A Closure-Based Semantics for Inconsistency Tolerant Agents

Linda Gutsche[1]

[1]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG, F-38000 Grenoble, France

Abstract

Physical agents have limited resources and cannot compute all the logical consequences of their beliefs. They therefore possess incomplete deductive abilities, and thus may fail to recognise inconsistencies. This does not prevent them from using and revising their beliefs, and from possibly resolving the undetected inconsistency after some time. However, most multi-agent logics of beliefs represent an agent's beliefs as a logically closed set, and therefore cannot represent inconsistent beliefs without explosion. Here we present a multi-agent logic of belief that can represent agents with incomplete deductive abilities and inconsistent beliefs.

Keywords

multi-agent logic, belief bases, incomplete deductive abilities, inconsistent beliefs

1. Introduction

Agents reason about the world they live in by maintaining an informational state and drawing inferences from it. Physical agents, whether human or artificial, are inherently bounded: they store only a finite amount of information and have limited deductive abilities. They cannot, and therefore do not, draw all the logical consequences of their beliefs. As a result, they may fail to detect that some of their beliefs are jointly inconsistent. Moreover, the presence of undetected inconsistencies does not trigger a belief revision process aimed at resolving them. Consequently, real-world agents may hold inconsistent beliefs.

Human reasoning provides a natural illustration of this phenomenon. People can entertain mutually inconsistent beliefs without also believing any arbitrary conclusion, unlike what most logically-complete models of belief would predict. For example, one might have the following belief: "By paying €10 per month for the subscription, I get a 50% reduction on everything. I am thus spending only €100 per year on purchases. The subscription is saving me money." This belief is logically inconsistent, yet someone may still believe it because they did not perform the computations required to detect the contradiction.

This situation does not fit the common assumption in logical models of belief that agents are logically omniscient and maintain globally consistent belief sets.

Inconsistency does not necessarily persist indefinitely. Agents may revise their beliefs in response to new evidence and restore consistency without ever being aware that their beliefs were initially inconsistent. This is evidenced by interaction-based ontology alignment repair [1]. In this setting, agents each classify objects following their own ontologies. They rely on alignments between their ontology and that of their interlocutor to interact, and revise these alignments when communication fails. Experimental results showed that, under the relaxation modality which ensures that an agent eventually makes use of all their alignments, agents always converge to fully consistent states.

This shows that, even in the absence of complete reasoning capabilities, iterative interaction and revision may be sufficient to eliminate inconsistency over time.

Interaction-based ontology repair has been modelled in a logical framework [2]. However, the resulting formalisation only applies to agents with consistent beliefs. Can this restriction be lifted? One possible approach to modelling inconsistent beliefs is to use context-dependent or non-monotonic logics. However, the two examples presented above do not involve situations in which truth depends

LINDA Workshop, July 24, 2026, Lisbon, Portugal
linda.gutsche@inria.fr (L. Gutsche)

on context or in which agents withdraw conclusions in response to new information. Instead, the issue lies in the agents' failure to derive the logical consequences that would reveal the inconsistency. We therefore aim to develop a multi-agent logic of belief that can represent agents with incomplete and potentially incorrect deductive abilities.

Starting with the Logic of Doxastic Attitudes [3], an existing multi-agent logic that distinguishes explicit beliefs (those stored in a belief base) and implicit beliefs (those derivable from it), we defined a semantics for multi-agent logics of explicit and implicit beliefs that can model agents with incorrect or incomplete reasoning. Our semantics integrates the deductive abilities of agents directly into the semantic structure by representing them as closure functions that may be logically incomplete. This openly separates the beliefs that an agent explicitly holds from their inference process, and provides a flexible framework for modelling agents whose reasoning is not necessarily complete nor correct with respect to the underlying logic. This approach captures both logically ideal agents and agents with imperfect deductive abilities, and supports heterogeneous reasoning within a unified multi-agent framework.

Using the Logic of Doxastic Attitudes as a basis for comparison with standard Kripke semantics, we showed that standard Kripke semantics correspond exactly to the special case in which agents are deductively complete: only such agents can be faithfully modelled using Kripke structures, and conversely, any situation involving only complete deductive abilities admits a Kripke representation.

In the remainder, we discuss some related works (§2): belief revision beyond classical consistency assumptions (§2.1), the Logic of Doxastic Attitudes (§2.2) which our work generalises, and logical frameworks modelling agents with not necessarily sound and complete deductive abilities (§2.3). We then introduce our formalism for logically incomplete agents by defining the language of explicit and implicit beliefs and the closure-based semantics (§3). We present a result showing that this closure-based semantics generalises Kripke semantics (§4), and conclude with an example illustrating how it can represent inconsistency tolerant agents (§5).

2. Related Work

Doxastic and epistemic logics are logical formalisms for expressing the beliefs and knowledge of agents. They model multi-agent settings and higher order beliefs. From Hintikka's model systems to Kripke structures, their usual semantics are based on a collection of possible worlds and accessibility relations between them [4, 5]. Informally, these relations capture which scenarios an agent considers possible.

2.1. Belief Revision under Inconsistent Beliefs

Belief revision studies how an agent should modify their beliefs when presented with new information, assumed to be true, that conflicts with their current beliefs [6, 7]. Classical belief revision, however, presupposes that such a conflict is detected. Moreover, it does not address situations in which an agent's belief state is already inconsistent.

Several extensions of belief revision relax these assumptions. Relevancy-based compartmentalisation only requires agents to maintain local consistency [8], while resource-bounded belief revision accounts for limitations in agents' reasoning capabilities [9]. Paraconsistent logics allow agents to reason coherently in the presence of inconsistency by blocking explosive inference [10], and paraconsistent belief revision studies rational belief change in such settings [11].

In this work, we take the approach of representing an agent's beliefs as the combination of a belief base and possibly incomplete deductive abilities, allowing inconsistencies to arise from the limitations in the agent's deductive abilities.

2.2. The Logic of Doxastic Attitudes: An Epistemic Logic with Belief Bases

The Logic of Doxastic Attitudes integrates the distinction between explicit and implicit beliefs into a multi-agent doxastic logic framework [12, 3]. It defines explicit and implicit belief operators, and interprets them by incorporating the notion of belief bases in its semantics. Its language and semantics are defined in two layers: a base layer for explicit beliefs, and an upper layer adding implicit beliefs. A world state consists of a belief base for each agent and a valuation representing the environment. Explicit beliefs are interpreted at the level of states: an agent explicitly believes the formulas contained in their belief base. A state is an epistemic alternative for an agent if it satisfies all their explicit beliefs. The full language, adding implicit beliefs, is then interpreted over a structure consisting of an actual state together with a set of possible states. An agent implicitly believes all formulas that hold true in all the possible states that they consider as epistemic alternatives. The Logic of Doxastic Attitudes has been extended with belief change operations, such as revision under constraint [13], and collective expansion through announcement [14].

The explicit and implicit belief operators used in the Logic of Doxastic Attitudes are particularly suited for our purpose of representing an agent's beliefs through their explicitly stored beliefs and their deductive abilities. We thus use this logic as the starting point of our generalisation, and as our basis for comparison with standard Kripke semantics.

2.3. Logical Representations of Agents with Imperfect Deduction

We aim to model agents with possibly incomplete or incorrect reasoning abilities. One way to model imperfect agents is the direct introduction of deduction functions in the semantics. They have already been considered as a solution to the omniscience problem.

The Deductive Model of Belief [15] interprets the beliefs of agents with a deductive system consisting of a set of base beliefs and deduction rules. Agents are assumed to have sound but incomplete reasoning abilities. An agent believes a formula if it is derivable in their deductive system: no distinction is made between explicit and implicit beliefs at the syntactic level. Through comparisons of the logical systems, the Deductive Model of Belief has been proven to be reducible to the classical doxastic logic based on Kripke structures when restricted to logically complete belief deduction rules [15].

However, this model assumes that all agents are sound reasoners.

Other frameworks consider agents whose beliefs are closed under logical consequence, but can contain more than what is logically entailed by the agent's certainties. It is the case for doxastic epistemic logics [5]. For instance, in those using plausibility-based semantics [16, 17], agents believe all formulas that hold true in all the worlds that they consider most plausible. These agents' beliefs can thus be considered as following from not only logical consequence but also additional assumptions.

In order to represent the variety of imperfect belief reasoners, we extend the Logic of Doxastic Attitudes by introducing a semantics that interprets the reasoning abilities of agents with deductive functions.

3. Closure-Based Semantics for Explicit and Implicit Beliefs

We define the language of explicit and implicit beliefs by introducing simultaneously the modal operators for explicit and implicit beliefs of the Logic of Doxastic Attitudes.

Let $\mathscr{P}$ and $\mathscr{A}$ respectively denote a countably infinite set of atomic propositions and a finite set of agents. $\triangle_a\varphi$ reads "agent a explicitly believes φ" and $\Box_a\varphi$ reads "agent a implicitly believes φ".

The language $\mathscr{L}$ of explicit and implicit beliefs is defined by

$$\mathscr{L} \overset{def}{=} \varphi ::= p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \triangle_a\varphi \mid \Box_a\varphi$$

such that $p \in \mathscr{P}$ and $a \in \mathscr{A}$.

The classical boolean connectives $\top$, $\bot$, $\vee$, $\Rightarrow$ and $\Leftrightarrow$ are defined in the standard way.

We now introduce a semantics that allows us to model agents with incomplete, incorrect, or heterogeneous reasoning abilities.

Explicit beliefs and atomic propositions are interpreted on the state of the world, which is composed of both the external environment and the agents' belief bases.

Definition 1 (State)**.** *A state s is a tuple $\langle (B_a^s)_{a\in\mathcal{A}}, V^s\rangle$ where $B_a^s \subseteq \mathcal{L}$ is agent a's belief base, and $V^s \subseteq \mathcal{P}$ is a valuation representing the environment at s.*

The set of all states is denoted **S**.

In the Logic of Doxastic Attitudes, implicit beliefs are determined by accessible worlds. Here, we instead represent them using deductive functions Cn_a, where $Cn_a(X)$ is the set of all formulas agent a can infer from the information set $X \subseteq \mathcal{L}$.

Since $Cn_a(X)$ represents what a can infer from X and deduction typically involves closure, we assume that each Cn_a is a closure function on $\mathcal{L}$.

Definition 2 (Closure function)**.** *A function $Cn : 2^{\mathcal{L}} \to 2^{\mathcal{L}}$ is a closure function if it verifies the three following properties:*

1. *Extensivity:* $\forall X \subseteq \mathcal{L},\ X \subseteq Cn(X)$
2. *Idempotence:* $\forall X \subseteq \mathcal{L},\ Cn(Cn(X)) = Cn(X)$
3. *Monotonicity:* $\forall X, X' \subseteq \mathcal{L},\ X \subseteq X' \ \Rightarrow\ Cn(X) \subseteq Cn(X')$

The closure functions used in the semantics are not further constrained. They can be defined through inference rules, or through any other means.

A closure family assigns such a function to each agent.

Definition 3 (Closure family)**.** *A closure family is a collection $Cn^{\mathcal{A}} = (Cn_a)_{a\in\mathcal{A}}$ of closure functions on $\mathcal{L}$, one for each agent $a \in \mathcal{A}$.*

Representing a's implicit beliefs with a closure function Cn_a, we say that agent a can derive a formula φ from its belief base B_a if and only if $\varphi \in Cn_a(B_a)$.

Formulas of the language of explicit and implicit beliefs are interpreted on pairs $\langle Cn^{\mathcal{A}}, s\rangle$ composed of a closure family $Cn^{\mathcal{A}}$ and a state $s = \langle (B_a^s)_{a\in\mathcal{A}}, V^s\rangle$. Such pairs are called *pointed closure families*.

Definition 4 (Satisfaction relation)**.** *Let $Cn^{\mathcal{A}}$ be a closure family and s a state.*

$$
\begin{array}{rll}
\langle Cn^{\mathcal{A}}, s\rangle \models p & \textit{iff} & p \in V^s \\
\langle Cn^{\mathcal{A}}, s\rangle \models \neg\varphi & \textit{iff} & \langle Cn^{\mathcal{A}}, s\rangle \not\models \varphi \\
\langle Cn^{\mathcal{A}}, s\rangle \models \varphi_1 \wedge \varphi_2 & \textit{iff} & \langle Cn^{\mathcal{A}}, s\rangle \models \varphi_1 \textit{ and } \langle Cn^{\mathcal{A}}, s\rangle \models \varphi_2 \\
\langle Cn^{\mathcal{A}}, s\rangle \models \triangle_a\varphi & \textit{iff} & \varphi \in B_a^s \\
\langle Cn^{\mathcal{A}}, s\rangle \models \Box_a\varphi & \textit{iff} & \forall a \in A : \Box_a\varphi \in Cn_a(B_a^s)
\end{array}
$$

4. Comparison with Kripke Semantics

We introduced a closure-based semantics for the language of explicit and implicit beliefs, using closure functions to represent the deductive abilities of agents. This semantics generalises Kripke semantics in the following sense: Kripke models capture exactly the class of agents whose implicit beliefs are closed under a deductive system that is sound and complete for the language of explicit and implicit beliefs under Kripke semantics. Such a deductive system is defined by the following axiom schemes and inference rules:

- All propositional tautologies
- Distribution (K) axiom for $\Box_a$: $\Box_a(\varphi \Rightarrow \psi) \Rightarrow (\Box_a\varphi \Rightarrow \Box_a\psi)$
- Modus ponens: from φ and $\varphi \Rightarrow \psi$, infer ψ
- Necessitation of $\Box_a$: from φ infer $\Box_a\varphi$
- Interaction axiom between $\triangle_a$ and $\Box_a$: $\triangle_a\varphi \Rightarrow \Box_a\varphi$

This comparison result also holds for a richer language that includes additional explicit and belief operators from the Logic of Doxastics Attitudes, such as common and shared belief operators.

$$
\begin{tikzcd}
C_{A,\ cheap} \arrow[r, leftrightarrow, "\perp"] \arrow[drr, dotted, dash, "\subseteq"'] & C_{A,\ expensive} \arrow[dr, dotted, dash, "\supseteq"] & \\
r \arrow[u, "\in"'] & & C_{B,\ cheap}
\end{tikzcd}
$$

Figure 1: Visual representation of the example of inconsistent beliefs. $\perp$ stands for $C_{A,\ cheap} \cap C_{A,\ expensive} = \emptyset$

5. Illustration and Perspectives

The closure-based semantics can model agents whose implicit beliefs are not closed under classical logical consequence. This makes it possibly to capture different forms of incomplete or selective reasoning. We illustrate this on a locally inconsistent scenario of ontology alignment (Figure 1).

Consider two agents, Alice and Bob, who classify all the restaurants of their city according to their own perspective. Let $C_{A,\ cheap}$ and $C_{A,\ expensive}$ represent the two classes in which Alice categorises restaurants according to how expensive she finds them. Let $C_{B,\ cheap}$ represent Bob's class of cheap restaurants. To keep the example concise, let us suppose that Alice and Bob are currently discussing one restaurant in particular. Alice finds the restaurant cheap, which we represent as $\Box_a c_{A,\ cheap}$. She considers that a restaurant cannot be both cheap and expensive: $\Box_a \neg(c_{A,\ cheap} \wedge c_{A,\ expensive})$. She believes that whenever she finds a restaurant to be cheap, Bob finds it cheap also: $\Box_a(c_{A,\ cheap} \Rightarrow c_{B,\ cheap})$. From experience, Alice believes that every time Bob describes a restaurant as cheap, she actually finds it to be expensive: $\Box_a(c_{B,\ cheap} \Rightarrow c_{A,\ expensive})$

These four beliefs of Alice are logically inconsistent. However, the closure-based semantics can still represent this situation. For example, Alice may reason by solely using modus ponens and the transitivity of $\Rightarrow$, with the restriction that she applies transitivity only when one of the two formulas uses only the vocabulary of one agent. This is natural in a situation where she uses alignments only to either classify a restaurant in one of Bob's classes or interpret his classifications in her own vocabulary: she never encounters a situation where she needs to reason with both the alignments towards and from Bob's ontology. If such a situation were however to arise, it could be represented as a change in her deductive abilities, modelled as Alice adopting a new closure function and updating her beliefs to resolve any inconsistencies she can now detect.

In general, closure functions generated from inference rules can be defined by limiting the agent to a subset of a general collection of inference rules, or by restricting the formulas on which the agent can apply each rules schemas. Such restriction can be done syntactically, but also structurally. For example, following [18], an agent's reasoning can be organised around a set of questions, with the agent only applying an inference rule if both the premises and the result are relevant for one same question.

Suppose that, in the example, Alice is only concerned with the questions Q_{rA} and Q_{rB} corresponding to respectively "which classes of Alice / Bob contain the restaurant or a class of Alice containing the restaurant?". Reasoning under Q_{rA} leads to Alice deducing that $r \notin C_{A,\ expensive}$ as she leverages the partial answers $r \in C_{A,\ cheap}$ and $C_{A,\ cheap} \perp C_{A,\ expensive}$. Similarly, reasoning under Q_{rB} leads Alice to deduce that $r \in C_{B,\ cheap}$. In fact, this is the only new information she derives from the relations illustrated in Figure 1. She therefore does not detect the inconsistency.

We defined a semantics for the language of explicit and implicit beliefs, interpreting the agents' deductive abilities with closure functions. This semantics separates the representation of stored beliefs from the representation of the reasoning process, making it possible to model agents with heterogeneous deductive abilities, including ones that may be incomplete or unsound.

Although we chose to use closure functions, the semantics could be further generalised to arbitrary functions.

For future work, we plan to study belief revision operators for agents with incomplete deductive abilities. We aim to investigate under which conditions repeated interaction between such agents enables them to converge towards fully consistent states.

References

[1] J. Euzenat, Interaction-based ontology alignment repair with expansion and relaxation, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 2017, pp. 185–191.

[2] L. van den Berg, M. Atencia, J. Euzenat, A logical model for the ontology alignment repair game, Autonomous Agents and Multi-Agent Systems 35 (2021) 1–32.

[3] E. Lorini, Rethinking epistemic logic with belief bases, Artificial Intelligence 282 (2020).

[4] K. J. J. Hintikka, Knowledge and Belief: An Introduction to the Logic of the Two Notions, Cornell University Press, Ithaca, NY, USA, 1962.

[5] H. van Ditmarsch, W. van der Hoek, J. Y. Halpern, B. Kooi (Eds.), Handbook of Epistemic Logic, College Publications, 2015.

[6] C. E. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet contraction and revision functions, The Journal of Symbolic Logic 50 (1985) 510–530.

[7] E. Fermé, S. O. Hansson, Belief Change: Introduction and Overview, Springer Verlag, Cham, 2018.

[8] S. O. Hansson, R. Wassermann, Local change, Studia Logica 70 (2002) 49–76.

[9] R. Wassermann, Resource bounded belief revision, Erkenntnis 50 (1999) 429–446.

[10] G. Priest, In Contradiction: A Study of the Transconsistent, Oxford University Press, New York, 2006.

[11] N. Schwind, S. Konieczny, R. P. Pérez, On paraconsistent belief revision in LP, in: AAAI Conference on Artificial Intelligence, 2022.

[12] E. Lorini, In Praise of Belief Bases: Doing Epistemic Logic without Possible Worlds, in: Proceedings of the 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, 2018, pp. 1915–1922.

[13] E. Lorini, F. Schwarzentruber, Multi-agent belief base revision, in: Proceedings of the 30th International Joint Conference on Artificial Intelligence, IJCAI 2021, 2021, pp. 1959–1965.

[14] E. Lorini, E. Rapion, Logical Theories of Collective Attitudes and the Belief Base Perspective, in: Proceedings of the 21st International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2022, ACM, 2022, pp. 833–841.

[15] K. G. Konolige, A deduction model of belief and its logics, Ph.D. thesis, Stanford University, 1984.

[16] A. Baltag, S. Smets, Dynamic belief revision over multi-agent plausibility models, Proceedings of LOFT (2006).

[17] A. Baltag, S. Smets, A Qualitative Theory of Dynamic Interactive Belief Revision, Springer International Publishing, Cham, 2016, pp. 813–858.

[18] D. Hoek, Minimal rationality and the web of questions, in: P. van Elswyk, D. Kindermann, C. D. Kirk-Giannini, A. Egan (Eds.), Unstructured Content, Oxford University Press, 2025.