] Jan De Houwer and Dirk Hermans. *Cognition and emotion: Reviews of current research and theories*. Psychology Press, 2010.

[DHM10] Mehdi Dastani, Koen V Hindriks, and John-Jules Meyer. *Specification and verification of multi-agent systems*. Springer New York, NY, 2010.

[Die10] Cindy Dietrich. "Decision making: Factors that influence decision making, heuristics used, and decision outcomes". In: *Inquiries Journal* 2.02 (2010).

[Dig04] Virginia Dignum. "A model for organizational interaction: based on agents, founded in logic". PhD thesis. Utrecht University, 2004.

[Dig17] Virginia Dignum. "Social agents: bridging simulation and engineering". In: *Communications of the ACM* 60.11 (2017), pp. 32–34.

[Dig+20] Frank Dignum, Virginia Dignum, Paul Davidsson, Amineh Ghorbani, Mijke van der Hurk, Maarten Jensen, Christian Kammler, Fabian Lorig, Luis Gustavo Ludescher, Alexander Melchior, René Mellema, Cezara Pastrav, Loïs Vanhee, and Harko Verhagen. "Analysing the combined health, social and economic impacts of the corovanvirus pandemic using agent-based social simulation". In: *Minds and Machines* 30.2 (2020), pp. 177–194.

[DKJ18] Ali Dorri, Salil S Kanhere, and Raja Jurdak. "Multi-agent systems: A survey". In: *IEEE Access* 6 (2018), pp. 28573–28593.

[DKR99] Edward L Deci, Richard Koestner, and Richard M Ryan. "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation." In: *Psychological bulletin* 125.6 (1999), pp. 627–668.

[DMJ06] Guillaume Deffuant, Scott Moss, and Wander Jager. "Dialogues concerning a (possibly) new science". In: *Journal of Artificial Societies and Social Simulation* 9.1 (2006).

[Dru01] James N Druckman. "Evaluating framing effects". In: *Journal of economic psychology* 22.1 (2001), pp. 91–101.

[DS98] Norman R Draper and Harry Smith. *Applied regression analysis*. 3rd. Vol. 326. John Wiley & Sons, 1998.

[DT04] Mehdi Dastani and Leendert van der Torre. "Programming BOID-plan agents: Deliberating about conflicts among defeasible mental attitudes and plans". In: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2*. IEEE Computer Society. 2004, pp. 706–713.

[Dug13] Julie Dugdale. "Human behaviour modelling in complex socio-technical systems : an agent based approach". PhD thesis. Université Joseph-Fourier - Grenoble I, Dec. 2013.

[Dur+17] Sylvain Duranton, Agnès Audier, Joël Hazan, Mads P Langhorn, and Vincent Gauche. "The 2017 European Railway Performance Index". In: *The Boston Consulting Group* (2017).

[DW13] John D DeLamater and Amanda Ward, eds. *Handbook of social psychology*. 2nd. Springer, 2013.

[EA96] Joshua M Epstein and Robert Axtell. *Growing artificial societies: social science from the bottom up*. 1st. Brookings Institution Press, 1996.

[EC93] Alice H Eagly and Shelly Chaiken. *The psychology of attitudes.* 1st. Harcourt brace Jovanovich college publishers, 1993.

[Edm15] Bruce Edmonds. "Using qualitative evidence to inform the specification of agent-based models". In: *Journal of Artificial Societies and Social Simulation* 18.1 (2015), p. 18.

[Edm17a] Bruce Edmonds. "Different modelling purposes". In: *Simulating social complexity*. Springer, 2017, pp. 39–58.

[Edm17b] Bruce Edmonds. "The post-truth drift in social simulation". In: *Proceedings of Social Simulation* (2017).

[EK02] Alice H Eagly and Steven J Karau. "Role congruity theory of prejudice toward female leaders." In: *Psychological review* 109.3 (2002), pp. 573–598.

[Ell62] Albert Ellis. *Reason and emotion in psychotherapy.* Lyle Stuart, 1962.

[EPH11] Joshua M Epstein, Ramesh Pankajakshan, and Ross A Hammond. "Combining computational fluid dynamics and agent-based modeling: A new approach to evacuation planning". In: *PloS one* 6.5 (2011), e20139.

[Eps01] Joshua M Epstein. "Learning to be thoughtless: Social norms and individual computation". In: *Computational economics* 18.1 (2001), pp. 9–24.

[Esc+16] Daniel J Escudero, Mark N Lurie, Kenneth H Mayer, Caleb Weinreb, Maximilian King, Sandro Galea, Samuel R Friedman, and Brandon DL Marshall. "Acute HIV infection transmission among people who inject drugs in a mature epidemic setting". In: *AIDS (London, England)* 30.16 (2016), pp. 2537–2544.

[EWD00] Alice H Eagly, Wendy Wood, and Amanda B Diekman. “Social role theory of sex differences and similarities: A current appraisal”. In: *The developmental social psychology of gender* 12 (2000), pp. 123–174.

[FA75] Martin Fishbein and Icek Ajzen. *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley, 1975.

[FB10] Giuseppe Feola and Claudia R Binder. “Towards an improved understanding of farmers’ behaviour: The integrative agent-centred (IAC) framework”. In: *Ecological economics* 69.12 (2010), pp. 2323–2333.

[FC01] Rino Falcone and Cristiano Castelfranchi. “Social trust: A cognitive approach”. In: *Trust and deception in virtual societies*. Springer, 2001, pp. 55–90.

[Fer+05] Neil M Ferguson, Derek AT Cummings, Simon Cauchemez, Christophe Fraser, Steven Riley, Aronrag Meeyai, Sopon Iamsirithaworn, and Donald S Burke. “Strategies for containing an emerging influenza pandemic in Southeast Asia”. In: *Nature* 437.7056 (2005), pp. 209–214.

[Fer+20] Neil Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, et al. “Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand”. In: *Imperial College London* 10.77482 (2020), pp. 491–497.

[Fer92] Innes A Ferguson. *TouringMachines: An architecture for dynamic, rational, mobile agents*. Tech. rep. University of Cambridge, Computer Laboratory, 1992.

[Fes57] Leon Festinger. *A theory of cognitive dissonance*. Vol. 2. Stanford university press, 1957.

[FH05] Alan Page Fiske and Nick Haslam. “The Four Basic Social Bonds: Structures for Coordinating Interaction.” In: *Interpersonal Cognition* (2005), pp. 267–298.

[Fis91] Alan Page Fiske. *Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing*. Free Press, 1991.

[FMP96] Klaus Fischer, Jörg P Müller, and Markus Pischel. “AGenDAA general testbed for distributed artificial intelligence applications”. In: *Foundations of distributed artificial intelligence*. 1996, pp. 401–427.

[For+19] Giancarlo Fortino, Fabrizio Messina, Domenico Rosaci, and Giuseppe ML Sarné. “Using blockchain in a reputation-based model for grouping agents in the Internet of Things”. In: *IEEE Transactions on Engineering Management* 67.4 (2019), pp. 1231–1243.

[FR14] Keith Frankish and William M Ramsey. *The Cambridge handbook of artificial intelligence*. Cambridge University Press, 2014.

[Fre98] Barbara L Fredrickson. “What good are positive emotions?” In: *Review of general psychology* 2.3 (1998), pp. 300–319.

[FST01] “Experiments in building experiential trust in a society of objective-trust based agents”. In: *Trust in Cyber-societies*. Ed. by Rino Falcone, Munindar Singh, and Yao-Hua Tan. Springer, 2001, pp. 111–132.

[GAD06] Davide Grossi, Huib Aldewereld, and Frank Dignum. “Ubi lex, ibi poena: Designing norm enforcement in e-institutions”. In: *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*. Ed. by Pablo Noriega, Javier Vázquez-Salceda, Guido Boella, Olivier Boissier, Virginia Dignum, Nicoletta Fornara, and Eric Matson. Springer. 2006, pp. 101–114.

[GAE20] Lutiele M Godois, Diana F Adamatti, and Leonardo R Emmendorfer. “A multi-agent-based algorithm for data clustering”. In: *Progress in Artificial Intelligence* 9.4 (2020), pp. 305–313.

[Gam+00] Diego Gambetta et al. “Can we trust trust”. In: *Trust: Making and breaking cooperative relations* 13 (2000), pp. 213–237.

[GDS15] Amineh Ghorbani, Gerard Dijkema, and Noortje Schrauwen. “Structuring qualitative data for agent-based modelling”. In: *Journal of Artificial Societies and Social Simulation* 18.1 (2015), p. 2.

[Geo+98] Michael Georgeff, Barney Pell, Martha Pollack, Milind Tambe, and Michael Wooldridge. “The belief-desire-intention model of agency”. In: *International workshop on agent theories, architectures, and languages*. Springer. 1998, pp. 1–10.

[GHFSS04] Alejandro Guerra-Hernández, Amal El Fallah-Seghrouchni, and Henry Soldano. “Learning in BDI multi-agent systems”. In: *Computational Logic in Multi-Agent Systems*. Ed. by Jürgen Dix and João Leite. Springer Berlin Heidelberg, 2004, pp. 218–233.

[GI89] Michael Georgeff and Felix Ingrand. “Decision-making in an embedded reasoning system”. In: *International joint conference on artificial intelligence*. 1989.

[Gil04] Nigel Gilbert. “Agent-based social simulation: dealing with complexity”. In: *The Complex Systems Network of Excellence* 9.25 (2004), pp. 1–14.

[Gil09] Itzhak Gilboa. *Theory of decision under uncertainty*. Vol. 45. Econometric Society Monographs. Cambridge University Press, 2009.

[Gil19] Nigel Gilbert. *Agent-based models*. 2nd. Vol. 153. Quantitative Applications in the Social Sciences. Sage Publications, 2019.

[Gil98] Daniel T Gilbert. “Ordinary personology”. In: *The handbook of social psychology* 2 (1998), pp. 89–150.

[Git] *BedDeM source code*. https://github.com/SiLab-group/$beddem_simulator$. [Last accessed 12-Sept-2019].

[GKO05] Paul H Garthwaite, Joseph B Kadane, and Anthony O'Hagan. "Statistical methods for eliciting probability distributions". In: *Journal of the American statistical Association* 100.470 (2005), pp. 680–701.

[GN10] Miguel A Gual and Richard B Norgaard. "Bridging ecological and social systems coevolution: A review and proposal". In: *Ecological economics* 69.4 (2010), pp. 707–717.

[Gow71] John C Gower. "A general coefficient of similarity and some of its properties". In: *Biometrics* (1971), pp. 857–871.

[GPL03] Nicholas Mark Gotts, J Gareth Polhill, and Alistair NR Law. "Aspiration levels in a land use simulation". In: *Cybernetics & Systems* 34.8 (2003), pp. 663–683.

[GPT17] Volker Grimm, Gary Polhill, and Julia Touza. "Documenting social simulation models: the ODD protocol as a standard". In: *Simulating Social Complexity - A Handbook, Second Edition*. Ed. by Bruce Edmonds and Ruth Meyer. Springer, 2017, pp. 349–365.

[Gra08] Ann M Graybiel. "Habits, rituals, and the evaluative brain". In: *Annual review of neuroscience* 31 (2008), pp. 359–387.

[Gra+15] Jones Granatyr, Vanderson Botelho, Otto Robert Lessing, Edson Emílio Scalabrin, Jean-Paul Barthès, and Fabrício Enembreck. "Trust and reputation models for multiagent systems". In: *ACM Computing Surveys (CSUR)* 48.2 (Oct. 2015), pp. 1–42.

[Gra+17] Jones Granatyr, Nardine Osman, João Dias, Maria Augusta Silveira Netto Nunes, Judith Masthoff, Fabrício Enembreck, Otto Robert Lessing, Carles Sierra, Ana Maria Paiva, and Edson Emílio Scalabrin. "The need for affective trust applied to trust and reputation models". In: *ACM Computing Surveys (CSUR)* 50.4 (2017), pp. 1–36.

[Gre+13] John J. Grefenstette, Shawn T. Brown, Roni Rosenfeld, Jay DePasse, Nathan TB Stone, Phillip C. Cooley, William D. Wheaton, Alona Fyshe, David D. Galloway, Anuroop Sriram, Hasan Guclu, Thomas Abraham, and Donald S. Burke. "FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations". In: *BMC public health* 13.1 (2013), pp. 1–14.

[Gre+97] Shaw Green, Leon Hurst, Brenda Nangle, Pádraig Cunningham, Fergal Somers, and Richard Evans. *Software agents: A review*. Tech. rep. Trinity College Dublin, 1997.

[Gro+17] J. Groeneveld, B. Müller, C.M. Buchmann, G. Dressler, C. Guo, N. Hase, F. Hoffmann, F. John, C. Klassert, T. Lauf, V. Liebelt, H. Nolzen, N. Pannicke, J. Schulze, H. Weise, and N. Schwarz. "Theoretical foundations of human decision-making in agent-based land use models - A review". In: *Environmental Modelling & Software* 87 (2017), pp. 39–48.

[GSP97] Jeff Greenberg, Sheldon Solomon, and Tom Pyszczynski. "Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements". In: *Advances in experimental social psychology*. Vol. 29. Elsevier, 1997, pp. 61–139.

[GT00] Nigel Gilbert and Pietro Terna. "How to build and use agent-based models in social science". In: *Mind & Society* 1.1 (2000), pp. 57–72.

[HÅ05] Peter Hedström and Yvonne Åberg. "Quantitative research, agent-based modelling and theories of the social". In: *Dissecting the social* (2005), pp. 114–144.

[Hah+20] Ulf JJ Hahnel, Gilles Chatelain, Beatrice Conte, Valentino Piana, and Tobias Brosch. "Mental accounting mechanisms in energy decision-making and behaviour". In: *Nature Energy* 5.12 (2020), pp. 952–958.

[Hal+02] M Elizabeth Halloran, Ira M Longini, Azhar Nizam, and Yang Yang. "Containing bioterrorist smallpox". In: *Science* 298.5597 (2002), pp. 1428–1432.

[HB11a] Dirk Helbing and Stefano Balietti. "From social data mining to forecasting socio-economic crises". In: *The European Physical Journal Special Topics* 195.1 (2011), pp. 3–68.

[HB11b] Dirk Helbing and Stefano Balietti. "How to Do Agent-Based Simulations in the Future: From Modeling Social Mechanisms to Emergent Phenomena and Interactive Systems Design". In: *Urban Economics & Regional Studies eJournal* (2011).

[HB99] John Heywood and Barbara Bush. "From pride and prejudice to persuasion satisficing in mate search". In: *Simple heuristics that make us smart* (1999), pp. 287–308.

[Hel+20] Joel Hellewell, Sam Abbott, Amy Gimma, Nikos I Bosse, Christopher I Jarvis, Timothy W Russell, James D Munday, Adam J Kucharski, W John Edmunds, Sebastian Funk, and Rosalind M Eggo. "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts". In: *The Lancet Global Health* 8.4 (2020), e488–e496.

[Her+17] Andreas Herzig, Emiliano Lorini, Laurent Perrussel, and Zhanhao Xiao. "BDI logics for BDI architectures: old problems, new perspectives". In: *KI-Künstliche Intelligenz* 31.1 (2017), pp. 73–83.

[Hew76] John P Hewitt. *Self and Society: A Symbolic Interactionist Social Psychology*. Allyn and Bacon, 1976.

[Hew77] Carl Hewitt. "Viewing control structures as patterns of passing messages". In: *Artificial intelligence* 8.3 (1977), pp. 323–364.

[HG03] David A Hensher and William H Greene. "The mixed logit model: the state of practice". In: *Transportation* 30.2 (2003), pp. 133–176.

[HH88] Steven J Henkind and Malcolm C Harrison. "An analysis of four uncertainty calculi". In: *IEEE Transactions on Systems, Man, and Cybernetics* 18.5 (1988), pp. 700–714.

[Hig+19] Julian PT Higgins, James Thomas, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew J Page, and Vivian A Welch. *Cochrane handbook for systematic reviews of interventions*. 2nd. John Wiley & Sons, 2019.

[Hig87] E Tory Higgins. "Self-discrepancy: a theory relating self and affect." In: *Psychological review* 94.3 (1987), pp. 319–340.

[HK04] Geoffrey M Hodgson and Thorbjørn Knudsen. "The complex evolution of a simple traffic convention: the functions and implications of habit". In: *Journal of Economic Behavior & Organization* 54.1 (2004), pp. 19–47.

[HL98] Stephen G Harkins and Bibb Latané. "Population and political participation: A social impact analysis of voter responsibility." In: *Group Dynamics: Theory, Research, and Practice* 2.3 (1998), pp. 192–207.

[HMNK17] Elizabeth Hunter, Brian Mac Namee, and John D Kelleher. "A taxonomy for agent-based models in human infectious disease epidemiology". In: *Journal of Artificial Societies and Social Simulation* 20.3 (2017), p. 2.

[HMS09] Peter de Haan, Michel G Mueller, and Roland W Scholz. "How much do incentives affect car purchase? Agent-based microsimulation of consumer choice of new cars - Part II: Forecasting effects of feebates based on energy-efficiency". In: *Energy Policy* 37.3 (2009), pp. 1083–1094.

[HN06] Martie Haselton and Daniel Nettle. "The Paranoid Optimist: An Integrative Evolutionary Model of Cognitive Biases". In: *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc* 10 (Feb. 2006), pp. 47–66.

[Hoe+20] Nicolas Hoertel, Martin Blachier, Carlos Blanco, Mark Olfson, Marc Massetti, Marina Sánchez Rico, Frédéric Limosin, and Henri Leleu. "A stochastic agent-based model of the SARS-CoV-2 epidemic in France". In: *Nature medicine* 26.9 (2020), pp. 1417–1421.

[Hof14] Robert R Hoffman. *The psychology of expertise: Cognitive research and empirical AI*. Psychology Press, 2014.

[HS94] Cindy Hazan and Phillip R Shaver. "Attachment as an organizational framework for research on close relationships". In: *Psychological inquiry* 5.1 (1994), pp. 1–22.

[Hub+18] Robert Huber, Martha Bakker, Alfons Balmann, Thomas Berger, Mike Bithell, Calum Brown, Adrienne Grêt-Regamey, Hang Xiong, Quang Bao Le, Gabriele Mack, et al. "Representation of decision-making in European agricultural agent-based models". In: *Agricultural Systems* 167 (2018), pp. 143–160.

[Hub+22] Robert Huber, Hang Xiong, Kevin Keller, and Robert Finger. "Bridging behavioural factors and standard bio-economic modelling in an agent-based modelling framework". In: *Journal of Agricultural Economics* 73.1 (2022), pp. 35–63.

[Ins84] Chester A Insko. "Balance theory, the Jordan paradigm, and the Wiest tetrahedron". In: *Advances in experimental social psychology*. Ed. by Leonard Berkowitz. Vol. 18. Advances in Experimental Social Psychology. Elsevier, 1984, pp. 89–140.

[Jac05] Tim Jackson. "Motivating sustainable consumption". In: *Sustainable Development Research Network* 29.1 (2005), pp. 30–40.

[Jac93] Jay W Jackson. "Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature". In: *The Psychological Record* 43.3 (1993), p. 395.

[Jag00] Wander Jager. "Modelling consumer behaviour". PhD thesis. University of Groningen, 2000.

[JJ02] Wander Jager and Marco Janssen. "The need for and development of behaviourally realistic agents". In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Ed. by Jaime Simão Sichman, Françcois Bousquet, and Paul Davidsson. Springer Berlin Heidelberg, 2002, pp. 36–49.

[JJ03] Marco A Janssen and Wander Jager. "Simulating market dynamics: Interactions between consumer psychology and social networks". In: *Artificial Life* 9.4 (2003), pp. 343–356.

[JJ12] Wander Jager and Marco Janssen. "An updated conceptual framework for integrated modeling of human decision making: The Consumat II". In: *Workshop complexity in the Real World*. 2012, pp. 1–18.

[JJB14] Wander Jager, Marco Janssen, and Marija Bockarjova. "Diffusion dynamics of electric cars and adaptive policy: Towards an empirical based simulation". In: *Advances in social simulation*. Ed. by Bogumił Kamiński and Grzegorz Koloch. Springer, 2014, pp. 259–270.

[JJM13] Yousef Jamali, Tahereh Jamali, and Mohammad RK Mofrad. "An agent based model of integrin clustering: Exploring the role of ligand clustering, integrin homo-oligomerization, integrin - ligand affinity, membrane crowdedness and ligand mobility". In: *Journal of Computational Physics* 244 (2013), pp. 264–278.

[JJV01] Wander Jager, Marco A Janssen, and Charles Viek. “Experimentation with household dynamics: the consumat approach”. In: *International Journal of Sustainable Development* 4.1 (2001), pp. 90–100.

[JJV99] Wander Jager, Marcus Alexander Janssen, and Charles AJ Vlek. “Consumats in a common dilemma. Testing the behavioural rules of simulated consumers”. In: *COV* COV99-01 (1999).

[Jon90] Edward E Jones. *Interpersonal perception.* WH Freeman/Times Books/Henry Holt & Co, 1990.

[Jøs02] Audun Jøsang. “The consensus operator for combining beliefs”. In: *Artificial Intelligence* 141.1-2 (2002), pp. 157–170.

[Jøs16] Audun Jøsang. *Subjective logic*. Springer, 2016.

[Jøs97] Audun Jøsang. “Artificial reasoning with subjective logic”. In: *Proceedings of the second Australian workshop on commonsense reasoning*. Vol. 48. 1997, p. 34.

[JV06] Hong Jiang and Jose M Vidal. “From rational to emotional agents”. In: *Proceedings of the AAAI Workshop on Cognitive Modeling and Agent-based Social Simulation*. 2006.

[Kah79] Daniel Kahneman. “Prospect theory: An analysis of decisions under risk”. In: *Econometrica* 47 (1979), pp. 263–292.

[Kam19] Subbarao Kambhampati. “Synthesizing Explainable Behavior for Human-AI Collaboration”. In: *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems. 2019, pp. 1–2.

[Kan14] A Kangur. “Simulating the transition to electric cars using the consumat agent rationale”. PhD thesis. Groningen, The Netherlands: University of Groningen, 2014.

[KAR18] Dimitris Kremmydas, Ioannis N Athanasiadis, and Stelios Rozakis. “A review of agent based modeling for agricultural policy evaluation”. In: *Agricultural systems* 164 (2018), pp. 95–106.

[Kau00] Vincent Kaufmann. “Modal Practices: From the rationales behind car & public transport use to coherent transport policies. Case studies in France & Switzerland”. In: *World Transport Policy and Practice* 6.3 (2000), pp. 8–17.

[KB15] Kalliopi Kravari and Nick Bassiliades. “A survey of agent platforms”. In: *Journal of Artificial Societies and Social Simulation* 18.1 (2015), p. 11.

[Kel95] JA Scott Kelso. *Dynamic patterns: The self-organization of brain and behavior*. MIT press, 1995.

[KGM01] Manuel Kolp, Paolo Giorgini, and John Mylopoulos. “A goal-based organizational perspective on multi-agent architectures”. In: *International Workshop on Agent Theories, Architectures, and Languages*. Ed. by John-Jules Ch. Meyer and Milind Tambe. Springer. 2001, pp. 128–140.

[Kim+11] Shintae Kim, Keeheon Lee, Jang Kyun Cho, and Chang Ouk Kim. "Agent-based diffusion model for an automobile market with fuzzy TOPSIS-based product adoption process". In: *Expert Systems with Applications* 38.6 (2011), pp. 7270–7276.

[Kla+87] David Klahr, Pat Langley, Robert Neches, and Robert T Neches. *Production system models of learning and development*. MIT press, 1987.

[KN03] Martin J Kollingbaum and Timothy J Norman. "Norm adoption and consistency in the NoA agent architecture". In: *International Workshop on Programming Multi-Agent Systems*. Ed. by Mehdi M. Dastani, Jürgen Dix, and Amal El Fallah-Seghrouchni. Springer Berlin Heidelberg, 2003, pp. 169–186.

[Koc+19] Geo Kocheril, Friedrich Krebs, Lukas Nacken, and Sascha Holzhauer. "Open and Integrative Modelling in Energy System Transitions - Conceptual Discussion About Model Reusability, Framework Requirements and Features". In: *Conference of the European Social Simulation Association*. Ed. by Petra Ahrweiler and Martin Neumann. Springer International Publishing, 2019, pp. 499–505.

[Kol05] Martin Josef Kollingbaum. "Norm-governed practical reasoning agents". PhD thesis. University of Aberdeen, 2005.

[Koo+20] Joel R Koo, Alex R Cook, Minah Park, Yinxiaohe Sun, Haoyang Sun, Jue Tao Lim, Clarence Tam, and Borame L Dickens. "Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study". In: *The Lancet Infectious Diseases* 20.6 (2020), pp. 678–688.

[KR76] Ralph Keeney and Howard Raiffa. *Decisions with multiple consequences: preferences and value tradeoffs*. 1976.

[Kre+20] Mirjam E Kretzschmar, Ganna Rozhnova, Martin CJ Bootsma, Michiel van Boven, Janneke HHM van de Wijgert, and Marc JM Bonten. "Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study". In: *The Lancet Public Health* 5.8 (2020), e452–e459.

[KRM93] Ralph L Keeney, Howard Raiffa, and Richard F Meyer. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.

[KS02] Daniel Kahneman and Vernon Smith. "Foundations of behavioral and experimental economics". In: *Nobel Prize in Economics Documents* 1.7 (2002).

[KS14] Áron Kiss and Gábor Simonovits. "Identifying the bandwagon effect in two-round elections". In: *Public Choice* 160.3 (2014), pp. 327–344.

[KT13] Daniel Kahneman and Amos Tversky. "Choices, values, and frames". In: *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 2013, pp. 269–278.

[Kuc+20] Adam J Kucharski, Petra Klepac, Andrew JK Conlan, Stephen M Kissler, Maria L Tang, Hannah Fry, Julia R Gog, W John Edmunds, Jon C Emery, Graham Medley, et al. "Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study". In: *The Lancet Infectious Diseases* 20.10 (2020), pp. 1151–1160.

[Lai12] John E Laird. *The Soar cognitive architecture*. MIT press, 2012.

[Lan+91] Pat Langley, Kathleen B McKusick, John A Allen, Wayne F Iba, and Kevin Thompson. "A design for the ICARUS architecture". In: *ACM Sigart Bulletin* 2.4 (1991), pp. 104–109.

[Lau+20] Max SY Lau, Bryan Grenfell, Michael Thomas, Michael Bryan, Kristin Nelson, and Ben Lopman. "Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA". In: *Proceedings of the National Academy of Sciences* 117.36 (2020), pp. 22430–22435.

[Laz63] Richard S Lazarus. "A laboratory approach to the dynamics of psychological stress". In: *Administrative science quarterly* (1963), pp. 192–213.

[LC06] Pat Langley and Dongkyu Choi. "A unified cognitive architecture for physical agents". In: *Proceedings of the National Conference on Artificial Intelligence*. Vol. 21. 2. AAAI Press. 2006, pp. 14691474.

[Le+08] Quang Bao Le, Soo Jin Park, Paul LG Vlek, and Armin B Cremers. "Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system. I. Structure and theoretical specification". In: *Ecological Informatics* 3.2 (2008), pp. 135–153.

[LeD98] Joseph LeDoux. *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster, 1998.

[Lee+10] Bruce Y Lee, Shawn T Brown, George W Korch, Philip C Cooley, Richard K Zimmerman, William D Wheaton, Shanta M Zimmer, John J Grefenstette, Rachel R Bailey, Tina-Marie Assi, et al. "A computer simulation of vaccine prioritization, allocation, and rationing during the 2009 H1N1 influenza pandemic". In: *Vaccine* 28.31 (2010), pp. 4875–4879.

[Lel+20] Jos Lelieveld, Frank Helleis, Stephan Borrmann, Yafang Cheng, Frank Drewnick, Gerald Haug, Thomas Klimach, Jean Sciare, Hang Su, and Ulrich Pöschl. "Model calculations of aerosol transmission and infection risk of COVID-19 in indoor environments". In: *International journal of environmental research and public health* 17.21 (2020).

[LL03] F López y López and Michael Luck. "Modelling norms for autonomous agents". In: *Proceedings of the Fourth Mexican International Conference on Computer Science, 2003. ENC 2003.* IEEE. 2003, pp. 238–245.

[LLR06] Jill Fain Lehman, John E. Laird, and Paul S. Rosenbloom. “A gentle introduction to soar, an architecture for human cognition: 2006 update”. In: *University of Michigan* (2006), pp. 1–37.

[LNR87] John E Laird, Allen Newell, and Paul S Rosenbloom. “Soar: An architecture for general intelligence”. In: *Artificial intelligence* 33.1 (1987), pp. 1–64.

[LRN86] John E Laird, Paul S Rosenbloom, and Allen Newell. “Chunking in Soar: The anatomy of a general learning mechanism”. In: *Machine learning* 1.1 (1986), pp. 11–46.

[Mal04] Bertram F Malle. *How the mind explains behavior: Folk explanations, meaning, and social interaction*. MIT press, 2004.

[Man14] Gianluca Manzo. “Potentialities and limitations of agent-based simulations: an introduction”. In: *Revue française de sociologie* 55.4 (2014), pp. 653–688.

[Mar+12] Brandon D. L. Marshall, Magdalena M. Paczkowski, Lars Seemann, Barbara Tempalski, Enrique Rodriguez Pouget, Sandro Galea, and Samuel R. Friedman. “A Complex Systems Approach to Evaluate HIV Prevention in Metropolitan Areas: Preliminary Implications for Combination Intervention Strategies”. In: *PLoS ONE* 7 (2012), pp. 1–16.

[Mar+14] Brandon DL Marshall, Samuel R Friedman, João FG Monteiro, Magdalena Paczkowski, Barbara Tempalski, Enrique R Pouget, Mark N Lurie, and Sandro Galea. “Prevention and treatment produced large decreases in HIV incidence in a model of people who inject drugs”. In: *Health Affairs* 33.3 (2014), pp. 401–409.

[Mas43] Abraham Harold Maslow. “A theory of human motivation”. In: *Psychological review* 50.4 (1943), pp. 370–396.

[MB17] Elham Majd and Vimala Balakrishnan. “A trust model for recommender agent systems”. In: *Soft computing* 21.2 (2017), pp. 417–433.

[MBS01] Rajatish Mukherjee, Bikramjit Banerjee, and Sandip Sen. “Learning mutual trust”. In: *Trust in Cyber-societies*. Springer, 2001, pp. 145–158.

[McF01] Daniel McFadden. “Economic choices”. In: *American economic review* 91.3 (2001), pp. 351–378.

[ME05] Scott Moss and Bruce Edmonds. “Sociology and simulation: Statistical and qualitative cross-validation”. In: *American journal of sociology* 110.4 (2005), pp. 1095–1131.

[Mic] *Comportement de la population en matière de transports*. FR. Neuchâtel, May 2017. URL: https://dam-api.bfs.admin.ch/hub/api/dam/assets/1840478/master.

[Min88] Marvin Minsky. *Society of mind*. Simon and Schuster, 1988.

[Mit13] Raktim Mitra. “Independent mobility and mode choice for school transportation: a review and framework for future research”. In: *Transport reviews* 33.1 (2013), pp. 21–43.

[MLT09] Matthias Meyer, Iris Lorscheid, and Klaus G Troitzsch. “The development of social simulation as reflected in the first ten years of JASSS: a citation and co-citation analysis”. In: *Journal of Artificial Societies and Social Simulation* 12.4 (2009), pp. 1–12.

[MN92] Manfred Max-Neef. *From the outside looking in: Experiences in barefoot economics*. Zed Books, 1992.

[Mni+15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. “Human-level control through deep reinforcement learning”. In: *Nature* 518.7540 (2015), pp. 529–533.

[Mor+21] Kevin Morisod, V Grazioli, V Schlüter, M Bochud, S Gonseth Nusslé, V D’Acremont, A Felappi, N Bühler, and P Bodenmann. “Prevalence of SARS-CoV-2 and associated risk factors among asylum seekers living in centres”. In: *European Journal of Public Health* 31.Supplement_3 (2021), pp. 164–250.

[MPT95] Jörg P Müller, Markus Pischel, and Michael Thiel. “Modeling reactive behaviour in vertically layered agent architectures”. In: *International Workshop on Agent Theories, Architectures, and Languages*. Ed. by Michael J. Wooldridge and Nicholas R. Jennings. Springer Berlin Heidelberg, 1995, pp. 261–276.

[Mur02] Megan Murray. “Determinants of cluster distribution in the molecular epidemiology of tuberculosis”. In: *Proceedings of the National Academy of Sciences* 99.3 (2002), pp. 1538–1543.

[MW02] Michael W Macy and Robert Willer. “From factors to actors: Computational sociology and agent-based modeling”. In: *Annual review of sociology* (2002), pp. 143–166.

[Nap18] Mauro Napoletano. “A short walk on the wild side: Agent-based models and their implications for macroeconomic analysis”. In: *Revue de l’OFCE* 157.3 (2018), pp. 257–281.

[NB18] Tung Doan Nguyen and Quan Bai. “A dynamic Bayesian network approach for agent group trust evaluation”. In: *Computers in human behavior* 89 (2018), pp. 237–245.

[Neu08] Martin Neumann. “Homo socionicus: a case study of simulation models of norms”. In: *Journal of Artificial Societies and Social Simulation* 11.4 (2008), pp. 1–6.

[New73] Allen Newell. "Production systems: Models of control structures". In: *Visual information processing*. Ed. by William G. Chase. Academic Press, 1973, pp. 463–526.

[New94] Allen Newell. *Unified theories of cognition*. Harvard University Press, 1994.

[Nic98] Raymond S Nickerson. "Confirmation bias: A ubiquitous phenomenon in many guises". In: *Review of general psychology* 2.2 (1998), pp. 175–220.

[NK09] Sunil Nautiyal and Harald Kaechele. "Natural resource management in a protected area of the Indian Himalayas: a modeling approach for anthropogenic interactions on ecosystem". In: *Environmental monitoring and assessment* 153.1 (2009), pp. 253–271.

[Nor+13] Michael J North, Nicholson T Collier, Jonathan Ozik, Eric R Tatara, Charles M Macal, Mark Bragen, and Pam Sydelko. "Complex adaptive systems modeling with Repast Simphony". In: *Complex adaptive systems modeling* 1.1 (2013), pp. 1–26.

[Not+16] Gali Noti, Effi Levi, Yoav Kolumbus, and Amit Daniely. "Behavior - Based Machine-Learning: A Hybrid Approach for Predicting Human Decision Making". In: *ArXiv* abs/1611.10228 (2016).

[NPS22] Khoa Nguyen, Valentino Piana, and René Schumann. *Simulating Bounded Rationality in Decision-Making: An Agent-Based Choice Modelling of Vehicle Purchasingy*. Accepted and to be published in the Proceedings of the Social Simulation Conference. Sept. 2022.

[NS19a] Khoa Nguyen and René Schumann. "An Exploratory Comparison of Behavioural Determinants in Mobility Modal Choices". In: *Conference of the European Social Simulation Association*. Ed. by Petra Ahrweiler and Martin Neumann. Springer International Publishing, 2019, pp. 569–581.

[NS19b] Khoa Nguyen and René Schumann. "On Developing a More Comprehensive Decision-Making Architecture for Empirical Social Research: Agent-Based Simulation of Mobility Demands in Switzerland". In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Ed. by Mario Paolucci, Jaime Simão Sichman, and Harko Verhagen. Springer International Publishing, 2019, pp. 39–54.

[NS20a] Khoa Nguyen and René Schumann. "A Socio-psychological Approach to Simulate Trust and Reputation in Modal Choices". In: *International Conference on Principles and Practice of Multi-Agent Systems*. Ed. by Takahiro Uchiya, Quan Bai, and Iván Marsá Maestre. Springer International Publishing, 2020, pp. 34–50.

[NS20b] Khoa Nguyen and René Schumann. "A socio-psychological modal choice approach to modelling mobility and energy demand for electric vehicles". In: *Energy Informatics* 3.1 (2020), pp. 1–18.

[NV98] Andrzej Nowak and Robin R Vallacher. *Dynamical social psychology*. Vol. 647. Guilford Press, 1998.

[NVJ44] Morgenstern O Neumann Von J. "Theory of games and economic behavior". In: *Princeton University Prress, Princeton* (1944).

[OASZ20] María Elena Orduña Alegría, Niels Schütze, and Samuel C Zipper. "A serious board game to analyze socio-ecological dynamics towards collaboration in agriculture". In: *Sustainability* 12.13 (2020).

[Ode+02] James J Odell, H Van Dyke Parunak, Mitch Fleischer, and Sven Brueckner. "Modeling agents and their environment". In: *International Workshop on Agent-Oriented Software Engineering*. Springer. 2002, pp. 16–31.

[OT04] Rina S Onorato and John C Turner. "Fluidity in the self-concept: the shift from personal to social identity". In: *European journal of social psychology* 34.3 (2004), pp. 257–278.

[Pad+14] Lin Padgham, Kai Nagel, Dhirendra Singh, and Qingyu Chen. "Integrating BDI Agents into a MATSim Simulation". In: *Proceedings of the Twenty-First European Conference on Artificial Intelligence*. ECAI'14. IOS Press, 2014, pp. 681–686.

[Par+19] German I Parisi, Ronald Kemker, Jose L Part, Christopher Kanan, and Stefan Wermter. "Continual lifelong learning with neural networks: A review". In: *Neural Networks* 113 (2019), pp. 54–71.

[Pas43] John Arthur Passmore. "Logical positivism (I)". In: *The Australasian Journal of Psychology and Philosophy* 21.2-3 (1943), pp. 65–92.

[PB+97] R Peter Bonasso, R James Firby, Erann Gat, David Kortenkamp, David P Miller, and Mark G Slack. "Experiences with an architecture for intelligent, reactive agents". In: *Journal of Experimental & Theoretical Artificial Intelligence* 9.2-3 (1997), pp. 237–256.

[PD02] Bernard Pavard and Julie Dugdale. "An introduction to complexity in social science". In: *Tutorial on complexity in social science, COSI project* (2002).

[Pea+17] Corey M Peak, Lauren M Childs, Yonatan H Grad, and Caroline O Buckee. "Comparing nonpharmaceutical interventions for containing emerging epidemics". In: *Proceedings of the National Academy of Sciences* 114.15 (2017), pp. 4023–4028.

[Per+05] David Pereira, Eugnio Oliveira, Nelma Moreira, and Lus Sarmento. "Towards an architecture for emotional BDI agents". In: *2005 portuguese conference on artificial intelligence*. IEEE. 2005, pp. 40–46.

[Per+20] T Alex Perkins, Sean M Cavany, Sean M Moore, Rachel J Oidtman, Anita Lerch, and Marya Poterek. "Estimating unobserved SARS-CoV-2 infections in the United States". In: *Proceedings of the National Academy of Sciences* 117.36 (2020), pp. 22597–22602.

[Plo+17] Ori Plonsky, Ido Erev, Tamir Hazan, and Moshe Tennenholtz. "Psychological forest: Predicting human behavior". In: vol. 31. 1. 2017.

[POM07] David Pereira, Eugénio Oliveira, and Nelma Moreira. "Formal modelling of emotions in BDI agents". In: *International Workshop on Computational Logic in Multi-Agent Systems*. Springer. 2007, pp. 62–81.

[Pop+06] Jennie Popay, Helen Roberts, Amanda Sowden, Mark Petticrew, Lisa Arai, Mark Rodgers, Nicky Britten, Katrina Roen, Steven Duffy, et al. "Guidance on the conduct of narrative synthesis in systematic reviews". In: *A product from the ESRC methods programme Version* 1.1 (2006), b92.

[PSG10] J Gary Polhill, Lee-Ann Sutherland, and Nicholas M Gotts. "Using qualitative evidence to enhance an agent-based modelling system for studying land use change". In: *Journal of Artificial Societies and Social Simulation* 13.2 (2010), p. 10.

[PWP05] Toan Phung, Michael Winikoff, and Lin Padgham. "Learning within the BDI framework: An empirical analysis". In: *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. Ed. by Rajiv Khosla, Robert J. Howlett, and Lakhmi C. Jain. Springer Berlin Heidelberg, 2005, pp. 282–288.

[QE20] Lorenzo Quolantoni and Ramon Egger. *Revue automobile*. https://revueautomobile.ch/. [Last accessed 20-Oct-2021]. 2020.

[Qi+18] Haokun Qi, Shengqiang Zhang, Xinzhu Meng, and Huanhe Dong. "Periodic solution and ergodic stationary distribution of two stochastic SIQS epidemic systems". In: *Physica A: Statistical Mechanics and its Applications* 508 (2018), pp. 223–241.

[Qui12] John Quiggin. *Generalized expected utility theory: The rank - dependent model*. Springer Science & Business Media, 2012.

[Rai21] Swiss Federal Railways. *SBB: Facts and Figures 2018*. https://reporting.sbb.ch/_file/733/sbb-cff-ffs-facts-and-figures-2021.pdf. Oct. 2021.

[RD00] Richard M Ryan and Edward L Deci. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." In: *American psychologist* 55.1 (2000), pp. 68–78.

[RDG21] Allison C Reilly, Robin L Dillon, and Seth D Guikema. "Agent-based models as an integrating boundary object for interdisciplinary research". In: *Risk Analysis* 41.7 (2021), pp. 1087–1092.

[Ren+20] Christopher T. Rentsch, Farah Kidwai-Khan, Janet P. Tate, Lesley S. Park, Joseph T. King Jr, Melissa Skanderson, Ronald G. Hauser, Anna Schultze, Christopher I. Jarvis, Mark Holodniy, Vincent Lo Re III, Kathleen M. Akgün, Kristina Crothers, Tamar H. Taddei, Matthew S. Freiberg, and Amy C. Justice. "Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: A nationwide cohort study". In: *PLoS medicine* 17.9 (2020), pp. 1–17.

[Rey87] Craig W Reynolds. “Flocks, herds and schools: A distributed behavioral model”. In: *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*. Vol. 21. 4. Association for Computing Machinery, 1987, pp. 25–34.

[RF20] Max Rollwage and Stephen M. Fleming. “Confirmation bias is adaptive when coupled with efficient metacognition”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 376 (2020).

[RG91] Anand S Rao and Michael P Georgeff. “Modeling rational agents within a BDI-architecture”. In: *KR* 91 (1991), pp. 473–484.

[RG+95] Anand S Rao, Michael P Georgeff, et al. “BDI agents: from theory to practice.” In: *ICMAS*. Vol. 95. 1995, pp. 312–319.

[RHJ04] Sarvapali D Ramchurn, Dong Huynh, and Nicholas R Jennings. “Trust in multi-agent systems”. In: *The Knowledge Engineering Review* 19.1 (2004), pp. 1–25.

[RK18] Ariel Rosenfeld and Sarit Kraus. “Predicting human decision-making: From prediction to action”. In: *Synthesis Lectures on Artificial Intelligence and Machine Learning* 12.1 (2018), pp. 1–150.

[RL12] G.E. Rusbult and Paul Lange. “Interdependence theory”. In: *The Blackwell encyclopedia of social psychology* 2 (Jan. 2012), pp. 329–334.

[RN10] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 3rd. USA: Prentice Hall Press, 2010.

[Roc+20] Rebecca J. Rockett, Alicia Arnott, Connie Lam, Rosemarie Sadsad, Verlaine Timms, Karen-Ann Gray, John-Sebastian Eden, Sheryl Chang, Mailie Gall, Jenny Draper, Eby M. Sim, Nathan L. Bachmann, Ian Carter, Kerri Basile, Roy Byun, Matthew V. OSullivan, Sharon C-A Chen, Susan Maddocks, Tania C. Sorrell, Dominic E. Dwyer, Edward C. Holmes, Jen Kok, Mikhail Prokopenko, and Vitali Sintchenko. “Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling”. In: *Nature medicine* 26.9 (2020), pp. 1398–1404.

[RRF18] Mohammad Raoufi and Aminah Robinson Fayek. “Fuzzy agent-based modeling of construction crew motivation and performance”. In: *Journal of Computing in Civil Engineering* 32.5 (2018), p. 04018035.

[Saa+18] Samaneh Saadat, Chathika Gunaratne, Nisha Baral, Gita Sukthankar, and Ivan Garibay. “Initializing agent-based models with clustering archetypes”. In: *International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation*. Ed. by Robert Thomson, Christopher Dancy, Ayaz Hyder, and Halil Bisgin. Cham: Springer International Publishing, 2018, pp. 233–239.

[San+19] Yanet Sanchez, Teresa Coma, Antonio Aguelo, and Eva Cerezo. "ABC-EBDI: An affective framework for BDI agents". In: *Cognitive Systems Research* 58 (2019), pp. 195–216.

[SB15] Rüdiger Schmitt-Beck. "Bandwagon effect". In: *The international encyclopedia of political communication* (2015), pp. 1–5.

[SC06] David K Sherman and Geoffrey L Cohen. "The psychology of self-defense: Self-affirmation theory". In: *Advances in experimental social psychology* 38 (2006), pp. 183–242.

[Sch+07] Stefan Schmidt, Robert Steele, Tharam S Dillon, and Elizabeth Chang. "Fuzzy trust evaluation and credibility development in multi-agent systems". In: *Applied Soft Computing* 7.2 (2007), pp. 492–505.

[Sch+17] Maja Schlüter, Andres Baeza, Gunnar Dressler, Karin Frank, Jürgen Groeneveld, Wander Jager, Marco A. Janssen, Ryan R.J. McAllister, Birgit Müller, Kirill Orach, Nina Schwarz, and Nanda Wijermans. "A framework for mapping and comparing behavioural theories in models of social-ecological systems". In: *Ecological Economics* 131 (2017), pp. 21–35.

[Sch+20a] Wolfgang Scherr, Patrick Manser, Chetan Joshi, Nathalie Frisch-knecht, and Denis Métrailler. "Towards agent-based travel demand simulation across all mobility choices–the role of balancing preferences and constraints". In: *European Journal of Transport and Infrastructure Research* 20.4 (2020), pp. 152–172.

[Sch+20b] Nina Schwarz, Gunnar Dressler, Karin Frank, Wander Jager, Marco Janssen, Birgit Müller, Maja Schlüter, Nanda Wijermans, and Jürgen Groeneveld. "Formalising theories of human decision-making for agent-based modelling of social-ecological systems: practical lessons learned and ways forward". In: *Socio-Environmental Systems Modelling* 2 (2020), pp. 16340–16340.

[Sch+21] Barry Schouten, Jan van den Brakel, Bart Buelens, Deirdre Giesen, Annemieke Luiten, and Vivian Meertens. *Mixed-Mode Official Surveys: Design and Analysis*. Chapman and Hall/CRC, 2021.

[Sch71] Thomas C Schelling. "Dynamic models of segregation". In: *Journal of mathematical sociology* 1.2 (1971), pp. 143–186.

[Sch89] Barry Schwartz. *Psychology of learning and behavior*. WW Norton & Co, 1989.

[SCM20] Carlos Dornels Freire de Souza, Rodrigo Feliciano do Carmo, and Michael Ferreira Machado. "The burden of COVID-19 in Brazil is greater in areas with high social deprivation". In: *Journal of travel medicine* (2020).

[SD02] Sandip Sen and Partha Sarathi Dutta. “The evolution and stability of cooperative traits”. In: *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 3*. ACM. 2002, pp. 1114–1120.

[SDM+10] Bas R Steunebrink, Mehdi Dastani, John-Jules Ch Meyer, et al. “Emotions to control agent deliberation.” In: *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*. AAMAS ’10. International Foundation for Autonomous Agents and Multiagent Systems, 2010, pp. 973980.

[SDSP06] Sebastian Sardina, Lavindra De Silva, and Lin Padgham. “Hierarchical planning in BDI agent programming languages: A formal approach”. In: *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*. AAMAS ’06. New York, NY, USA: Association for Computing Machinery, 2006, pp. 1001–1008.

[Sei14] Roman Seidl. “Social scientists, qualitative data, and agent-based modeling”. In: *Social Simulation Conference*. 2014.

[SFR00] Michael Schillo, Petra Funk, and Michael Rovatsos. “Using trust for detecting deceitful agents in artificial societies”. In: *Applied Artificial Intelligence* 14.8 (2000), pp. 825–848.

[Sha+13] Ehsan Shafiei, Hlynur Stefansson, Eyjolfur Ingi Asgeirsson, Brynhildur Davidsdottir, and Marco Raberto. “Integrated agent-based and system dynamics modelling for simulation of sustainable mobility”. In: vol. 33. 1. Routledge, 2013, pp. 44–70.

[She] *Swiss Household Energy Demand Survey (SHEDS) website*. https://www.sccer-crest.ch/research/swiss-household-energy-demand-survey-sheds/. [Last accessed 24-Jan-2021]. 2020.

[SHW88] Blair H. Sheppard, Jon Hartwick, and Paul R. Warshaw. “The Theory of Reasoned Action: A Meta-Analysis of Past Research with Recommendations for Modifications and Future Research”. In: *Journal of Consumer Research* 15.3 (Dec. 1988), pp. 325–343.

[Sil17] Samuel David Silvey. *Statistical inference*. Routledge, 2017.

[Sim55] Herbert A Simon. “A behavioral model of rational choice”. In: *The quarterly journal of economics* 69.1 (1955), pp. 99–118.

[Sim78] Herbert A Simon. “Rationality as process and as product of thought”. In: *The American economic review* 68.2 (1978), pp. 1–16.

[Sim87] Herbert A Simon. “Making management decisions: The role of intuition and emotion”. In: *Academy of Management Perspectives* 1.1 (1987), pp. 57–64.

[Sim96] Herbert A. Simon. *The Sciences of the Artificial*. 3rd. Cambridge, MA, USA: MIT Press, 1996. ISBN: 0262691914.

[SJDLRH94] William B Swann Jr, Chris De La Ronde, and J Gregory Hixon. "Authenticity and positivity strivings in marriage and courtship." In: *Journal of personality and social psychology* 66.5 (1994), pp. 857–869.

[SLC19] Yanet Sánchez-López and Eva Cerezo. "Designing emotional BDI agents: good practices and open questions". In: *The Knowledge Engineering Review* 34 (2019), e26.

[SMP98] Ron Sun, Edward Merrill, and Todd Peterson. "A bottom-up model of skill learning". In: *Proc. of 20th cognitive science society conference*. 1998, pp. 1037–1042.

[SN71] Herbert A Simon and Allen Newell. "Human problem solving: The state of the theory in 1970." In: *American Psychologist* 26.2 (1971), p. 145.

[Sob18] Pawel Sobkowicz. "Opinion dynamics model based on cognitive biases of complex agents". In: *Journal of Artificial Societies and Social Simulation* 21.4 (2018), p. 8.

[Sol80] Richard L Solomon. "The opponent-process theory of acquired motivation: the costs of pleasure and the benefits of pain." In: *American psychologist* 35.8 (1980), p. 691.

[Son+05] Shanshan Song, Kai Hwang, Runfang Zhou, and Y-K Kwok. "Trusted P2P transactions with fuzzy reputation aggregation". In: *IEEE Internet computing* 9.6 (2005), pp. 24–34.

[SPS02] Ron Sun, Todd Peterson, and Chad Sessions. "Beyond simple rule extraction: Acquiring planning knowledge from neural networks". In: *Neural Nets WIRN Vietri-01*. Springer, 2002, pp. 288–300.

[SR13] Constantinos I Siettos and Lucia Russo. "Mathematical modeling of infectious disease dynamics". In: *Virulence* 4.4 (2013), pp. 295–306.

[SS01] Jordi Sabater and Carles Sierra. "REGRET: reputation in gregarious societies". In: *Agents*. Vol. 1. 2001, pp. 194–195.

[SS15] Jan Schnellenbach and Christian Schubert. "Behavioral political economy: A survey". In: *European Journal of Political Economy* 40 (2015), pp. 395–417.

[SS91] Emile Servan-Schreiber. "The Competitive Chunking Theory: Models of Perception, Learning, and Memory". PhD thesis. Carnegie Mellon University, 1991.

[Str89] Timothy J Strauman. "Self-discrepancies in clinical depression and social phobia: Cognitive structures that underlie emotional disorders?" In: *Journal of Abnormal Psychology* 98.1 (1989), pp. 14–22.

[Sun06] Ron Sun. "The CLARION cognitive architecture: Extending cognitive modeling to social simulation". In: *Cognition and multi-agent interaction* (2006), pp. 79–99.

[Sun+16] Zhanli Sun, Iris Lorscheid, James D. Millington, Steffen Lauf, Nicholas R. Magliocca, Jürgen Groeneveld, Stefano Balbi, Henning Nolzen, Birgit Müller, Jule Schulze, and Carsten M. Buchmann. "Simple or complicated agent-based models? A complicated issue". In: *Environmental Modelling & Software* 86 (2016), pp. 56–67.

[TBS05] Peter M Todd, Francesco C Billari, and Jorge Simao. "Aggregate age-at-marriage patterns from individual mate-search heuristics". In: *Demography* 42.3 (2005), pp. 559–574.

[TG13] Robert J Taormina and Jennifer H Gao. "Maslow and the motivation hierarchy: Measuring satisfaction of the needs". In: *The American journal of psychology* 126.2 (2013), pp. 155–177.

[TK06] Samuel Thiriot and Jean-Daniel Kant. "A multi-agent cognitive framework to model human decision making under bounded rationality". In: *IAREPSABE 06 (International Conference on Behavioural Economics and Economic Psychology), Paris*. 2006.

[TK79] Amos Tversky and Daniel Kahneman. "Prospect theory: An analysis of decision under risk". In: *Econometrica* 47.2 (1979), pp. 263–291.

[TL03] Yaacov Trope and Nira Liberman. "Temporal construal." In: *Psychological review* 110.3 (2003), pp. 403–421.

[TLA06] Niels A Taatgen, Christian Lebiere, and John R Anderson. "Modeling paradigms in ACT-R". In: *Cognition and multi-agent interaction: From cognitive modeling to social simulation* (2006), pp. 29–52.

[Tri77] Harry Charalambos Triandis. *Interpersonal Behavior*. Brooks/Cole Pub. Co., 1977.

[Tur+94] John C Turner, Penelope J Oakes, S Alexander Haslam, and Craig McGarty. "Self and collective: Cognition and social context". In: *Personality and social psychology bulletin* 20.5 (1994), pp. 454–463.

[TW04] Seth Tisue and Uri Wilensky. "NetLogo: Design and implementation of a multi-agent modeling environment". In: *Proceedings of the Agent2004 Conference*. Vol. 2004. Springer Cham, Switzerland, 2004, pp. 7–9.

[Urb00] Christoph Urban. "PECS: A reference model for the simulation of multi-agent systems". In: *Tools and techniques for social science simulation*. Springer, 2000, pp. 83–114.

[VLKH12] Paul AM Van Lange, AW Kruglanski, and E Tory Higgins. "Theories of social psychology: An introduction". In: *Handbook of theories of social psychology* (2012), pp. 1–8.

[VW14] Robin R Vallacher and Daniel M Wegner. *A theory of action identification*. Psychology Press, 2014.

[WBW76] Elaine Walster, Ellen Berscheid, and G William Walster. "New directions in equity research". In: *Advances in experimental social psychology*. Vol. 9. Elsevier, 1976, pp. 1–42.

[WD92] Christopher JCH Watkins and Peter Dayan. "Q-learning". In: *Machine learning* 8.3 (1992), pp. 279–292.

[Web+17] Sylvain Weber, Paul Burger, Mehdi Farsi, Adan L Martinez-Cruz, Michael Puntiroli, Iljana Schubert, and Benjamin Volland. *Swiss household energy demand survey (SHEDS): Objectives, design, and implementation*. Tech. rep. IRENE Working Paper, 2017.

[Wey+04] Danny Weyns, H Van Dyke Parunak, Fabien Michel, Tom Holvoet, and Jacques Ferber. "Environments for multiagent systems state-of-the-art and research challenges". In: *International Workshop on Environments for Multi-Agent Systems*. Springer. 2004, pp. 1–47.

[Wil00] Edward O. Wilson. *Sociobiology: The New Synthesis, Twenty-Fifth Anniversary Edition*. Harvard University Press, 2000.

[Woo09] Michael Wooldridge. *An Introduction to MultiAgent Systems*. 2nd. John Wiley & Sons, 2009.

[Woo99] Michael Wooldridge. "Intelligent agents". In: *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* 1 (1999), pp. 27–73.

[WR15] Uri Wilensky and William Rand. *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo*. MIT Press, 2015.

[WS01] DJ Wu and Yanjun Sun. "The emergence of trust in multi-agent bidding: a computational approach". In: *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*. IEEE Computer Society, 2001, p. 1041.

[WS98] Stefan Wermter and Ron Sun. "An overview of hybrid neural systems". In: *International Workshop on Hybrid Neural Systems*. Ed. by Stefan Wermter and Ron Sun. Springer. 1998, pp. 1–13.

[WW13] Annie Waldherr and Nanda Wijermans. "Communicating social simulation models to sceptical minds". In: *Journal of Artificial Societies and Social Simulation* 16.4 (2013), p. 13.

[Yah12] Hamdi Yahyaoui. "A trust-based game theoretical model for web services collaboration". In: *Knowledge-Based Systems* 27 (2012), pp. 162–169.

[YFK94] R. Yager, M. Fedrizzi, and J. Kacprzyk. *Advances in the Dempster- Shafer Theory of Evidence*. John Wiley, 1994.

[Yin+19] Xicheng Yin, Hongwei Wang, Pei Yin, and Hengmin Zhu. "Agent-based opinion formation modeling in social network: A perspective of social psychology". In: *Physica A: Statistical Mechanics and its Applications* 532 (2019), p. 121786.

[YS02] Bin Yu and Munindar P Singh. "Distributed reputation management for electronic commerce". In: *Computational Intelligence* 18.4 (2002), pp. 535–549.

[Zel+08] Moira L Zellner, Thomas L Theis, Arunprakash T Karunanithi, Ahjond S Garmestani, and Heriberto Cabezas. "A new framework for urban sustainability assessments: Linking complexity, information and policy". In: *Computers, environment and urban systems* 32.6 (2008), pp. 474–488.

[Zil96] Dolf Zillmann. "Sequential dependencies in emotional experience and behavior". In: *Emotion: interdisciplinary perspectives* (1996), pp. 243–272.

[ZL17] Yanjun Zuo and Jigang Liu. "A reputation-based model for mobile agent migration for information search and retrieval". In: *International Journal of Information Management* 37.5 (2017), pp. 357–366.

Index