

# A formal account of opportunism based on the situation calculus

Jieting Luo<sup>1</sup> · John-Jules Meyer<sup>1</sup>

Received: 28 November 2015 / Accepted: 20 May 2016 / Published online: 11 June 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** In social interactions, it is common for individuals to possess different amounts of knowledge about a specific transaction, and those who are more knowledgeable might perform opportunistic behavior to others in their interest, which promotes their value but demotes others' value. Such a typical social behavior is called opportunistic behavior (opportunism). In this paper, we propose a formal account of opportunism based on the situation calculus. We first propose a model of opportunism that only considers a single action between two agents, and then extend it to multiple actions and incorporate social context in the model. A simple example of selling a broken cup is used to illustrate our models. Through our models, we can have a thorough understanding of opportunism.

**Keywords** Opportunism · Value · Situation calculus · Formalization

## 1 Introduction

Consider a common social scenario. A seller is trying to sell a cup to a buyer, and it is known by the seller beforehand that the cup is actually broken (e.g., there is a crack at the bottom of the cup). The buyer buys the cup for its good appearance, but of course gets disappointed when he fills it with water. In this example, the seller earns

money from the buyer by exploiting the opportunity of having more knowledge about the transaction, while the buyer just focuses on the appearance of the cup rather than being leaky or not. Such a social behavior intentionally performed by the seller is first named opportunistic behavior (or opportunism) by economist Williamson (1975). Opportunistic behavior commonly exists in business transactions and other types of social interactions in various forms such as deceit, lying, and betraying. This is because individuals working in different positions are capable of having access to different amounts of information, which provides the opportunity for them to gain personal advantage, regardless of the consequences to others. Since it has negative results for other individuals involved in the relationship and strongly affects the cooperative relationship once it is unveiled, it is prevented or eliminated by social laws and norms.

Over the years, a large amount of research from social science was done to investigate opportunistic behavior from its own perspective (Conner and Prahalad 1996; Cabon-Dhersin and Ramani 2007; Jiraporn et al. 2008), providing a descriptive theoretical foundation to the study of opportunism. However, it is difficult for them to offer a general model that can be applied in any context. This is mainly because the original definition from Williamson “self-interest seeking with guile” is relatively implicit. On the one hand, it strongly captures various behaviors and activities that are judged as opportunism. On the other hand, it makes people have different interpretations of the concept. Therefore, there exists no agreed general and scientific definition of opportunism, which makes the study on its emergence and constraint mechanism even more difficult.

Is the investigation of opportunism of interest to AI? Social concepts are often used to construct artificial

---

✉ Jieting Luo  
J.Luo@uu.nl

John-Jules Meyer  
J.J.C.Meyer@uu.nl

<sup>1</sup> Intelligent Systems Group, Utrecht University,  
P.O. Box 80089, 3508 TB Utrecht, The Netherlands

societies. Viewing individuals as agents, we might have similar phenomena in the context of multi-agent systems. Interacting agents were designed to behave in a human-like way with characteristics of self-interest. When such agents possess different amounts of relevant information about a specific transaction and try to maximize their own benefits, those who are more knowledgeable may probably perform opportunistic behavior to other agents in their own interest, which is against others' benefits. For example, lying or deception is one of the strategies in multi-agent negotiation (Ettinger and Jehiel 2010; Zlotkin and Rosenschein 1991), and its effects are investigated in practical Turing tests (Warwick and Shah 2016), and it is possible for agents to hide important information to his or her peers for increasing his own payoff. In order to perform the investigation about opportunism, we first need to have a formal specification of opportunism with widely applicable generalization. Through the specification, we can understand more clearly the elements in the definition, how they relate to each other, and derive interesting properties that are useful for our future research. We believe that such a research perspective can ease the debates about opportunism in social science. Moreover, future work on its emergence and constraint mechanism can be conducted based on our formal definition, rendering our study relevant for multi-agent system (MAS) research.

In this paper, we take the initiative to propose a formal account of opportunism. Aiming at the investigation about the different judgment on opportunistic behavior, we integrate the notion of value to represent agents' preference on situations. We then formalize opportunism using the situation calculus (McCarthy 1968; Reiter 2001) as our technical framework based on our extended definition. We first propose a model of opportunism that only considers a single action between two agents, indicating three basic concepts such as knowledge asymmetry, value opposition, and intention in the model, and then extend it to multiple actions and incorporate social context in the model. A simple example of selling a broken cup is used to illustrate our models.

The rest of the paper is organized as follows. Section 2 has a informal definition of opportunism extended from Williamson's, highlighting the key elements we need to model. Our technical framework of the situation calculus is briefly introduced in Sect. 3 together with the appendices for the semantics we use. Section 4 proposes a preliminary model of opportunism, which serves as a basis for the following extensions. Sections 5 and 6 extend the model to multiple actions and incorporate social context, respectively. Section 7 illustrates our models by a simple example. Section 8 addresses related work, and Sect. 9 concludes the paper.

## 2 Defining opportunism with value

In this section, we extend Williamson's definition of opportunism and suggest a more explicit one as a prelude and basis to proposing a formal model in the next section.

### 2.1 Definition of opportunism

The classical definition of opportunism is offered by Williamson as "self-interest seeking with guile" (Williamson 1975). While this definition has been used in a large amount of research, it only mentions two attributes, self-interest and guile, explicitly, leaving other attributes for researchers to interpret from different perspectives. For example, Das defined partner opportunism as "a behavior by a partner firm that is motivated to pursue its self-interest with deceit to achieve gains at the expense of the other alliance members" (Das and Rahman 2010). In game-theoretical setting, Seabright defines opportunism as "the behaviour of those who seek to benefit from the efforts of others without contributing anything themselves" (Seabright 2010). Even though those definitions are elaborated enough, they come from different theoretical settings. In this study, based on the definition of Williamson, we compare opportunistic scenarios with non-opportunistic ones and redefine this social behavior in a more explicit way.

*Opportunism is a behavior that is motivated by self-interest and takes advantage of relevant knowledge asymmetry<sup>1</sup> to achieve own gains, regardless of the principles.*

First of all, there has been reached consensus that opportunistic behavior is performed with self-interest motivation (Das and Rahman 2010). We admit that self-interested pursuit is the natural property of human beings, but opportunism is more than that: Individuals with opportunistic behavior do not care about the negative effects on others. Secondly, relevant knowledge asymmetry provides the chance to individuals to be opportunistic. Opportunistic individuals may break the contracts or the relational norms using the relevant knowledge that others do not have. It is important for opportunistic individuals to use cheating, deceit, or infidelity for hiding their self-interest motive. Therefore, individuals with more relevant knowledge will have more potential for being opportunistic. Thirdly, principles are ignored by opportunistic individuals. The reason to use "ignore" here is to distinguish

<sup>1</sup> Many papers in social science use information asymmetry to represent the situation where one party in a transaction knows more compared to another. We argue that once the information is stored in our mind and can be used appropriately, it becomes our knowledge. For this reason, we would rather revise the term as knowledge asymmetry in this paper, which is also consistent with our technical framework.

opportunism from accidentally bringing harm to others. Opportunistic behavior is performed intentionally without any compensation to the victims. Principles can be the value of others, or the contract rules or the relational norms that are used for balancing various interests and already agreed to by a majority of the individuals. Fourthly, even though we do not explicitly declare the result of performing opportunistic behavior in our extended definition, such a social behavior must result in gains at the expense of others. Any self-interested behavior that does not end up in affecting other individuals should not be considered as opportunism.

From the above elaboration, we can derive something interesting and important about opportunism: Opportunistic individuals ignore the interest of others, which means that it is already known by them that the behavior will cause harm to others; as opportunistic individuals intend to gain personal advantage, can we say that it is also their intention to cause harm to others? We will investigate this problem through our formal models of opportunism.

## 2.2 Integrating with value

Based on the informal definition of opportunism, the example about hiding important information from peers that we encountered in the introduction is opportunistic behavior, since it is against others' benefits or the norms of the system. However, if hiding is not forbidden by the norm, the agent could not be said to have done anything wrong. Or if hiding is accepted by peers, it may not be against their interest. We can see that both the system's norms and the agents' perspectives can influence the judgment of opportunism, and they are the representation of value systems at the collective level and individual level, respectively, which might be different among systems and agents.

Value is something that we think is important, and various types of values together with their orderings form a value system. By integrating the notion of value into our model, the result of performing opportunistic behavior is represented as the promotion of opportunistic individuals' value and the demotion of others' value. Furthermore, even though a value system is relatively stable within individuals, it may differ across different individuals and societies. For different societies, each has its own value system as part of the social context, and it serves as the basis for any judgment within the society. In this sense, some behaviors which are regarded as opportunistic in one society may not be considered as opportunistic in another society, if the two societies do not share the same value system. A similar idea, although more focusing on opportunistic propensity, can be found in Chen et al. (2002). Given the value system

of the society, opportunistic behavior promotes the self-interest which is in opposition with others' value.

## 3 Technical framework: situation calculus

The situation calculus provides a formal language for representing and reasoning about dynamical worlds based on first-order logic. Its idea is that we can represent any reachable states in terms of actions that are required to reach them, and that the reachable states are called situations. There are three elements: actions  $Act$  that can be performed by agents, situations  $S$  that represent a history of action occurrences, and fluents  $F$  that describe the properties of the situation. Situation  $S_0$  represents the initial situation that no action can result in. The properties of situations are specified through relational and functional fluents taking a situation term as their last argument, which means their truth value may vary from situation to situation. The relational fluents can be true or false, while the functional fluents can take a range of values. For instance,  $ontable(x, s)$  is a relational fluent which is true in situation  $s$  where object  $x$  is on the table, and  $temperature(s)$  is a functional fluent whose value in situation  $s$  is an integer representing the temperature of the environment.

To represent how situations change, one has to specify in which situation an action can be performed and how to reason about the changes in the world by performing an action. In the situation calculus, we use predicate symbol  $Poss(a, s)$  to denote the set of preconditions that action  $a$  is executable in situation  $s$ , and a distinguished binary function  $do(a, s)$  to denote the unique successor situation that results from the performance of action  $a$  in situation  $s$ . For example, in order to pick up object  $x$ , one must have an empty hand and object  $x$  must be on the table in situation  $s$ :

$$Poss(pick(x), s) \equiv handempty(s) \wedge ontable(x, s).$$

And  $do(pick(x), s)$  represents the situation that results from the performance of action  $pickup(x)$  in situation  $s$ . One more example: In order to repair object  $x$  in situation  $s$ , the object  $x$  must be broken, and there must be a glue available in situation  $s$ :

$$Poss(repair(x), s) \equiv broken(x, s) \wedge hasglue(s).$$

Other special predicates and functions can be introduced as needed. For instance, propositions  $P$  can be used as assertions from classical proposition logic instead of fluents, that is, their truth values are not dependent on the situation but consistent throughout all the situations.

With the situation calculus, we can reason about how the world changes as the result of the available actions. The effects of actions are specified through *successor state*

axioms. For example, the effect on fluent *broken* of object  $x$  is:

$$\text{broken}(x, \text{do}(a, s)) \equiv \text{broken}(x, s) \vee (\exists r)\text{fragile}(x, s) \wedge a = \text{drop}(r, x),$$

which is saying that object  $x$  will be broken in the successor situation  $\text{do}(a, s)$  if and only if  $x$  is fragile in  $s$  and the action that takes us to the successor situation is someone  $r$  dropping  $x$ , or  $x$  is already broken in  $s$ .

This is a brief overview of the situation calculus, which is the technical preliminary of our formalization. However, this language can only provide information about the history of a situation, and there is no way to represent the future of a situation. For example, propositions like “I shall sell the cup now” cannot be represented by situation calculus. Since this representation is of great importance to our formalization, we extend the situation to one step further in the future. An extended situation is a pair  $(s, s')$  such that  $s$  is a situation and  $s'$  is the next situation of  $s$  connected with an action, and *occur* is a relation between actions and situations. Here is the semantics of *occur*:

$$(s, s') \models \text{occur}(a, s) \text{ iff } s' = \text{do}(a, s). \text{ That is, } \text{occur}(a, s) \text{ holds if action } a \text{ occurs in situation } s.$$

From now on, the situation calculus we are using as our technical framework will be extended with the semantics above.

After John McCarthy’s introduction of this theory, people made extensions capable of representing knowledge, belief, intention, and obligation in order to better reason about actions and their effects on the world (Shapiro et al. 2011; Scherl and Levesque 2003; Demolombe and Parra 2009). We will introduce and adopt those extensions in the following sections as appropriate. Since in the situation calculus the last argument is always a situation, we will follow this convention in this paper for any definition of fluents and predicates.

## 4 Formalizing opportunism

For better understanding, we first propose a preliminary model of opportunism that only considers a single action between two agents, without any legal or moral evaluation. It serves as a basis for the extensions of multiple actions and social context in the following sections. We will use normal possible-world semantics to define knowledge and neighborhood semantics to define intention. Ones who are unfamiliar with the two types of semantics can refer to “Appendices 2 and 3” for their introductions.

### 4.1 Knowledge asymmetry

We adopt the approach of Scherl to formalizing knowledge, which is to add an agents’ possible-world model of knowledge to situation calculus (Scherl and Levesque 2003). To treat knowledge as a fluent, we have a binary relation  $K(s', s)$ , reading as situation  $s'$  is epistemically accessible from situation  $s$ . It is reflexive ( $K(s, s)$  holds for all  $s \in S$ ), transitive ( $K(s, s') \wedge K(s', s'')$  implies  $K(s, s'')$  for all  $s, s', s'' \in S$ ) and symmetric ( $K(s, s')$  implies  $K(s', s)$  for all  $s, s' \in S$ ).

Definition 1 (*Knowledge*)

$$\text{Know}(i, \phi, s) \stackrel{\text{def}}{=} (\forall s') K_i(s', s) \rightarrow \phi[s']$$

This definition shows that agent  $i$  has knowledge about  $\phi$  if and only if  $\phi$  holds in all the epistemic possible situations of the agent. Then, we can have the definition of knowledge asymmetry.

Definition 2 (*Knowledge Asymmetry*)

$$\text{KnowAsym}(i, j, \phi, s) \stackrel{\text{def}}{=} \text{Know}(i, \phi, s) \wedge \neg \text{Know}(j, \phi, s) \wedge \text{Know}(i, \neg \text{Know}(j, \phi, s), s)$$

*KnowAsym* is a fluent in situation  $s$  where agent  $i$  has knowledge about  $\phi$  while agent  $j$  does not have it, and this is also known by agent  $i$ . It can be the other way around with  $i$  and  $j$ . But for simplicity of our model, we limit this definition to one case. Note that  $\phi$  can represent any proposition in this definition.

### 4.2 Value opposition

From the definition of opportunism, we know that agents have different evaluations on the same state transition. For agent  $i$  who performs opportunistic behavior, his value gets promoted, while the value of agent  $j$  gets demoted. We argue that this is because agents always have the evaluation from their perspective, which is part of their value system. This property of state transition is named value opposition in this study. In order to extend our technical framework with value theory, we define a symbol  $V$  to represent agents’ value system and a binary relation  $<$  over situations to represent agents’ preference, where  $s <_V s'$  denotes “ $s'$  is preferred to  $s$  based on value system  $V$ ”.

In the situation calculus, situations can be described in terms of propositions  $P$ , which are structured with objects and their properties. For having preferences on situations, we argue that agents evaluate the truth value of specific propositions, which are called perspectives in this study, based on their value systems. For instance, the buyer tries

to see whether the cup has good quality or not in order to have a preference on the situations before and after the transaction. In order to specify agents’ preference on situations, we first define a function *EvalRef* that represents agents’ perspective for evaluation:

**Definition 3 (Evaluation Reference)**

$$EvalRef : V \times S \times S \rightarrow P$$

It returns a proposition that an agent refers to for specifying his preference on two situations based on his value system. It is worth noting that in real life, agents’ specification of preferences on situations is based on a set of propositions  $2^P$  rather than a single proposition. For instance, both whether the cup has good quality and appearance are important to the buyer. For simplicity, here we restrict the return value to only one proposition without loss of generality.

We then specify agents’ preferences on situations, where *V* is restricted to perspective-based value:

$$s <_{V_i} s' \equiv \neg p(s) \wedge p(s') \text{ where } p = EvalRef(V_i, s, s')$$

$$s >_{V_i} s' \equiv p(s) \wedge \neg p(s') \text{ where } p = EvalRef(V_i, s, s')$$

It means that agent *i*’s value gets promoted/demoted from *s* to *s*’ when the truth value of the proposition *p* that he refers to based on his value system *V<sub>i</sub>* changes. As for the example about selling the broken cup, the seller’s value gets promoted when he has earned money from the transaction, whereas the buyer’s value gets demoted when the cup he bought is broken. Because of having different value systems, they refer to different propositions and thereby evaluate different propositions for specifying their preferences. Similar to knowledge asymmetry, we only limit the specification to one case in terms of the truth value of *p*.

**Definition 4 (Value Opposition)**

$$ValueOppo(i, j, s, s') \stackrel{def}{=} s <_{V_i} s' \wedge s >_{V_j} s'$$

We define value opposition as a property of a state transition where a state transition from *s* to *s*’ can promote the value of agent *i* but demote the value of agent *j*. In other words, agent *i* has positive effects from the state transition, while agent *j* has negative effects. Again, we only limit the definition to one case for simplicity.

### 4.3 Intention

Opportunistic behavior is performed by intent rather than by accident. In order to suggest this aspect in our formal model, we adopt the logic of *intention to do something for being something* in our framework. The notion of *Intend* is defined through neighborhood semantics instead of Kripke

semantics. This is because agents need not intend all the expected side effects of their intentions as Bratman argued (Bratman 1987). For example, an agent has a toothache and is going to see the dentist with intention to get his tooth fixed. Although the agent believes that it will cause him much pain, we surely cannot say that he *intends* to get the pain. The formal definition of *Intend* is given as follows:

**Definition 5 (Intention)**

$$Intend(i, a, \phi, s) \stackrel{def}{=} ||A|| \in N_I(i, s),$$

where

$$||A|| = \{s' \in S \mid occur(a, s') \wedge \phi[s', do(a, s')]\}$$

*N<sub>I</sub>(s)* is an intentional neighborhood function of an agent that returns a set of subsets of *S*, meaning that what is the case in the neighborhood is intended to have in situation *s*. *occur(a, s')* is true when action *a* is performed in situation *s*’, and *ϕ* is true in the state transition. An intention of agent *i* *Intend(i, a, ϕ, s)* holds if and only if the truth set of *occur(a, s')* and *ϕ[s', do(a, s')]* is an intentional neighborhood in *s*. Based on this definition of intention, we have two instances for value promotion *pro(j) = s' <<sub>v<sub>j</sub></sub> do(a, s')* and value demotion *de(j) = s' ><sub>v<sub>j</sub></sub> do(a, s')* by action *a*, which will be later used for providing the final definition and proving its properties

$$Intend(i, a, pro(j, v), s) \stackrel{def}{=} ||A|| \in N_I(i, s),$$

where

$$||A|| = \{s' \in S \mid occur(a, s') \wedge s' <_{v_j} do(a, s')\}$$

$$Intend(i, a, de(j, v), s) \stackrel{def}{=} ||A|| \in N_I(i, s),$$

where

$$||A|| = \{s' \in S \mid occur(a, s') \wedge s' >_{v_j} do(a, s')\}$$

*Intend(i, a, pro(j), s)* denotes that agent *i* intends to promote the value of agent *j* by action *a* in situation *s*. Similar for *Intend(i, a, de(j), s)*. When *i = j*, agent *i* intends to promote/demote his own value by action *a*.

### 4.4 Opportunistic behavior

The above definitions are basic ingredients that we need for having the formal model of opportunism: knowledge asymmetry as the precondition, value opposition as the effect, and intention as the mental state. Besides, based on the informal definition we gave in Sect. 2, there are two more aspects that should be suggested in the definition. Firstly, the knowledge that the performer has while others do not have should be relevant to the state transition. Secondly, the performer is aware of value opposition for



the state transition beforehand but still ignores it. Opportunism is defined as follows:

**Definition 6 (Opportunism)** Let  $D$  be a situation calculus  $BAT^2$ ,  $K$  and  $I$  be the axioms for knowledge and intention representation in the situation calculus, respectively,  $V$  be the value system of agents,  $EvalRef$  be the reference function representing the object for an agent's evaluation on situations, and  $<_V$  be a preference ordering on situations. Then,  $(D \cup K \cup I, V, EvalRef, <_V)$  is a situation calculus  $BAT$  extended with knowledge, intention, value, and preference. Within this system, we have

$$Opportunism(i, j, a, s) \stackrel{def}{=} Poss(i, j, a, s) \wedge Intend(i, a, pro(i), s) \wedge \phi$$

where  $Poss(i, j, a, s) \equiv KnowAsym(i, j, \phi, s)$   
 $\phi = ValueOppo(i, j, s, do(a, s))$ .

This formula defines a predicate *Opportunism* where action  $a$  is opportunistic behavior performed by agent  $i$  to agent  $j$  in the situation  $s$ . In this concise formula, the precondition of action  $a$  is knowledge asymmetry about the state transition from  $s$  to  $do(a, s)$ , and action  $a$  is performed by intent and results in value opposition.

One observation from the model is about the subjectivity of opportunism. We can see through the functional fluent *EvalRef* that agents always evaluate the situations and consequently the state transition from their own perspectives, which are part of their value systems. If the value systems upon which they have evaluation change to other ones, the property of value opposition may become false. Opportunism is presented as a problem in most research about it. However, the above formal model of opportunism implies that it depends on from which perspective, or more generally value system, we evaluate the state transition. It is positive from the perspective of agent  $i$ , while it is negative from the perspective of agent  $j$ . In reality and multi-agent systems, people usually take the established norms into consideration when they decide whether it should be prevented, and the result may be different from society to society and from system to system.

After having the formal model of opportunism, we show how the propositions we informally suggest in text at the beginning is captured by our formalization.

**Proposition 1** *Given an opportunistic behavior a performed by agent  $i$  to agent  $j$ , each agent evaluates the behavior from a different perspective, which is formalized as:*

$$\begin{aligned} &\models Opportunism(i, j, a, s) \rightarrow EvalRef(V_i, s, do(a, s)) \\ &\neq EvalRef(V_j, s, do(a, s)) \end{aligned}$$

**Proof** If *Opportunism*( $i, j, a, s$ ) holds, the property *ValueOppo*( $i, j, s, do(a, s)$ ) also holds. Following the definition of value opposition, we have

$$s <_{V_i} do(a, s) \wedge s >_{V_j} do(a, s).$$

The specification of  $s <_{V_i} do(a, s)$  is

$$\neg p(s) \wedge p(do(a, s)) \text{ where } p = EvalRef(V_i, s, do(a, s)) \quad (1)$$

The specification of  $s >_{V_j} do(a, s)$  is

$$q(s) \wedge \neg q(do(a, s)) \text{ where } q = EvalRef(V_j, s, do(a, s)) \quad (2)$$

Sentences (1) and (2) hold together. Since any formula has only one truth value given a situation, we have  $p \neq q$ , that is,

$$EvalRef(V_i, s, do(a, s)) \neq EvalRef(V_j, s, do(a, s)). \quad \square$$

**Proposition 2** *Given an opportunistic behavior a performed by agent  $i$  to agent  $j$ , agent  $i$  knows the performance of this behavior demotes agent  $j$ 's value, but needs not intend to get this result for agent  $j$ , which is characterized by:*

$$\begin{aligned} &\models Opportunism(i, j, a, s) \rightarrow Know(i, s >_{V_j} do(a, s), s) \\ &\neq Opportunism(i, j, a, s) \rightarrow Intend(i, a, de(j, v), s) \end{aligned}$$

**Proof** The first formula is already in the definition of opportunism, so we are going to prove the second one. In our model, opportunistic behavior is performed with intention and opportunistic behavior implies agent  $j$ 's value gets demoted, then definitely agent  $j$ 's value gets demoted in agent  $i$ 's intentional neighborhood where opportunistic behavior holds (denoted as set  $O$ ). In neighborhood semantics, if it holds that agent  $i$  intends to demote agent  $j$ 's value, then the truth set that agent  $j$ 's value gets demoted (denoted as set  $D$ ) must be an intentional neighborhood of agent  $i$ . However, we only know that  $O$  is an intentional neighborhood of agent  $i$  and  $D$  might be bigger than  $O$  ( $O \subseteq D$ ) so that  $D$  might not necessarily be an intentional neighborhood. Therefore, we can theoretically conclude that agent  $i$  might not intend to demote agent  $j$ 's value.

We can also empirically prove it. Free riding is one of the classic models about opportunism, and it occurs when someone benefits from resources, goods, or services but does not contribute to them, which results in either an

<sup>2</sup> See "Appendix 1" for an introduction of Reiter's Basic Action Theories.

under-provision of those goods or services, or in an overuse or degradation of a common property resource (Baumol 1967). Suppose agent  $i$  is a free rider, it is rather weird to say that agent  $i$  intends to reduce others’ share of public goods. □

The proposition shows that it is not the intention of opportunistic individuals to harm others even though opportunism is deliberate with self-interest motive. The ignored principles are a specific kind of knowledge about the interest of others that cannot be considered as an intention to be opportunistic. □

### 5 Opportunistic behavior for multiple actions

In the previous section, we only consider one single action as opportunistic behavior. But in the real life, it is common that opportunistic behavior consists of multiple actions. For instance, unlike the simple selling example at the beginning of this paper, commerce transactions between businesses usually have a couple of actions, each of which ends up in a situation. In this context, the whole sequence of actions is opportunistic behavior instead of any single action within. Of course, a sequence of actions can be seen as one action if we only look at the precondition of the first action and the effect of the last action, but we may also be interested in what properties we can derive from opportunistic behavior when considering multiple actions.

In situation calculus, a binary function  $do(a, s)$  is used to denote the situation resulting from performing action  $a$  in situation  $s$ , so for a finite sequence of actions  $[a_1, \dots, a_n]$ , the situation resulting from performing the sequence of actions in situation  $s$  is denoted as  $do(a_n, do(a_{n-1}, \dots, do(a_1, s)))$ . Each action within the sequence brings about a new situation that satisfies certain properties. Formally, based on Definition 6, opportunism for multiple actions is defined as below:

**Definition 7 (Opportunism for Multiple Actions)** Let  $D$  be a situation calculus  $BAT$ ,  $K$  and  $I$  be the axioms for knowledge and intention representation in the situation calculus, respectively,  $V$  be the value system of agents,  $EvalRef$  be the reference function representing the object for an agent’s evaluation on situations, and  $<_V$  be a preference ordering on situations. Then,  $(D \cup K \cup I, V, EvalRef, <_V)$  is a situation calculus  $BAT$  extended with knowledge, intention, value, and preference. Within this system, we have

$$\begin{aligned}
 & Opportunism(i, j, [a_1, \dots, a_n], s_1) \stackrel{def}{=} \\
 & \bigwedge_{1 \leq k \leq n} Poss(i, j, a_k, s_k) \wedge Intend(i, a_k, pro(i), s_k) \wedge \phi \\
 & \text{where } Poss(i, j, a_k, s_k) \equiv KnowAsym(i, j, \phi, s_k) \\
 & \phi = ValueOppo(i, j, s_1, do \\
 & (a_n, do(a_{n-1}, \dots, do(a_1, s_1)))) \\
 & s_k = do(a_{k-1}, \dots, do(a_1, s_1)) (1 < k \leq n).
 \end{aligned}$$

Because each action in the sequence must be possible to be performed and it is the property of intention to be persistent (Bratman 1987), knowledge asymmetry and intention is true in  $s_k$  for  $1 \leq k \leq n$ . Value opposition is the property of the state transition by the sequence of actions. A finite sequence of actions  $[a_1, \dots, a_n]$ , which is performed by agent  $i$  to agent  $j$  in situation  $s_1$ , is opportunistic behavior if and only if each action is possible to be performed with the intention to promote agent  $i$ ’s value and the whole sequence results in value opposition for agent  $i$  and  $j$ .

Regarding the effects of opportunistic behavior, agent  $j$ ’s value gets demoted by the behavior, which can be permanent or repairable. In the former case, it is impossible to compensate the negative effect on agent  $j$  (e.g., somebody dies from it), while in the latter case, it is possible in some forms (e.g., a broken cup can be returned). Since opportunistic behavior is performed by intent, we argue that agent  $i$  will not actively compensate agent  $j$ ’s loss, no matter it is permanent or repairable. For this reason, we introduce the following definition *non-compensation* for agent  $j$ , which is an essential property of opportunism:

**Definition 8 (Non-compensation)** Given a sequence of actions  $Seq = [a_1, \dots, a_n]$  as opportunistic behavior  $Opportunism(i, j, Seq, s_1)$  and  $q = EvalRef(V_j, s, do(a, s))$ , we say that  $Seq$  is non-compensated for agent  $j$  iff  $\exists k : a_k \in Seq$  such that for the subsequence of actions  $Seq_B = [a_1, \dots, a_k]$

$$q(s_1) \wedge \neg q(do(Seq_B, s_1))$$

and for the subsequence of actions  $Seq_R = [a_{k+1}, \dots, a_n]$ ,  $\forall m : a_m \in Seq_R$

$$q(do(a_m, s_m)) \equiv q(s_m).$$

By this definition, we separate the sequence of actions into two parts:  $Seq_B$  that brings about  $\neg q$ , and  $Seq_R$  that retains  $\neg q$ . Note that  $Seq_R$  can be empty, which implies that the whole sequence brings about  $\neg q$  and the situation

transition is permanent and irreversible. Moreover, as the whole sequence of actions is performed by agent  $i$ , the compensation for agent  $j$ 's loss comes from agent  $i$  rather than agent  $j$  itself or someone else.

Definition 7 together with its property of non-compensation captures some interesting properties, which cannot be derived from Definition 6. First of all,

Proposition 3 For a sequence of actions  $Seq = [a_1, \dots, a_n]$  being opportunistic behavior  $opportunism(i, j, Seq, s)$ , we have

$$\models Opportunism(i, j, Seq, s) \rightarrow (\exists a \notin Seq_R) \neg s >_{V_j} do([Seq_B, a], s)$$

It implies that the negative effect of opportunistic behavior on agent  $j$  could have been compensated but is not done by agent  $i$ . Typically, when  $Seq_R$  is empty, it is meaningless to talk about action  $a$ , because the negative effect is permanent.

Proposition 4 Given a finite sequence of actions  $[a_1, \dots, a_n]$  as opportunistic behavior, we can prove that

$$\models Opportunism(i, j, [a_1, \dots, a_n], s_1) \rightarrow KnowAsym(i, j, \phi, s_k) \equiv KnowAsym(i, j, \phi, do(a_k, s_k)) \quad (1 < k < n)$$

Proof Each action in the sequence is possible to be performed and also

$$Poss(i, j, a_k, s_k) \equiv KnowAsym(i, j, \phi, s_k) \quad (1 \leq k \leq n)$$

$$s_k = do(a_{k-1}, \dots, do(a_1, s_1)) \quad (1 < k \leq n)$$

Combining these two formulas, we can easily get

$$KnowAsym(i, j, \phi, s_k) \equiv KnowAsym(i, j, \phi, do(a_k, s_k)) \quad (1 \leq k < n). \quad \square$$

This proposition shows that, when opportunistic behavior consists of a sequence of actions, property knowledge asymmetry is preserved throughout the whole sequence.

Proposition 5 Given a finite sequence of actions  $[a_1, \dots, a_n]$  as opportunistic behavior, we can prove action  $a_i$  needs not be opportunistic, which is characterized by

$$\not\models Opportunism(i, j, [a_1, \dots, a_n], s_1) \quad (n > 1) \rightarrow Opportunism(i, j, a_k, s_k) \quad (1 \leq k \leq n)$$

Proof In order to prove this proposition, we are going to find a counterexample of opportunistic behavior which satisfies condition  $n > 1$ , but each action does not satisfy all the properties of opportunism.

Freeriding is still a nice model to prove this property. Since freeriding is one form of opportunistic behavior,  $Opportunism(i, others, freeride, s_1)$  is true in our model. Now, we are going to split it into a sequence of actions  $[a_1, \dots, a_n]$  and suppose a free rider exist in a society with a large population and benefits from the public goods without paying. Since the amount that the free rider is supposed to pay is shared by a large population, other agents do not notice (or even not care about) the small change in the current situation thus not getting their value demoted for little amount of freeriding. That is, for action  $a_k$ ,

$$EvalRef(V_{others}, s_k, do(a_k, s_k)) = \top$$

so that  $s <_{V_{others}} do(a, s)$  does not hold any more. Therefore, it is not true that

$$Opportunism(i, j, a_k, s_k) \quad (1 \leq k \leq n). \quad \square$$

However, once the amount that the free rider is supposed to pay accumulates to be large enough for getting other agents' value demoted (the whole sequence of actions is considered), will it be regarded as opportunistic behavior. By theoretical comparison, this example is quite similar to Sorites paradox, where grains are individually removed from a heap of sands and the heap stops being a heap when the process is repeated for enough times Hyde (2014). So it is also interesting to think about when the behavior starts to be regarded as opportunistic.

In the example above, the negative effect associated with agents' preference is ignored for its small change. It is also possible that the information associated with an agent's preference is blocked such that he cannot specify his preference on the situations and consequently cannot evaluate the actions. Only when he receives the specific information and compares his current situation with previous situations can the sequence of actions be considered as opportunistic behavior.

## 6 Opportunistic behavior with social context

In the previous sections, we made an assumption for the sake of simplicity that there is no legal or moral evaluation being made or implied to opportunistic behavior such that it is not necessarily good or bad. However, agents in MAS are residing in a social context which provides obligations, permissions, and other types of norms for guiding agents' behaviors. The setting of those norms reflects the value system of a MAS. To have a formal model of opportunism with social context, we can of course replace the agent  $j$  in



our previous models with a society (in this way, we see the whole society as an agent) and get similar properties as in last two sections, but now we are more interested in putting opportunism in a deontic-based social context to see how it relates to social norms. Thus, in this section, we are going to place opportunistic behavior into a social context with norms and propose a formal model of opportunism from this perspective.

For defining opportunistic behavior with a social context, we adopt the definition of knowledge asymmetry and intention in previous sections but redefine value opposition. Firstly, we have three normative statuses, which are similar to deontic logic.

- it is obligatory that (OB)
- it is permissible that (PE)
- it is forbidden that (FO)

Secondly, we define the above deontic notions for specifying the normative propositions  $\Pi$ .

**Definition 9 (Obligatory, Permissible, and Forbidden)**

$$OB(i, a, s) \stackrel{def}{=} (\forall s') R_i(s', s) \rightarrow occur(a, s')$$

$$PE(i, a, s) \stackrel{def}{=} (\exists s') R_i(s', s) \wedge occur(a, s')$$

$$FO(i, a, s) \stackrel{def}{=} (\forall s') R_i(s', s) \rightarrow \neg occur(a, s')$$

where  $R_i(s', s)$  denotes the deontic accessibility relation of agent  $i$ , meaning that what is the case in situation  $s'$  is ideal for situation  $s$ , and  $occur(a, s')$  is true when action  $a$  is performed in situation  $s'$ . R-relation is serial, which means for all situations  $s$ , there is at least one possible situation  $s'$  such that  $R_i(s', s)$  holds. This property of R-relation ensures the validity  $\models OB(i, a, s) \rightarrow PE(i, a, s)$  to be hold, which is also consistent with our intuition. Each modality can be taken as a basic to define the other two modalities.

We then specify the social preference on situations, where  $V$  is restricted to deontic-based social value.

$$s <_{V_A} s' \equiv (\exists a, i) s' = do(a, s) \wedge OB(i, a, s)$$

$$s >_{V_A} s' \equiv (\exists a, i) s' = do(a, s) \wedge FO(i, a, s)$$

Here, symbol  $A$  represents the whole society, which is a set of agents. The first one means that the social value gets promoted if there exists an action whose performance complies with the social norm, while the second one means that the social value gets demoted if there exists an action whose performing violates the social norm.

Together with the specification of agents' preferences on situations, we have the definition of value opposition between an agent and the whole society.

**Definition 10 (Value Opposition with Social Context)**

$$ValueOppo(i, A, s, s') \stackrel{def}{=} s <_{V_A} s' \wedge s >_{V_A} s'$$

For the state transition from  $s$  to  $s'$ , the value of agent  $i$  gets promoted, whereas the social value gets demoted. Again, we only limit the definition to one case excluding the other way around for simplicity.

Therefore, similar to Definition 6, we have the definition of opportunistic behavior with social context.

**Definition 11 (Opportunism with Social Context)** Let  $D$  be a situation calculus  $BAT$ ,  $K$  and  $I$  be the axioms for knowledge and intention representation in the situation calculus, respectively,  $V$  be the value system of agents,  $EvalRef$  be the reference function representing the object for an agent's evaluation on situations,  $\Pi$  be a finite set of normative propositions, and  $<_V$  be a preference ordering on situations. Then,  $(D \cup K \cup I, V, EvalRef, \Pi, <_V)$  is a situation calculus  $BAT$  extended with knowledge, intention, value, norms, and preference. Within this system, we have

$$Opportunism(i, A, a, s) \stackrel{def}{=} Poss(i, A, a, s) Intend(i, a, pro(i), s) \wedge \phi$$

where  $Poss(i, A, a, s) \equiv KnowAsym(i, A, \phi, s)$   
 $\phi = ValueOppo(i, A, s, do(a, s)).$

Action  $a$  performed by agent  $i$  is regarded as opportunistic behavior if and only if it is performed with the asymmetric knowledge  $\phi$  about the state transition from  $s$  to  $do(a, s)$  and the intention of self-interest, and results in value opposition against the society  $A$  where he is staying.

The definition of opportunistic behavior with a social context shows that given the value system of a society, opportunistic behavior is considered to be bad since its performance results in demoting the social value. Further, it implies the moral dilemma concerning the conflict between desire and obligation. More precisely, an agent has the desire "to do what he wants," while the social context where the agent is residing gives the obligation "to do what one ought to do." Opportunistic agents follow their desire but ignore the obligation. Hence, it should be prohibited by laws or social norms from the perspective of the whole society.

Since we assume a social context with norms in this section, it is worth investigating the relation between deontic notions and mental states. Our formalization governs Proposition 6 regarding opportunistic agents having knowledge about the relevant norms, and Proposition 7 and Proposition 8 about the intention of opportunistic behavior not being derived from the obligation.

**Proposition 6** *Let action  $a$  be opportunistic behavior performed by agent  $i$  within society  $A$  in situation  $s$ , for the normative proposition associated with action  $a$   $FO(i, a, s) \in \Pi$  we have*

$$\models \text{Opportunism}(i, A, a, s) \rightarrow \text{Know}(i, FO(i, a, s), s)$$

**Proof** Since  $\text{Opportunism}(i, A, a, s)$  holds, by Definition 11, agent  $i$  must have knowledge about the effect of performing action  $a$ , that is,  $\text{Know}(i, \phi, s)$  holds, where  $\phi$  represents value opposition. By Definition 10,  $\phi = s <_{V_i} do(a, s) \wedge s >_{V_A} do(a, s)$ . Therefore,  $\text{Know}(i, s >_{V_A} do(a, s), s)$  holds. Because  $V$  is restricted to deontic-based social value in our model,  $s >_{V_A} do(a, s) \equiv FO(a, s)$  holds, thereby  $\text{Know}(i, FO(i, a, s), s)$  holds as well.  $\square$

Agents have the knowledge about the relevant norms in the society and decide whether and which to comply with based on their own analysis. Typically, opportunistic agents behave in their interest, regardless of the social norms they are supposed to follow.

Moreover, as Broersen and his colleagues indicate in their BOLD architecture (Broersen et al. 2005), intention might be derived from obligation (e.g., I ought to go to work this morning, so I intend to go to work this morning), or might just come from agents' own desire (e.g., I feel thirsty, so I intend to get some water). In a given situation, agents intend to perform opportunistic behavior, which is motivated by self-interest. In order to prove this property rigorously, we should first prove the disobedience of opportunistic behavior.

**Proposition 7** *Let action  $a$  be opportunistic behavior performed by agent  $i$  within society  $A$  in situation  $s$ , and  $V_i$  be agent  $i$ 's value system, we can prove*

$$\models \text{Opportunism}(i, A, a, s) \rightarrow (V_i \neq \text{Obedience})$$

**Proof** By contradiction, we assume that  $V_i = \text{Obedience}$ . Because agent  $i$  obeys to the social norm in order to promote his value, action  $a$  should not be forbidden by the society, that is,  $FO(a, s)$  does not hold. Consequently,  $s >_{V_A} do(a, s)$  and  $\text{Opportunism}(i, A, a, s)$  do not hold, either. Therefore,  $V_i = \text{Obedience}$  is false for opportunistic behavior.  $\square$

Using Proposition 7, we are going to prove it is not the case for opportunistic behavior that the intention is derived from the obligation.

**Proposition 8** *Let action  $a$  be opportunistic behavior performed by agent  $i$  within society  $A$  in situation  $s$ , for the social norm associated with action  $a$   $OB(i, a, s) \in \Pi$ , we can prove*

$$\not\models \text{Opportunism}(i, A, a, s) \rightarrow (OB(i, a, s) \rightarrow \text{Intend}(i, a, pro(i), s))$$

**Proof** We can prove this proposition by contradiction. Suppose action  $a$  is opportunistic behavior and sentence  $OB(i, a, s) \rightarrow \text{Intend}(i, a, pro(i), s)$  holds in our model, which means the intended situations of agent  $i$  are the subset of ideal situations, formalized as  $(\forall s')(A) \rightarrow R_i(s', s)$ , where  $(A) = s' \in S \mid occur(a, s')$ . Therefore, agent  $i$  intends to promote his own value and the social value by action  $a$ . Of course, when agent  $i$ 's value is obedience, both agent  $i$ 's value and the social value are promoted. But we have already proved in 7 that this possibility does not exist. So our assumption at the beginning is wrong. Therefore,  $OB(i, a, s) \rightarrow \text{Intend}(i, a, pro(i), s)$  does not hold in our model.  $\square$

## 7 Example: selling a broken cup

Recall the example that we used to introduce opportunism at the beginning of the paper. The scenario is simple but enough to illustrate our formal specification of opportunism. We label the seller and the buyer as  $s$  and  $b$ , who can be in one of the situations:  $S_0$  (the initial situation, before the transaction) and  $do(a, S_0)$  (after the transaction). The seller can either sell the cup ( $a = sell(x)$ ) or keep it. If the seller performs the action  $sell(x)$  in  $S_0$ , then situation will go to  $do(sell(x), S_0)$ .

In situation  $S_0$ , the asymmetric knowledge owned by the seller but not the buyer is not only about the broken cup, but also the state transition: Once the transaction finishes, the situation will go from  $S_0$  to  $do(sell(x), S_0)$ , which gets the value of the seller promoted, whereas the value of the buyer demoted. That is, the precondition  $\text{KnowAsym}(s, b, \phi, S_0)$  holds. Now consider the value for both parties. Apparently, both parties go for economic value. However, they have different and contradictory perspectives about the economic value. What the seller looks at is how much money he earns from the transaction. When the broken cup has already been sold, his value gets promoted ( $S_0 <_{V_s} do(sell(x), S_0)$  holds). Conversely, what the buyer looks at is whether the cup has good quality or not. So once the buyer knows the cup is broken, his value gets demoted ( $S_0 >_{V_b} do(sell(x), S_0)$  holds). The above two sentences ensure sentence  $\text{ValueOppo}(s, b, S_0, do(sell(x), S_0))$  holds. Further, since it is the seller's intention to sell the broken cup to the buyer for promoting his value, sentence  $\text{Intend}(s, sell(x), pro(s), S_0)$  also holds. With the above formalization, the formula for this example  $\text{Opportunism}(s, b, sell(x), S_0)$  holds.

We now discuss two interesting situations extended from the simple example. Firstly, if the buyer buys the cup only for decoration without using it, he will never know the cup is broken or even cares about it. That is, the buyer's perspective is revised to  $EvalRef(V_b, S_0, do(sell(x), S_0)) = appearance$  and then sentence  $S_0 >_{v_b} do(sell(x), S_0)$  does not hold any more. In this case, because the two perspectives are not contradictory, the seller's behavior is not opportunistic from the perspective of the buyer, if the social norms are not taken into account. It is already proved in Proposition 1 that agents have different perspectives about the same state transition if there is opportunistic behavior between those two agents. But the above discussion shows that having different perspectives does not necessarily lead to opportunistic behavior: they must be contradictory. The subjectivity of opportunism is reflected by the different judgments on the same action.

Secondly, if there is nothing, the seller can do except sell the broken cup when being in state  $S_0$ , it will be regarded as opportunistic behavior based on Definition 6, with the nature of self-defense, which might be allowed by the society. It is because there is no moral or legal evaluation in this definition, thus no matter whether the behavior is good or bad. However, it will be different if we analyze it with Definition 11. Suppose self-defense behavior is allowed by the society ( $PE(i, a, S_0)$ ). Then,  $S_0 >_{v_A} do(sell(x), S_0)$  does not hold, and then selling a broken cup is not opportunistic behavior from the perspective of the society. In our example, however, the options available to the buyer in state  $S_0$  are  $\{sell, keep\}$ , which means it is not the only choice for the seller to sell the broken cup. Moreover, sometimes it is our intention to put ourselves in a situation where we only have one option to choose. In this case, the action with the nature of self-defense might not be opportunistic, but the whole sequence of actions that illustrates how the situation arrives in one option available might be opportunistic.

Further, with the help of our model, we can gain practical insights into constraint mechanism of opportunism. In our case, one important reason why the seller's behavior is seen as opportunistic is that the seller and the buyer evaluate the state transition from two contradictory perspectives based on their value systems. In other words, even though they both go for economic value, they look at different things for evaluation. When applying this approach in collaborative relationship, it is much easier to understand how the relationship ends in defection. Therefore, one deterrence mechanism for partner opportunism is to avoid having contrasted value systems in the relationship. As for the precondition of opportunism, even though it is difficult to prevent knowledge asymmetry in business transactions, we still need to think about how much

information we can provide to our partners, especially during negotiation, and how they are going to use the information.

## 8 Discussion

As we try to propose a simple but elegant model of opportunism for different context settings, our specification might not manage to capture every possible scenario. For instance, in Sect. 4 we only talk about the interaction between two agents and investigate the evaluation on the state transition based on the value system of the two agents who are involved in the transaction. But actually such evaluation can also be done by others. This is because in the specification of value promotion and demotion, the proposition evaluated based on an agent's value system is not necessarily related to the transactions he is involved. Assume that a friend of the buyer knows the story about the broken cup. He may get angry with the seller for the unfair transaction, and then, the behavior performed by the seller is regarded as opportunistic from his perspective, even though he is not involved. In other words, the judgment of opportunism is subjective not only for the agents involved, but also for anybody who evaluate the action based on his or her own value system. Further, our models only consider intentional actions. However, opportunistic behavior can also be about intentional inactions such as withholding information. In this case, the social value gets demoted for agent  $i$ 's not performing an obligatory action instead of performing a forbidden action. Of course, our models can capture this scenario in a way that doing nothing can be seen as a particular way of doing something. Interesting insights can be gained from further study on this part.

We also propose that the asymmetric knowledge obtained by opportunistic agents is value opposition about the state transition, which is out of our intuition. The reason can be shown by the example in Sect. 7. Intuitively, the asymmetric knowledge that the seller has is about the broken cup. Now, we assume that both the seller and the buyer know the cup is broken and the seller sells it with a high price. Once the buyer knows that the broken cup is not worth that price, his value will get demoted. From that, we can see it does not matter whether the fact about the broken cup is only known by one party beforehand, but whether value opposition about the transaction is only known by one party beforehand. In other words, the asymmetric knowledge is not about the objective fact, but about agents' evaluation on the state transition.

The definition of non-compensation for multiple actions is introduced, based on the fact that the negative effect of opportunistic behavior can be permanent or repairable.

Given a social context, the norm that opportunistic behavior triggers (violates) can be repaired or not based on the same fact. In the former case, we can eliminate opportunistic behavior by imposing punishment or sanction on the norm. For instance, the problem of free riding can be handled through fine and social embarrassment. Opportunistic agents may be forced to repair the norm by regimented norms after the opportunistic behavior is detected. When the norm cannot be repaired once being violated, such an opportunistic behavior is supposed to be prevented from happening. In other words, the norm should be implemented in the environment or by designing norm-abiding agents.

## 9 Related work

Opportunism is not a new topic in social science. Since it was proposed by economist Williamson, scholars have studied this typical social behavior of economic players from various perspectives, i.e., transaction cost economics (Williamson 1985), resource-based view (Conner and Prahalad 1996), game theory (Cabon-Dhersin and Ramani 2007), agency theory (Jiraporn et al. 2008), and strategic management (Yaqub 2009). Even though they are indeed all worthwhile, it is difficult to directly apply their conclusions to MAS for improving the system's behavior because most of them are informal, which makes reasoning about this behavior in MAS impossible, and also not commonly accepted even in their own area.

In the field of artificial intelligence, there is a tradition to devise intelligent artifacts and construct intelligent system using symbolic representation of all factors involved (Segerberg et al. 2009). Especially for mathematical logic, it is a greatly important approach to this field due to its highly abstract representation and reasoning about social reality. Therefore, a lot of work on logic formalism has been designed for representing and reasoning about dynamical domains such as situation calculus (McCarthy and Hayes 1968), event calculus Kowalski and Sergot (1989) and fluent calculus (Thielscher 1998). We chose to use the situation calculus as our basic framework because it has been well developed and extended with knowledge (Scherl and Levesque 2003), belief (Shapiro et al. 2011), and other model semantics. For instance, in Scherl and Levesque (2003), an epistemic fluent  $Know(P, s)$  is proposed by adapting the standard possible-world model of knowledge. We use this approach to define knowledge asymmetry where agents possess different amounts of knowledge.

Formalization of opportunism is new in MAS, but there is some work on logic of lying, deception, and dishonesty (Sakama et al. 2010, 2015; Van Ditmarsch

et al. 2012), which are forms of opportunism. In their work, modalities for belief and intention are commonly used for formalizing different types of dishonest communication, which is similar to our work. However, Sakama's work (Sakama et al. 2010, 2015) only formalizes one agent's communication to another agent and his mental states, regardless of the effect on another agent, which means that we cannot reason about the state transition based on the approach. The primary goal of van Ditmarsch's work (Van Ditmarsch et al. 2012) was to model lying by modeling how agents' beliefs change from the communications. It analyzes the effect of lying in public discourse and explains how lying can be used as an optimal strategy through a game-theoretical analysis. For providing a formal model of opportunism, we not only need to formalize the mental states of interacting agents, but also need to reason about how the physical situations are changed by opportunistic behavior, both of which are related to the above work.

We also integrated the notion of value into the situation calculus to represent agents' preference on situations. However, in logical formalization, people usually use goals rather than value (e.g., Cohen and Levesque 1990; Rao and Georgeff 1991) for the same purpose. Only some work in the area of argumentation reasons about agents' preferences and decision making by value (e.g., Bench-Capon et al. 2012; van der Weide 2011). Even though both goals and values can be used to represent agents' preferences about situations, they have different features. Goals are concrete and should be specified with time, place, and objects. For example, to earn 1000 euros next month is a goal. If one agent's goal is achieved in one situation, then he has high evaluation on that situation. Value is described by Schwartz as trans-situational (Schwartz 1992), which means that value is relatively stable and not limited to be applied in a specific situation. For instance, if honesty is a value of somebody, he will be honest for a long period of time. Since state transition results from the performing of actions, we can evaluate actions by whether our value is promoted or demoted in the state transition, as what we do in this study. For representing agents' evaluation on situations, Keeney and Raiffa proposed multi-attribute utility theory (MAUT) in which situations are described in terms of a set of attributes and the utilities of the situations are calculated by the sum of the scores on each attribute based on agents' value system (Keeney and Raiffa 1993). In this study, we use a similar approach in which situations are represented through propositions and agents refer to a specific proposition based on their value systems to evaluate a state transition. Apparently, different agents may refer to different propositions, thus having different evaluations on the same state transition.



## 10 Conclusions

Agents situated with information asymmetry might perform opportunistic behavior to others in their interest. Numerous works about such a social behavior have been done in social science due to its negative effect on the relationship between people. However, most conclusions are based on a given form of opportunism, making it hard to build a fundamental theory that can be applied in any context. This study takes the initiative to propose a formal model of opportunism in the multi-agent system context based on the extended informal definition from Williamson. The modeling work is done based on the situation calculus integrating the notion of value. We first have a preliminary model that only considers a single action between two agents and then extend it for multiple actions with social context. Each model captures interesting properties that are useful for our future research. It is important to keep in mind that the aim of this paper is not to find out where opportunistic behavior comes from and how to eliminate it, but rather to have a thorough understanding of the nature of opportunism before exploring those issues. Therefore, the main strength of this study is defining such a behavior from our specific perspective in a formal way, so as to represent the elements in the definition and their relations and reason about the state transition by the behavior.

Based on our understanding of the concept of opportunism, we can study where and when opportunism arises in a social setting. Evaluation based on different value systems is the reason for value opposition of a state transition. So considerable insights can be achieved from the investigation about the compatibility of different value systems and the coevolution of agents' value systems with social context or environmental changes. Further, as opportunism is a self-interested behavior that conflicts with social norms, its emergence might come from the way in which agents resolve the conflicts between beliefs, obligations, intentions, and desires. For instance, an agent whose desires always overrule obligations might behave opportunistically in his interest. Those conflicts and their resolutions corresponding to different agent types are investigated in BOID architecture (Broersen et al. 2001, 2005). Similar to lie detection (Ohmoto et al. 2009), a well-designed monitoring mechanism can be used to automatically detect opportunism in (computer-based) human interactions, providing ways to protect agents' values from being demoted. Another important topic is designing constraint mechanisms to eliminate or prevent opportunism from happening.

**Acknowledgments** The research was supported by China Scholarship Council. We would like to thank anonymous reviewers, Marlo

Souza, Hein Duijf, Max Knobbout, and other colleagues for their helpful comments.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix 1: Basic action theories

Based on the situation calculus, Reiter's basic action theories are of the following form. A basic action theory  $\mathcal{D}$  (Reiter 2001; Levesque et al. 1998) is defined as a set of axioms:

$$D = \Sigma \cup D_{ap} \cup D_{ss} \cup D_{so} \cup D_{una}$$

- $\Sigma$ : the set of foundational axioms,
  - $do(a_1, s_1) = do(a_2, s_2) \rightarrow a_1 = a_2 \wedge s_1 = s_2$ ; Two situations are the same if and only if they are the same sequence of actions.
  - $(\forall Q)Q(S_0) \wedge (\forall s, a)[Q(s) \rightarrow Q(do(a, s))] \rightarrow (\forall s)Q(s)$ ; This is a second-order induction axiom saying that for any property  $Q$ , if  $Q(S_0)$  and, for any situation  $s$  and action  $a$ , property  $Q$  remains the same, then we have  $(\forall s)Q(s)$ .
  - $s \sqsubset do(a, s') \equiv s \sqsubseteq s'$ ;
  - $\neg s \sqsubset S_0$ ; The relation  $\sqsubset$  provides an ordering relation on situations.  $s \sqsubset s'$  means that the action sequence  $s$  is a subsequence of that of  $s'$ . Thus,  $s$  is a subsequence of  $do(a, s')$  if and only if  $s$  is a subsequence of  $s'$  or they have the same action sequence. And no situation is before initial situation  $S_0$ .
- $D_{ap}$ : the set of actions preconditions,
 
$$Poss(a(x), s) \equiv \pi(x, s)$$
 where  $\pi(x, s)$  is a formula uniform in  $s$  and whose free variables are among  $x$  and  $s$ . Thus, whether  $a(x)$  can be performed in situation  $s$  depends entirely on  $s$ .
  - $D_{ss}$ : the set of successor state axioms,
 
$$F(do(a, s)) \equiv \gamma_F^+(a, s) \vee (F(s) \wedge \neg \gamma_F^-(a, s))$$
 Here,  $\gamma_F^+(a, s)$  and  $\gamma_F^-(a, s)$  are two formulas expressing the conditions for the fluent  $F$  becoming true and false, respectively;
    - $D_{so}$ : the sentences uniform in  $S_0$  describing the initial situation;
    - $D_{una}$ : the unique name axioms for actions.



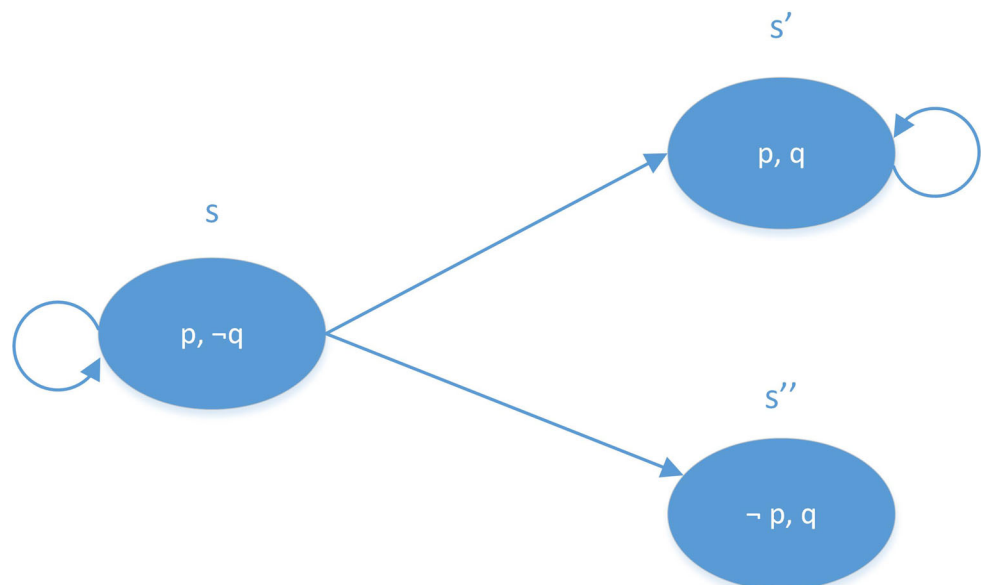
## Appendix 2: Possible-world structure

Possible-world structure (or Kripke structure) is the model that people adopt to formalize knowledge, belief, intention, and obligation in the situation calculus. Therefore, we will briefly introduce this model before we use those modalities in the next section. A Kripke structure is proposed by Kripke (1963) and has become the standard type of the models in modal logic and related non-classical logics. Basically, it is a graph whose nodes represent the possible situations of the system and whose edges represent accessibility relations. A valuation function maps each node to a set of properties hold in the corresponding state. Formally, let  $\Phi$  be a set of atomic propositions. A Kripke structure over  $\Phi$  is defined as a tuple  $M = (S, R, \pi)$ , where

- $S$  denotes a set of situations;
- $R \subseteq S \times S$  is a set of accessibility relations;
- $\pi : S \rightarrow 2^\Phi$  denotes a valuation function, meaning that for each situation  $s \in S$ , the set  $\pi(s)$  of atomic propositions hold in  $s$ . Therefore, *fluents* in the situation calculus can be interpreted as: Given a proposition  $p$ , fluent  $p(s)$  holds iff  $p \in \pi(s)$ .

By means of a Kripke structure, we can represent exactly an agent's mental state in a certain situation. Figure 1 is an example of a Kripke structure. Suppose that the actual situation is that  $p$  is true and  $q$  is false, represented by situation  $s \in S$  for which it holds that  $p(s)$  and  $\neg q(s)$ . Now, the model can be represented by  $S = \{s, s', s''\}$ , where  $s$  is as above,  $s'$  is  $p(s')$  and  $q(s')$ , and  $s''$  is  $\neg p(s'')$  and  $q(s'')$ . The accessibility relation  $R$  is illustrated in Fig. 1.

**Fig. 1** Example of a Kripke structure



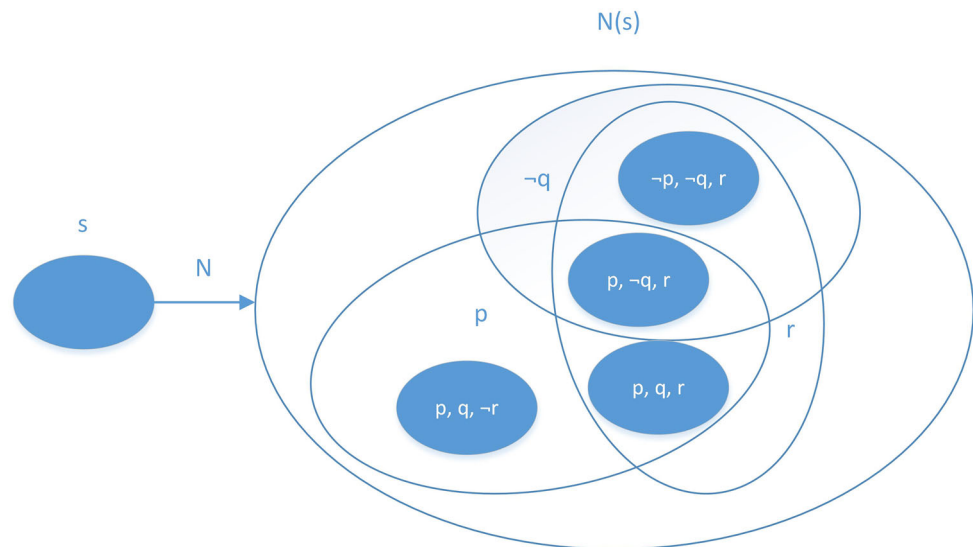
Kripke structures are adopted by the situation calculus to represent knowledge, belief, intention, and obligation. Taking knowledge as an example, we assume that there is an accessibility relation over situations, where situation  $s'$  is accessible from situation  $s$  if an agent residing in situation  $s$  thinks he might be in situation  $s'$ . So something is known in situation  $s$  if it holds in situation  $s$  and every situation  $s'$  accessible from  $s$ , and something is not known if it does not hold in at least one accessible situation.

## Appendix 3: Neighborhood semantics

Neighborhood semantics (Pacuit 2007), also known as Scott–Montague semantics, is another formal semantics for modal logics compared to normal possible-world semantics. It is developed by Dana Scott and Richard Montague. The basic idea behind a neighborhood model is that: At each situation, list all the sets that are considered “necessary.” That is, given a non-empty set of situations  $S$ , each situation  $s$  is assigned a set of subsets of  $S$  (these subsets are called neighborhoods). Formally, let  $\Phi$  be a set of atomic propositions. A neighborhood model over  $\Phi$  is defined as tuple  $M = \{S, N, v\}$ , where

- $S$  denotes a set of situations;
- $N$  is a neighborhood function  $N : S \rightarrow 2^{2^S}$  which assigns a collection of sets of situations to each situation in  $S$ ;
- $v : \Phi \rightarrow 2^S$  denotes a valuation function assigning a set of possible worlds to each atomic proposition.

**Fig. 2** Example of a neighborhood model



Therefore, *fluents* in the situation calculus can be interpreted as: Given a proposition  $p$ , fluent  $p(s)$  holds iff  $s \in v(p)$ .

Similar to Kripke structures, we can represent exactly an agent's mental state in a certain situation by neighborhood semantics. Figure 2 is an example of a neighborhood model. Suppose that  $s$  is the actual situation and  $S$  consists of the following situations:  $S = \{(p, q, r), (p, \neg q, r), (p, q, \neg r), (\neg p, \neg q, r)\}$ . Neighborhood function  $N(s)$  returns a set of subsets of  $S$  that are the neighborhoods in  $s$ . Set  $\{(p, q, r), (p, \neg q, r), (p, q, \neg r)\}$  is called the truth set of  $p$  and it is a neighborhood in  $s$ . The same with  $\neg q$  and  $r$ . The model is illustrated as below:

In this paper, we adopt neighborhood semantics to define *Intention*. Suppose we have a set of situations labeled with propositions. Proposition  $p$  is intended to be in the actual situation  $s$  if and only if the truth set of  $p$  is an intentional neighborhood in  $s$ .

## References

- Baumol WJ (1967) Welfare economics and the theory of the state. Harvard University Press, Cambridge
- Bench-Capon T, Atkinson K, McBurney P (2012) Using argumentation to model agent decision making in economic experiments. *Auton Agent Multi Agent Syst* 25(1):183–208
- Bratman M (1987) Intentions, plans, and practical reason. Harvard University Press, Cambridge, MA
- Broersen J et al (2001) The BOID architecture: conflicts between beliefs, obligations, intentions and desires. In: Proceedings of the fifth international conference on autonomous agents. ACM
- Broersen J, Dastani M, van der Torre L (2005) Beliefs, obligations, intentions, and desires as components in an agent architecture. *Int J Intell Syst* 20(9):893–919
- Cabon-Dhersin M-L, Ramani SV (2007) Opportunism, trust and cooperation a game theoretic approach with heterogeneous agents. *Ration Soc* 19(2):203–228
- Chen CC, Peng MW, Saporito PA (2002) Individualism, collectivism, and opportunism: A cultural perspective on transaction cost economics. *J Manag* 28(4):567–583
- Cohen PR, Levesque HJ (1990) Intention is choice with commitment. *Artif Intell* 42(2):213–261
- Conner KR, Prahalad CK (1996) A resource-based theory of the firm: Knowledge versus opportunism. *Organ Sci* 7(5):477–501
- Das TK, Rahman N (2010) Determinants of partner opportunism in strategic alliances: A conceptual framework. *J Bus Psychol* 25(1):55–74
- Demolombe R, Parra PP (2009) Integrating state constraints and obligations in situation calculus. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial* 13(41):54–63
- Ettinger D, Jehiel P (2010) A theory of deception. *Am Econ J Microecon* 2(1):1–20
- Hyde D (2014) Sorites paradox. In: Zalta EN (ed) The stanford encyclopedia of philosophy (Winter 2014 edn). <http://plato.stanford.edu/archives/win2014/entries/sorites-paradox/>
- Jiraporn P et al (2008) Is earnings management opportunistic or beneficial? An agency theory perspective. *International Review of Financial Analysis* 17(3):622–634
- Keeney RL, Raiffa H (1993) Decisions with multiple objectives: Preferences and value trade-offs. Cambridge University Press, Cambridge
- Kowalski R, Sergot M (1989) A logic-based calculus of events. *Foundations of knowledge base management*. Springer, Berlin Heidelberg
- Kripke SA (1963) Semantical analysis of modal logic I normal modal propositional calculi. *Math Log Q* 9.56:67–96
- Levesque H, Pirri F, Reiter R (1998) Foundations for the situation calculus. *Linköping Electron Artic Comput Inf Sci* 3(18)
- McCarthy J (1968) Situations, actions, and causal laws, Technical report, DTIC Document
- McCarthy J, Hayes P (1968) Some philosophical problems from the standpoint of artificial intelligence. Stanford University, Stanford
- Ohmoto Y, Ueda K, Ohno T (2009) Real-time system for measuring gaze direction and facial features: Towards automatic discrimination of lies using diverse nonverbal information. *AI&Soc* 23(2):187–200

- Pacuit E (2007) Neighborhood semantics for modal logic. Notes of a course on neighborhood structures for modal logic. [http://staff.science.uva.nl/~epacuit/nbhd\\_essli.html](http://staff.science.uva.nl/~epacuit/nbhd_essli.html)
- Rao AS, Georgeff MP (1991) Modeling rational agents within a BDI-architecture. KR 91:473–484
- Reiter R (2001) Knowledge in action: logical foundations for specifying and implementing dynamical systems, vol 16. MIT Press, Cambridge
- Sakama C, Caminada M, Herzig A (2010) A logical account of lying. Logics in artificial intelligence. Springer, Berlin
- Sakama C, Caminada M, Herzig A (2015) A formal account of dishonesty. Log J IGPL 23(2):259–294
- Scherl RB, Levesque HJ (2003) Knowledge, action, and the frame problem. Artif Intell 144(1):1–39
- Schwartz SH (1992) Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. Adv Exp Soc Psychol 25(1):1–65
- Seabright P (2010) The company of strangers: a natural history of economic life. Princeton University Press, Princeton
- Segerberg K, Meyer J-J, Kracht M (2009) The logic of action. In: Zalta EN (ed) The Stanford Encyclopedia of Philosophy (Winter 2013 Edition). <http://plato.stanford.edu/archives/win2013/entries/logic-action/>
- Shapiro S, Pagnucco M, Lesprance Y, Levesque HJ (2011) Iterated belief change in the situation calculus. Artif Intell 175(1):165–192
- Thielscher M (1998) Introduction to the fluent calculus. Electron Trans Artif Intell 3. <http://www.etaij.org>.
- van der Weide TL (2011) Arguing to motivate decisions. SIKS Dissertation Series 2011
- Van Ditmarsch H et al (2012) On the logic of lying. Games, actions and social software. Springer, Berlin
- Warwick K, Shah H (2016) Effects of lying in practical Turing tests. AI&Soc 31(1):5–15
- Williamson OE (1975) Markets and hierarchies, analysis and antitrust implications: a study in the economics of internal organization. Free Press, New York
- Williamson OE (1985) The economic institutions of capitalism. Free Press, New York
- Yaqub MZ (2009) Antecedents, consequences and control of opportunistic behavior in strategic networks. J Bus Econ Res JBER) 7(2):15–32
- Zlotkin G, Rosenschein JS (1991) Incomplete information and deception in multi-agent negotiation. IJCAI 91:225–231