Annals of Mathematics and Artificial Intelligence
https://doi.org/10.1007/s10472-023-09902-w

# Modeling and shadowing paraconsistent BDI agents

**Barbara Dunin-Kęplicz[1,2] · Andrzej Szałas[1,3]**

Accepted: 6 October 2023


**Abstract**

The BDI model of rational agency has been studied for over three decades. Many robust multiagent systems have been developed, and a number of BDI logics have been studied. Following this intensive development phase, the importance of integrating BDI models with inconsistency handling and revision theory have been emphasized. There is also a demand for a tighter connection between BDI-based implementations and BDI logics. In this paper, we address these postulates by introducing a novel, paraconsistent logical BDI model close to implementation, with building blocks that can be represented as SQL/rule-based databases. Importantly, tractability is achieved by reasoning as querying. This stands in a sharp contrast to the high complexity of known BDI logics. We also extend belief shadowing, a shallow and lightweight alternative to deep and computationally demanding belief revision, to encompass agents' motivational attitudes.

**Keywords** Beliefs-Desires-Intentions models · Paraconsistent reasoning · Doxastic reasoning · Shadowing · Reasoning by querying

**Mathematics Subject Classification (2010)** 93A16 · 68T42 · 03B42 · 03B50 · 03B53 · 68T27

## 1 Introduction

The BDI model of agency has been intensively discussed over years, starting from the seminal research of Bratman [8], and creatively developed by many others, just to mention some of them [7, 9, 11, 14, 15, 18, 24, 25, 32, 35, 37, 41, 46]. Taking into account the experience both from the implemented systems as well as extensive formal studies, it is now time to pay closer attention to a more computationally friendly approach to modeling and practical reasoning about BDI agents. Tackling these agents within a logical framework, while employing a

✉ Andrzej Szałas
andrzej.szalas@mimuw.edu.pl

Barbara Dunin-Keplicz
keplicz@mimuw.edu.pl

1 Institute of Informatics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

2 QED Software, Warsaw, Poland

3 Department of Computer and Information Science, Linköping University, SE-581 83 Linköping, Sweden

Published online: 23 November 2023

database perspective, has been presented in [42]. Beliefs and intentions are organized there in two temporal databases providing means for further reasoning. Within the realm of classical truth values $\mathbb{t}$ (true) and $\mathbb{f}$ (false), beliefs and intentions are required there to be internally consistent, at the same time preserving the consistency of intentions with beliefs.

A similar in spirit shift in perspective is presented in [17], where instead of reasoning about beliefs in modal logics or other high-complexity formalisms, a tractable approach based on querying belief bases has been advocated for. That is, rather than being concerned with general properties that apply to all models, agents are seen to be primarily concerned with how to act in the specific world (model) they are embedded in. For such purposes reasoning-by-querying is more appropriate. As realistic world models are generally acquired from multiple information sources, they typically include inconsistencies and/or gaps of knowledge. To ensure the required expressiveness and modeling convenience, two additional truth values are adopted: $\mathbb{i}$ (inconsistent), and $\mathbb{u}$ (unknown). This adds the dimensions of paraconsistent and paracomplete reasoning, that are central to our approach.

Real-world AI applications often need agents to switch between different roles or groups. A well-controlled and computationally-friendly adjustment of relevant beliefs is required, particularly when the structures and organizations evolve dynamically. In such volatile circumstances, a full-fledged reasoning about beliefs is too demanding. Instead, a shallow belief change may be realized by *belief shadowing* [5]. Since beliefs influence agents' motivational attitudes, shadowing needs to be extended to include goals and intentions. Our formalism covers this case: when a belief (respectively, desire or intention) base $\mathcal{B}$ of an agent is shadowed by a belief (desire, intention) base $\mathcal{B}$', the latter becomes dominant and the agent behaves according to $\mathcal{B}$' unless $\mathcal{B}$' has no established attitude towards a given matter in which case the agent acts according to $\mathcal{B}$.[1]

In the context of BDI systems, in [25] the authors postulated to provide:

- a logical semantics underlying the implementations;
- a link between logical approach and revision theory;
- a paraconsistent reasoning about desires,

while stressing the importance of tractability: "one reason explaining the gap between theory and practice is the too high complexity of existing logics." In this paper we address these postulates and ensure the desideratum of tractability by reporting original developments that include:

- a paraconsistent and paracomplete approach to beliefs, desires and intention;
- shadowing mechanism: a lightweight form of transient revision of BDI attitudes;
- an intuitive and tractable formalism used in querying and shadowing BDI bases.

To our best knowledge, no paraconsistent and paracomplete approach to BDI systems has been provided so far. Further, a tractable BDI revision/update technique has not yet been developed. Our shadowing mechanism is the first that achieves this complexity level for full BDI. In addition, no other formalism dealing with beliefs, desires and intentions provides a tractable reasoning machinery that can be implemented and used in real-world scenarios. In order to distinguish it from traditional BDI models and emphasize its four-valued nature, the introduced formalism is referred to as 4BDI.

The rest of the paper is structured as follows. In Section 2, we discuss our approach to modeling BDI agents. In Section 3 we present the intended architecture of 4BDI systems.

[1] The term *shadowing* is used due to an analogy to shadowing in modular programming, where entities defined in the inner scope shadow entities with the same name, occurring in the outer scope.

Section 4 introduces the language for reasoning about beliefs, desires and intentions. Section 5 is devoted to semantical structures underlying 4BDI. In Section 6 belief shadowing is extended to desires and intentions. Complexity of the approach is addressed in Section 7, showing tractability of 4BDI querying and shadowing wrt size of the databases involved. Section 8 summarizes our design choices while Sections 9 and 10 present related work and conclusions.

## 2 Modeling BDI agents

BDI model of rational agency combines agents informational and motivational attitudes: beliefs, intentions and desires. Intuitively, beliefs represent information an agent has about the world, including environment, other agents and itself. If choosing a potentially elaborated and nuanced logical representation of beliefs, typically in dedicated multimodal logics, we know pretty well how to reason about them. But this comes at a price of high complexity.

On the other hand, desires stand for those states of affairs the agent wishes to bring about. Finally, intentions represent the chosen desires that it has committed to achieve. This part - a deliberation process - aims at figuring out *what* to do next. The next phase of deciding *how* to achieve it is called means-end-reasoning or planning. The entire process leading from an initial setting of beliefs and desires to the relevant plan is a subject of practical reasoning [8]: the process of deciding, step by step, which action to perform next. To put it simply, it consists of:

- the *deliberation phase* of choosing the proper intention to focus on for the time being;
- the *planning phase*, resulting in a specific plan that (most of the time) is chosen from the precompiled library of plan skeletons.

Following [46]: "[...] a fundamental problem in developing such a theory of rational agency is to give an account of the relationships that exist between an agents' mental states. In particular, a complete agent theory would explain how an agent's mental states lead it to select and perform rational actions, thus realizing the mapping from perception to action."

To approach this goal, instead of playing with a family of multimodal logics, we decided to follow the database perspective: modeling beliefs, desires, and intentions requires appropriate semantical structures that are simple to use and implement. Following this line of modeling, the next question is how to effectively represent potentially nuanced deliberation processes? In order to make or solution generic, we decided to encapsulate them in so-called bridging functions between various 4BDI components. These, potentially complex and application-specific functions, are supposed to reflect the nature of the particular transitions. Finally, relevant plans are selected from a precompiled plan library.

Let us first discuss the components informally before providing rigorous definitions.

The basic building blocks of 4BDI are *ground literals* (possibly negated atomic formulas without variables) used to construct:

- *world aspects*: finite sets of ground literals;
- *clusters*: finite sets of world aspects.

**Remark 1** World aspects represent perspectives on the world as perceived by heterogeneous information sources, like sensors, classifiers, and so on. When compared to the worlds in Kripke structures, aspects can additionally represent complementary views.

They can be implemented as rule-based/SQL-like databases. The aspects may also be computed as models of rule-based languages, such as variants of DATALOG [1] (two-valued models), ASP [22, 23] (three-valued models with the third truth value $\mathtt{u}$), paraconsistent 4QL [34] (four-valued models, additionally with $\mathtt{i}$), etc. □

To illustrate the 4BDI model we will use a running example.

**Example 1** (Running example) Eve is a medicine student who has to complete an anatomy course. To pass the course she needs to pass a test, and give a presentation or write an essay. She enjoys traveling but being committed to the course, she resigns from traveling for the time being. All in all, the related part of her belief base contains the following world aspects:

$$\{prepare(pres), \neg prepare(essay), pass(test), \neg travel\}, \tag{1}$$

$$\{\neg prepare(pres), prepare(essay), pass(test), \neg travel\}. \tag{2}$$

Each world aspect, (1) and (2), is internally consistent. However, at the level of cluster consisting of (1) and (2) the literals *prepare*(*pres*) and *prepare*(*essay*) are inconsistent until Eve makes her choice and, in effect, one of the aspects is removed. □

In general, in 4BDI structures, inconsistencies can occur at two structural levels:

- inside a single world aspect, when it contains both a literal and its negation;
- inside a single cluster, when different aspects contain contradictory literals.

Furthermore, information gathered is frequently incomplete, what calls for introducing means for paraconsistent and paracomplete reasoning.

Clusters provide an abstraction for databases used in 4BDI frames and structures. As shown in Fig. 1, 4BDI frames and structures include the following components:

- *belief base* ($\mathcal{B}$) containing aspects of the current world, possibly obtained form multiple information sources;
- *desire base* ($\mathcal{D}$) containing agent's desired world aspects;
- *intention base* ($\mathcal{I}$) containing world aspects the agent has chosen to accomplish;
- *goal-desires* bridging function ($\mathcal{G}$) used to determine desires;
- *desires-intentions* bridging function ($\mathcal{A}$) used to determine intentions.

Notice that we assume that *goals* are delivered externally and specified by formulas. It is important to distinguish externally delivered goals from internally generated desires.[2] For example, when a team of agents is formed to accomplish a particular mission, one has to supply goals to be achieved during the mission. Then such external goals are transformed into agents' internal desires (options) and, further, to intentions.

## 3 The architecture of 4BDI systems

The intended architecture of 4BDI systems is shown in Fig. 2. We assume that a system designer provides bridging functions $\mathcal{G}$ and $\mathcal{A}$ tailored to a specific application domain. When a goal and a belief base are provided, these functions return a desire base, followed by an intention base. This way, a frame enables the creation of a 4BDI structure containing belief, desire, and intention bases. Both functions, defined by a user, encapsulate a number of database related strategies to establish desires and, especially, intentions. On the other hand, other strategies ensuring persistence of achieving intentions, and reconsideration methods, are realized in the main body of an agent. From the agents' perspective, the 4BDI framework provides 4BDI structure construction with query answering services.

[2] In the literature, desires are also sometimes called goals. However, in opposite to goals (desires) being an internal part of BDI models, when we refer to goals in 4BDI, we mean external goals, supplied from outside, e.g., by a mission coordinator.

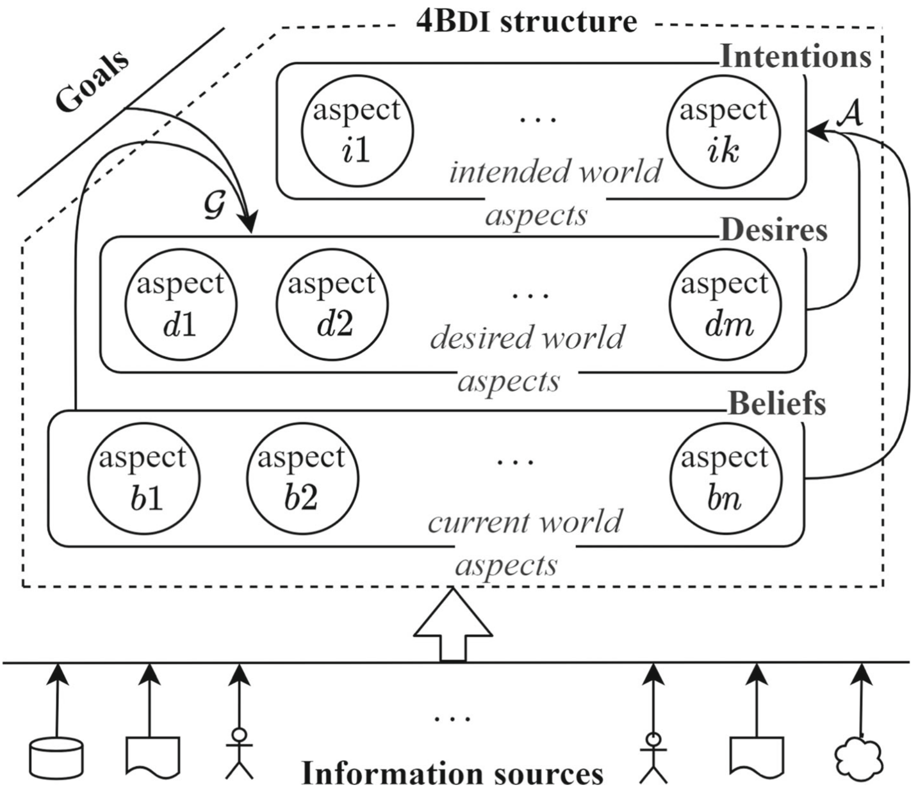


**Fig. 1** High level architecture of 4BDI structures. Rounded arrows represent goal-desires ($\mathcal{G}$) and desires-intentions ($\mathcal{A}$) bridging functions

Notice that in Fig. 2 the 4BDI model, encompassed in a dashed line box, is embedded in a larger agent's architecture with the following components external to 4BDI:

- *Environment*: a natural, typically largely uncontrolled, environment the agents operate within;
- *Perception*: a layer through which agents observe the environment and establish their beliefs about it, consisting of hardware ingredients like sensors, cameras, etc., as well as software components like classifiers, noise filters, etc.;
- *Information sources*: subjects, like agents, people, databases, providing the agent with supplementary knowledge/beliefs useful for completing a given mission, e.g., related to coordination, task allocation, etc.;
- *Goal*: a goal to be achieved, supplied externally, and then transformed into internal 4BDI desires;
- *Reconsideration, plan selection*: functionalities related to monitoring and executing plans, where replanning may be necessary in response to newly acquired information and corresponding updates to 4BDI components;
- *Action*: an action selected for execution, possibly affecting the environment.

In 4BDI belief, desire and intention bases are relatively independent of other components of 4BDI agents: only query and update interfaces serve the purpose of interaction between them.

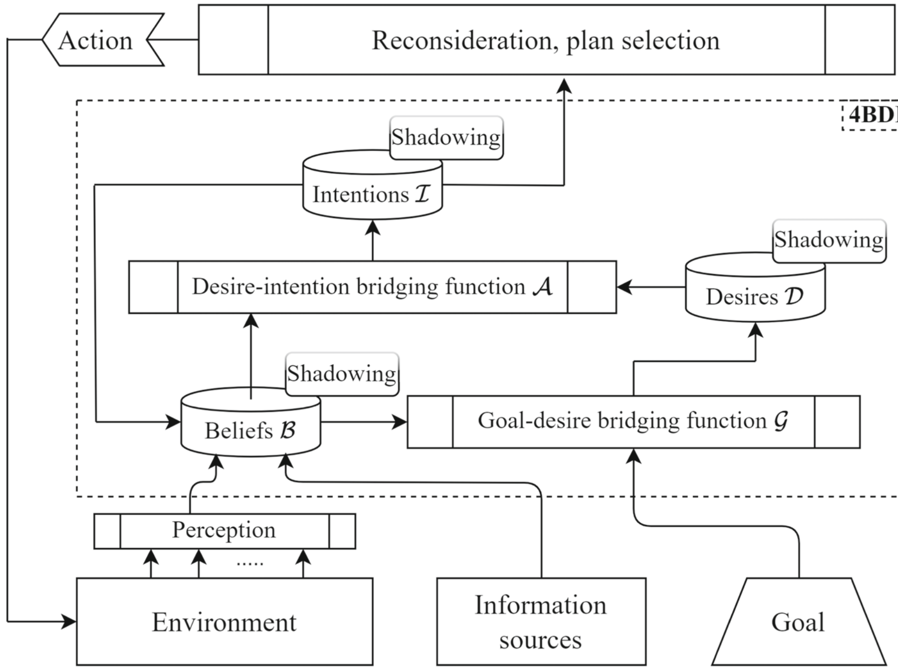


**Fig. 2** The architecture of 4BDI-based agents

This independence assumption is similar to the modular architecture of [37]. The capabilities component of [37] which "serves as a technical shortcut for encapsulating implementation of agent's capabilities wrt its environment", in our architecture is encapsulated in desires-intentions and possibly goal-desires bridging function. In the literature desires-intentions bridging function is sometimes called a *filter*. The BDI standard *option generation function* corresponds to our goal-desires bridging function, except that in our approach an external goal also participates in the desire generation process.

The 4BDI agent's loop is shown in Fig. 3: first a goal for the agent is set. Next, the agent uses a predefined frame to determine desires and intentions based on its beliefs and goal. The loop restarts whenever the goal or relevant beliefs are updated.

## 4 The 4BDI language

Let *Ag* be a finite set of *agent names*, uniquely identifying agents. We will use the syntax of classical first-order logic extended by operators $\mathbb{B}_a(A)$ ("$A$ is in $a$'s beliefs"), $\mathbb{D}_a(A)$ ("$A$ is desired by $a$") and $\mathbb{I}_a(A)$ ("$A$ is intended by $a$").

Let us now define the signature and language of the considered logic.

**Definition 1** (Signature) Let $\mathcal{P}$ and $\mathcal{C}$ be finite sets of relation symbols and constants.[3] The pair $\Sigma = \langle \mathcal{P}, \mathcal{C} \rangle$ is called the language *signature*. □

In the rest of the paper, the set of first-order variables is denoted by $\mathcal{V}$.

[3] As standard in query languages, we exclude function symbols.

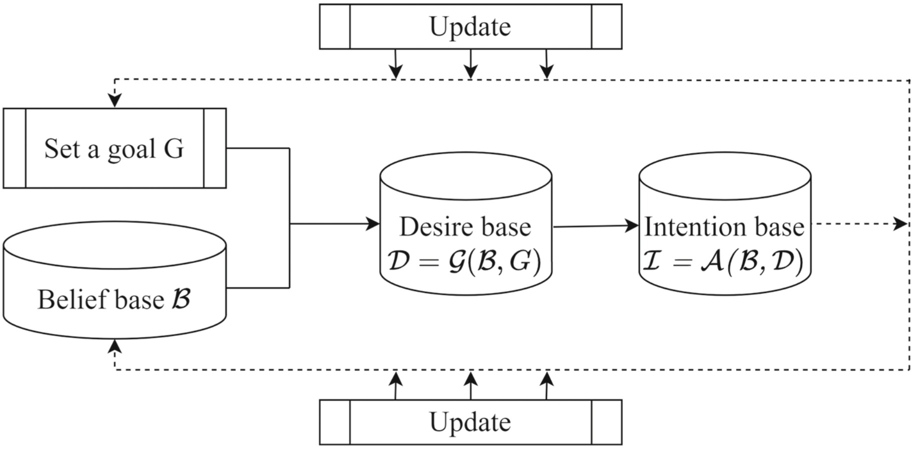


**Fig. 3** The loop of 4BDI agents

**Definition 2** (Formulas, literals, ground literals) The following BNF grammar defines syntax of formulas, where $\langle\mathcal{T}\rangle$ and $\langle\mathcal{F}\rangle$ are nonterminal symbols denoting truth values and formulas, respectively. Let $p \in \mathcal{P}$ be a $k$-argument relation symbol, $\overline{arg} \in (\mathcal{C} \cup \mathcal{V})^k$ be a tuple of constants and/or variables, $x \in \mathcal{V}$ be a variable, and $a \in Ag$ be an agent name. The formulas are defined by:

$$\langle\mathcal{F}\rangle ::= \ \langle\mathcal{T}\rangle \mid p(\overline{arg}) \mid \neg\langle\mathcal{F}\rangle \mid \langle\mathcal{F}\rangle \wedge \langle\mathcal{F}\rangle \mid \langle\mathcal{F}\rangle \vee \langle\mathcal{F}\rangle \mid \tag{3}$$

$$\forall x\big(\langle\mathcal{F}\rangle\big) \mid \exists x\big(\langle\mathcal{F}\rangle\big) \mid \tag{4}$$

$$\mathbb{B}_a\big(\langle\mathcal{F}\rangle\big) \mid \mathbb{D}_a\big(\langle\mathcal{F}\rangle\big) \mid \mathbb{I}_a\big(\langle\mathcal{F}\rangle\big) \tag{5}$$

The set of all formulas is denoted by $\mathcal{F}$. Formulas of the form $p(\overline{arg})$, $\neg p(\overline{arg})$ are called *literals* (respectively, *positive* and *negative* ones). The literals with $\overline{arg}$ containing only constants are called *ground*. □

**Definition 3** (Plain formulas) Formulas defined by the grammar specified in lines (3)–(4) are called *plain*. The set of plain formulas is denoted by $\mathcal{F}^I$.

By a *free occurrence of a variable x in a formula A* we mean an occurrence of $x$ outside of the scope of any quantifier, $\forall x$, $\exists x$, binding $x$. A formula without free variables is called *closed*. □

**Definition 4** (4BDI formulas) A 4BDI *formula* is any formula of the form $\mathbb{O}_a(A)$, where $\mathbb{O} \in \{\mathbb{B}, \mathbb{D}, \mathbb{I}\}$, $a \in Ag$ and $A \in \mathcal{F}$. The set of all 4BDI formulas is denoted by $\mathcal{F}^{bdi}$. □

All formulas under consideration are assumed to be closed. Since we deal with finite domains, generality is not lost in this manner. Rather than computing queries as sets of tuples satisfying the formula, we verify the truth value of formulas by substituting constants for free variables. This simplifies definitions without influencing the complexity classes of queries [1].

## 5 Semantics of 4Bdi

### 5.1 Semantics of plain formulas

To define the semantics of plain formulas we use a four-valued logic with truth values and designated truth values (see, e.g., [44]) defined as follows.

**Definition 5** (Truth values, designated truth values) We assume the set of truth values $\mathcal{T} \stackrel{\text{def}}{=} \{\mathtt{t}, \mathtt{f}, \mathtt{i}, \mathtt{u}\}$ ordered by $\leq_l$ shown in Fig. 4. The set of *designated truth values* that act as true, denoted by $\mathcal{T}^d$, is a set $\mathcal{T}^d \subset \mathcal{T}$ such that $\mathtt{t} \in \mathcal{T}^d$ and $\mathtt{f} \notin \mathcal{T}^d$. □

In many-valued logics one separates the set $T^d$ of designated truth values to accept arguments and conclusions being "close enough to truth" or "useful enough" to participate in reasoning and decision making, as clarified in the context of 4Bdi in the following remark.

**Remark 2** Note that the set $\mathcal{T}^d$ permits to accept rational intentions that are not fully logically justified, i.e., not necessarily being true. Therefore, $\mathcal{T}^d$ can be used to fine-tune 4Bdi in this respect. For example, there may be no exact plan to accomplish a goal "save a victim after a disaster" by making it true. However, there may be a plan whose effect on saving the victim is inconsistent or unknown. It is rational to carry out such a plan rather than doing nothing. In 4Bdi this can be achieved by setting $\mathcal{T}^d = \{\mathtt{t}, \mathtt{i}, \mathtt{u}\}$. Depending on the strategy, one could set $\mathcal{T}^d$ to be $\{\mathtt{t}, \mathtt{i}\}$, $\{\mathtt{t}, \mathtt{u}\}$, or remain with $\{\mathtt{t}\}$. □

The semantics of connectives and operators will be defined using orderings on truth values. When multiple information sources are present, two orderings on truth values are typically considered to address distinct needs:

- *truth ordering* for evaluating the truth value of formulas to reflect their "truth contents" (as traditionally done in logics);
- *information ordering* for fusing information from multiple sources.

Sample orderings on the values $\mathtt{t}, \mathtt{f}, \mathtt{i}, \mathtt{u}$, are shown in Fig. 4. The ordering $\leq_i$ is commonly used in modeling knowledge gathering:

- initially there is no knowledge about a given fact what is represented by $\mathtt{u}$;
- then, evidence for the fact or against it, making it $\mathtt{t}$ or $\mathtt{f}$;
- finally, evidence both for and against the fact may be collected resulting in $\mathtt{i}$.

In paraconsistent reasoning, different truth orderings are chosen depending on the application area. While the most popular is perhaps $\leq_t$ [2, 3], some other approaches advocate for $\leq_l$ [13, 34]. Following [13, 34], here we choose $\leq_l$ since it provides more intuitive results in the intended application domains than other orderings. For example, consider two information sources providing information whether "an observed object is a car" (denoted by $ic_1$ and $ic_2$, respectively). Assuming that the first source does not know and the second having inconsistent information whether this is the case, with $\leq_t$ one derives that $\underbrace{ic_1}_{\mathtt{u}} \vee \underbrace{ic_2}_{\mathtt{i}}$ is $\mathtt{t}$, what is hardly intuitive. On the other hand, when $\leq_l$ is used, the value of $ic_1 \vee ic_2$ becomes $\mathtt{i}$, what actually is expected in this case. Of course, $\leq_l$ can easily be replaced by $\leq_t$ or other ordering adequate for the application in question.[4]

[4] In such a case, rather than using min, max, one should use the greatest lower bound glb and least upper bound lub, as discussed in Section 5.5.

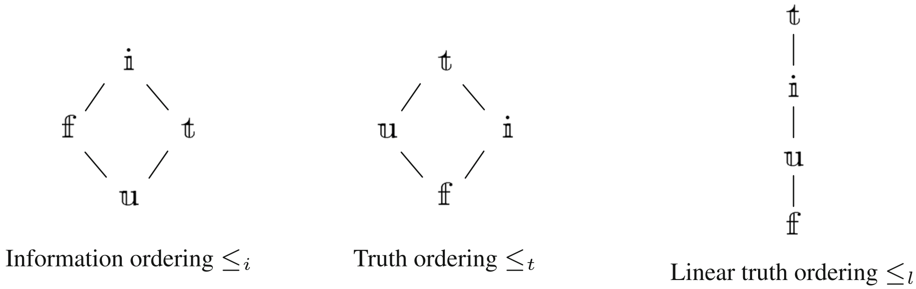


**Fig. 4** Sample four-valued orderings on truth values

The negation of truth values is defined by:

$$\neg\mathbb{t} \overset{\text{def}}{=} \mathbb{f}, \quad \neg\mathbb{f} \overset{\text{def}}{=} \mathbb{t}, \quad \neg\mathbb{i} \overset{\text{def}}{=} \mathbb{i}, \quad \neg\mathbb{u} \overset{\text{def}}{=} \mathbb{u}. \tag{6}$$

To define the semantics of plain formulas we consider an arbitrary signature $\Sigma = \langle \mathcal{R}, \mathcal{C} \rangle$. Semantically, plain formulas are evaluated in world aspects. By $\mathbb{W}$ we denote the set of world aspects. For $w \in \mathbb{W}$, the *truth value of a plain formula* $F \in \mathcal{F}^I$ *in* $w$, denoted by $w(F)$, is defined in Table 1.

## 5.2 From beliefs and goals to desires

Typically, goals in planning and agent systems are expressed by plain formulas. Given current beliefs and a goal, one can determine desires, representing desired world aspects. This, might be done in many different ways, for example exploiting the correspondence between a disjunctive normal form of the goal, the beliefs and the resulting world aspects, as shown in the following example.

**Example 2** (Example 1 continued) Assume that the current Eve's goal is:

$$\exists x\big(prepare(x)\big) \land pass(test). \tag{7}$$

Given that the domain of $x$ is $\{pres, essay\}$, (7) is equivalent to:

$$\big(prepare(pres) \lor prepare(essay)\big) \land pass(test),$$

whose disjunctive normal form is:

$$\big(prepare(pres) \land pass(test)\big) \lor \big(prepare(essay) \land pass(test)\big).$$

**Table 1** The semantics of plain formulas, where $w \in \mathbb{W}$, $\ell$ is a positive literal and $A, B \in \mathcal{F}^I$

| | |
|---|---|
| $w(\tau) \overset{\text{def}}{=} \tau,\ \text{for } \tau \in \mathcal{T};$ | $w(\neg A) \overset{\text{def}}{=} \neg w(A);$ |
| $w(\ell) \overset{\text{def}}{=} \begin{cases} \mathbb{t} & \text{when } \ell \in w \text{ and } \neg\ell \notin w; \\ \mathbb{f} & \text{when } \ell \notin w \text{ and } \neg\ell \in w; \\ \mathbb{i} & \text{when } \ell \in w \text{ and } \neg\ell \in w; \\ \mathbb{u} & \text{when } \ell \notin w \text{ and } \neg\ell \notin w; \end{cases}$ | $w(A \land B) \overset{\text{def}}{=} \min_{\leq_l}\{w(A), w(B)\};$ <br> $w(A \lor B) \overset{\text{def}}{=} \max_{\leq_l}\{w(A), w(B)\};$ <br> $w\big(\forall x(A(x)\big) \overset{\text{def}}{=} \min_{\leq_l}\{w(A(c)) \mid c \in \mathcal{C}\};$ <br> $w\big(\exists x(A(x)\big) \overset{\text{def}}{=} \max_{\leq_l}\{w(A(c)) \mid c \in \mathcal{C}\}.$ |

Since Eve believes that traveling does not leave her a sufficient amount of time for preparing her presentation ($prepare(pres) \implies \neg travel$) nor writing an essay ($prepare(essay) \implies \neg travel$), the resulting desire base can contain two desired world aspects:

$$\{prepare(pres), pass(test), \neg travel\}, \{prepare(essay), pass(test), \neg travel\}.$$

□

Desire determination is a broad topic. In the rest of the paper we abstract from particular methods by assuming that a goal-desires bridging function is provided as an inherent part of 4Bdi agent-specific frame.

### 5.3 From beliefs and desires to intentions

In general, intentions are chosen desires to which the agent has committed in a course of practical reasoning. In its first phase, the agent deliberates what intention to achieve, and then, during means-end reasoning, plans how to achieve it. The resulting plan is typically chosen from a pre-assembled collection of plans constituting a plan library that creates an independent plan base external to the 4Bdi framework. Otherwise, a plan leading from the current world state to a state satisfying the intention, can be constructed from scratch. The whole process of practical reasoning must be somehow addressed by an agent. The first phase of filtering intentions from desires is encapsulated in a rather complex desires-intentions bridging function given as a part of an agent-specific 4Bdi frame, as illustrated in the following example.

**Example 3** (Example 2 continued) In order to accomplish the desired world aspects, Eve has to examine her capabilities. In the scenario three actions are explicitly referred to: $prepare(pres)$, $prepare(essay)$, $pass(test)$. Given that:

- Eve does not have enough time to write an essay (a corresponding precondition of action $prepare(essay)$ is violated);
- she is capable of preparing her presentation if she does not spend time on traveling;
- she can attempt the test, what is a part of a plan to accomplish a desire $pass(test)$,

the resulting intention base may, among others, contain the world aspect:

$$\{prepare(pres), attempt(test), \neg travel\}. \tag{8}$$

□

**Remark 3** An interesting phenomenon implicitly occurring in Example 3 is typically neglected in planning scenarios: the plan extracted from (8) does not necessarily guarantee the initial goal (7), involving $pass(test)$. Indeed, taking a test does not imply passing it. However, in realistic scenarios, one occasionally accepts plans with unknown or inconsistent results when no plan ensuring success is available. These aspects are addressed in [6], sharing with 4Bdi the same logical background.[5] □

### 5.4 4Bdi frames and structures

Let us now define 4Bdi frames and structures as outlined in Sections 5.2– 5.3.

[5] For an open source implementation of the paraconsistent rule language 4QL with belief bases, belief shadowing and actions, see the `inter4QL` open-source interpreter available via `4ql.org`.

Let the signature $\Sigma = \langle \mathcal{P}, \mathcal{C} \rangle$ be fixed. We assume that the domain consists of constants of $\mathcal{C}$. Recall that a *world aspect* is a finite set of ground literals over $\Sigma$ and a *cluster* is a finite set of world aspects. By $\mathbb{C}$ we denote the set of clusters. In the next definition we use notation $\mathcal{D} \stackrel{\text{def}}{=} \mathcal{G}(\mathcal{B}, G)$ and $\mathcal{I} \stackrel{\text{def}}{=} \mathcal{A}(\mathcal{B}, \mathcal{D}) = \mathcal{A}(\mathcal{B}, \mathcal{G}(\mathcal{B}, G))$.

**Definition 6** (4BDI frames) A 4BDI *frame* is a pair $\mathbb{F} = \langle \mathcal{G}, \mathcal{A} \rangle$, where:

- $\mathcal{G} : \mathbb{C} \times \mathcal{F}^I \longrightarrow \mathbb{C}$ is a *goal-desires bridging function* (takes a belief base and a goal as arguments and returns a desire base);
- $\mathcal{A} : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$ is a *desires-intentions bridging function* (takes a belief base and a desire base as arguments and returns an intention base),

such that:

1. intentions are selected from desires: for all $G \in \mathcal{F}^I$ and $\mathcal{B} \in \mathbb{C}$, $\underbrace{\mathcal{A}(\mathcal{B}, \mathcal{G}(\mathcal{B}, G))}_{\mathcal{I}} \subseteq \underbrace{\mathcal{G}(\mathcal{B}, G)}_{\mathcal{D}}$;
2. for each positive literal $\ell$,
   - generalization of intention consistency at the world aspects level: for all $w \in \mathcal{I}$, $w(\ell) \in \mathcal{T}^d$;
   - generalization of intention consistency at the intention bases level: $\text{lub}_{\leq_i}\{w(\ell) \mid w \in \mathcal{I}\} \in \mathcal{T}^d$. □

Given particular goals, 4BDI frames are used to determine 4BDI structures.

**Definition 7** (4BDI structures) Let $G \in \mathcal{F}^I$ be a plain formula representing a goal,[6] and $\mathcal{B} \in \mathbb{C}$ be a belief base. Let $\mathbb{F} = \langle \mathcal{G}, \mathcal{A} \rangle$ be a 4BDI frame. A 4BDI *structure for* $G$ *over* $\mathbb{F}$ *and* $\mathcal{B}$ is a triple $\mathbb{F}(G) \stackrel{\text{def}}{=} \langle \mathcal{B}, \mathcal{D}, \mathcal{I} \rangle$, where:

1. $\mathcal{D} = \mathcal{G}(\mathcal{B}, G)$: desires are obtained from the goal assuming current beliefs;
2. $\mathcal{I} = \mathcal{A}(\mathcal{B}, \mathcal{D})$: intentions are obtained from desires assuming current beliefs. □

4BDI agents' mental attitudes are represented by 4BDI structures. That is, for each 4BDI agent $a \in Ag$, there is a goal $G_a$, an associated 4BDI frame $\mathbb{F}_a$ and a 4BDI structure $\mathbb{S}_a \stackrel{\text{def}}{=} \mathbb{F}_a(G_a) = \langle \mathcal{B}_a, \mathcal{D}_a, \mathcal{I}_a \rangle$.

## 5.5 Semantics of 4BDI formulas

4BDI formulas are evaluated in clusters provided by 4BDI structures. To define the semantics of 4BDI formulas we will need the greatest lower bound (glb) and the least upper bound (lub) wrt ordering $\leq_i$ shown in Fig. 4. We assume that:

$$\begin{aligned} &\text{glb}_{\leq_t}(\emptyset) \stackrel{\text{def}}{=} \mathbb{t};\ \text{lub}_{\leq_t}(\emptyset) \stackrel{\text{def}}{=} \mathbb{f}; \\ &\text{glb}_{\leq_i}(\emptyset) \stackrel{\text{def}}{=} \mathbb{i};\ \text{lub}_{\leq_i}(\emptyset) \stackrel{\text{def}}{=} \mathbb{u}. \end{aligned} \tag{9}$$

Let $\bar{\mathbb{S}} = \langle \mathbb{S}_a \mid a \in Ag \rangle$ be the tuple of 4BDI structures associated with agents in $Ag$. For every $\mathbb{O} \in \{\mathbb{B}, \mathbb{D}, \mathbb{I}\}$, $a \in Ag$ and $\tau \in \mathcal{T}$, we set $\mathbb{O}_a(\tau) \stackrel{\text{def}}{=} \tau$. For any 4BDI formula of the

[6] If there are a (finite) number of goals, we consider $G$ to be the conjunction of them.

form $\mathbb{O}_a(A)$, where $A \in \mathcal{F}^I$ is a plain formula, its *truth value in* $\bar{\mathbb{S}}$ is defined in Table 2. In order to calculate the *truth values of an arbitrary* 4BDI *formula* $A \in \mathcal{F}^{bdi}$ *in* $\bar{\mathbb{S}}$ of the form $\mathbb{O}_a(B)$, we use Algorithm 1 which successively computes truth values of subformulas of the form $\mathbb{O}_a(F)$, where $F \in \mathcal{F}^I$, and substitutes them with the calculated truth values.

**Algorithm 1** Calculating truth values of 4BDI formulas.

**Input**: • A 4BDI formula $A \in \mathcal{F}^{bdi}$ of the form $\mathbb{O}_a(B)$
• A tuple of 4BDI structures $\bar{\mathbb{S}} = \langle \mathbb{S}_a \mid a \in Ag \rangle$
**Output**: The truth value of $A$ in $\bar{\mathbb{S}}$

**set** $C = A$;
**repeat**
  **chose** a subformula of $C$ of the form $\mathbb{O}'_b(F)$, where $F \in \mathcal{F}^I$;
  **set** $\tau$ to be the truth value of $\mathbb{O}'_b(F)$ in $\bar{\mathbb{S}}$ by applying a suitable clause shown in Table 2;
  **replace** in $C$ all occurrences of $\mathbb{O}'_b(F)$ with $\tau$
**until** $C$ is reduced to a truth value;
**return** $C$;

The following proposition demonstrates that the 4BDI semantics preserves commonly assumed logical properties of beliefs, desires, and intentions. Note that, in many-valued logics, implication can be defined in a variety of ways. Rather than implication, we employ truth ordering $\leq_l$ which reflects the semantics of classical implication on the truth values $\mathbb{t}$, $\mathbb{f}$. Below $\pm$ denotes the empty string or the negation connective.

**Proposition 1** *For any formula* $A \in \mathcal{F}^I$, *ground literal* $\ell$, *any tuple of* 4BDI *structures,* $\bar{\mathbb{S}} = \langle \mathbb{S}_a \mid a \in Ag \rangle$ *and any* $a \in Ag$,

$$\bar{\mathbb{S}}\big(\pm\mathbb{B}_a(A)\big) \leq_l \bar{\mathbb{S}}\big(\mathbb{B}_a\big(\pm\mathbb{B}_a(A)\big)\big); \tag{10}$$

$$\bar{\mathbb{S}}\big(\pm\mathbb{D}_a(A)\big) \leq_l \bar{\mathbb{S}}\big(\mathbb{B}_a\big(\pm\mathbb{D}_a(A)\big)\big); \tag{11}$$

$$\bar{\mathbb{S}}\big(\pm\mathbb{I}_a(A)\big) \leq_l \bar{\mathbb{S}}\big(\mathbb{B}_a\big(\pm\mathbb{I}_a(A)\big)\big); \tag{12}$$

$$\bar{\mathbb{S}}\big(\neg\mathbb{O}_a(\mathbb{f})\big) = \mathbb{t},\ \textit{for}\ \mathbb{O} \in \{\mathbb{B}, \mathbb{D}, \mathbb{I}\}; \tag{13}$$

$$\bar{\mathbb{S}}\big(\mathbb{I}_a(\ell)\big) \in \mathcal{T}^d; \tag{14}$$

$$\bar{\mathbb{S}}\big(\mathbb{I}_a(A)\big) \leq_l \bar{\mathbb{S}}\big(\mathbb{D}_a(A)\big). \tag{15}$$

Note that (10)–(12) represent positive and negative introspection, (13)–(14) generalizes the consistency laws for $\mathbb{B}$, $\mathbb{D}$, $\mathbb{I}$, while (15) reflects that intentions are selected from desires. In particular, (14)–(15) reflect requirements stated in points 1– 2 of Definition 6.

## 6 Shadowing beliefs, desires and intentions

Belief revision/update/merging is an important task in agents' activities [38]. However, it typically requires deep and/or complex adjustments of belief bases even when adaptations

**Table 2** Semantics of 4BDI formulas, where $\bar{\mathbb{S}} = \langle \mathbb{S}_a \mid a \in Ag \rangle$ and $A \in \mathcal{F}^I$

$$\bar{\mathbb{S}}\big(\mathbb{B}_a(A)\big) \stackrel{\text{def}}{=} \text{lub}_{\leq_i}\{w(A) \mid w \in \mathcal{B}_a\};$$

$$\bar{\mathbb{S}}\big(\mathbb{D}_a(A)\big) \stackrel{\text{def}}{=} \text{lub}_{\leq_i}\{w(A) \mid w \in \mathcal{D}_a\};$$

$$\bar{\mathbb{S}}\big(\mathbb{I}_a(A)\big) \stackrel{\text{def}}{=} \text{lub}_{\leq_i}\{w(A) \mid w \in \mathcal{I}_a\}.$$

happen to be transient. In [5], a new kind of beliefs' update, belief shadowing, is introduced. It depends on a swap of beliefs when a part of one belief base is to be shadowed by a belief base of superior agents or roles, without changing both belief bases. This substantially improves the complexity of reasoning. Because beliefs influence agents' motivational attitudes, this approach is expanded for 4BDI agents.

As an agent may play dynamically assigned roles, let $Rl$ denote a (finite) set of *role names* ($Rl \cap Ag = \emptyset$). We assume that roles have their associated beliefs, desires and intentions, and they can also extend the abilities of agents. For example, a role "coordinator" may have a belief base that includes aspects of mission coordination, as well as desires and intentions that reflect a role-associated goal. When an agent takes on this role, it not only adds its own beliefs, but it also expands its capabilities of coordinating the activities of other agents. The beliefs and goal-desires/desires-intentions bridging functions of the role become dominant.

**Remark 4** In 4BDI agents' beliefs, desires and intentions are to be structured and distributed among agents and roles. The intended 4BDI methodology of their use is similar to the object-oriented one, where functionalities allocated to a class reflect its responsibilities. When an agent takes on a role, it acquires beliefs/desires/intentions the role is responsible for, perhaps even contradicting the agent's own mental attitudes. When the agent leaves the role, it no longer needs to remember mental attitudes specific to the role. Moreover, keeping them could, in certain cases, lead to violations of security, procedures, role hierarchies, etc. A manager role, for instance, may store beliefs specific to managing a division of a company, associated rights, procedures, etc. When a regular employee fills the position, perhaps temporarily, such "higher level" beliefs are now in effect. It could even be detrimental to maintain the manager specific beliefs after the person leaves the managerial position

When keeping selected mental attitudes is justified, related updates to the agent's bases are responsibility of a system designer. Such updates are application-dependent so are not included in the general 4BDI shadowing machinery. □

**Definition 8** (Shadowing expressions) By a *shadowing expression* we mean any expression of the form:

$$a \,\mathsf{as}\, r_1 \,\mathsf{as}\, r_2 \,\mathsf{as}\, \ldots \,\mathsf{as}\, r_{m-1} \,\mathsf{as}\, r_m, \tag{16}$$

where $a \in Ag$, $m \geq 1$ and $r_1, \ldots, r_m \in Rl$. For $\epsilon$ denoting the empty sequence, we define $a \,\mathsf{as}\, \epsilon \stackrel{\text{def}}{=} a$. The *set of shadowing expressions* is denoted by $\mathcal{E}$. □

The intuition behind shadowing expression $a \,\mathsf{as}\, r$ is that agent $a$ takes on the role $r$. That is, mental attitudes of $a$ become those of $r$, unless $r$ is ignorant about the status of a given attitude, in which case the attitude of $a$ is binding. In the general form (16), $a$ takes on the roles $r_1, \ldots, r_{m-1}, r_m$. Mental attitudes of $a$ become those of $r_m$, unless $r_m$ is ignorant about the status of a given attitude, in which case, inductively, the attitude of $a \,\mathsf{as}\, r_1 \,\mathsf{as}\, \ldots \,\mathsf{as}\, r_{m-1}$ is binding.

To make use of roles in the logical language, the $\mathbb{B}$, $\mathbb{D}$, $\mathbb{I}$ operators introduced in Definition 2 have to be extended by assuming that Line (5) of the BNF grammar is replaced by:

$$\mathbb{B}_e\big(\langle\mathcal{F}\rangle\big) \mid \mathbb{D}_e\big(\langle\mathcal{F}\rangle\big) \mid \mathbb{I}_e\big(\langle\mathcal{F}\rangle\big), \text{ for } e \in \mathcal{E}. \tag{17}$$

That is, rather than using agent names, we use shadowing expressions to indicate the context of a given operator.

**Definition 9** [Extended operators and formulas] For $e$ of the form (16) with $m \geq 1$, the operators $\mathbb{B}_e()$, $\mathbb{D}_e()$, $\mathbb{I}_e()$ are respectively called *extended belief, extended desire*, and *extended* *intention operators*. By an *extended formula* we mean any formula defined by (17). An *extended* 4BDI *formula* is a formula defined as in Definition 4 with $a \in Ag$ replaced by $e \in \mathcal{E}$. The set of extended 4BDI formulas is denoted by $\mathcal{F}^{ext}$. □

Notice that the language offers a rich expressiveness. For example, one can formulate properties not directly available in other logical formalisms, like:

$$\mathbb{B}_{jack\ \text{as}\ coord}\big(\mathbb{I}_{eve\ \text{as}\ vln\ \text{as}\ med}(assist)\big),$$

stating that *jack*, acting as a mission coordinator (*coord*), believes that *eve*, acting as a volunteer (*vln*) being a medicine student (*med*), intends to assist medical staff in basic medical care.

The semantics of extended belief operators $\mathbb{B}_e$, with $e \in \mathcal{E}$, is provided in Table 3. Notice that for $i \in Ag \cup Rl$, $\mathbb{B}_i()$ is defined as in Table 2 since in these cases, $\mathcal{B}_i$ is a belief base in the sense assumed in Section 5.

The following example illustrates belief shadowing.

**Example 4** (Example 3 continued) Assume a large-scale natural disaster occurred, resulting in the displacement of a large number of evacuees. As a medicine student, Eve joined a volunteer team that was formed at her university to support evacuees in everyday matters or assist medical staff in basic medical care. The belief base associated with the role volunteer (*vln*) consists of two world aspects:

$$\{support, \neg assist, travel\}, \{\neg support, assist, travel\}. \tag{18}$$

When Eve takes the role *vln*, her beliefs are expressed by $\mathbb{B}_{eve\ \text{as}\ vln}()$. For example, though Eve believes she should not travel, her belief as a volunteer, $\mathbb{B}_{eve\ \text{as}\ vln}(travel)$, obtains the value $\mathbb{t}$, since the volunteer belief base consisting of world aspects (18) shadows Eve's original beliefs. Using the contents of (18), the beliefs $\mathbb{B}_{eve\ \text{as}\ vln}(support)$ and $\mathbb{B}_{eve\ \text{as}\ vln}(assist)$ obtain the value $\mathbb{i}$.

Consider now a role of a medical assistant (*med*) with its belief base containing only:

$$\{\neg support, assist, travel\}. \tag{19}$$

According to the contents of (19):

- the value of $\mathbb{B}_{eve\ \text{as}\ med}(assist)$ and $\mathbb{B}_{eve\ \text{as}\ vln\ \text{as}\ med}(assist)$ become $\mathbb{t}$;
- the value of $\mathbb{B}_{eve\ \text{as}\ med}(support)$ and $\mathbb{B}_{eve\ \text{as}\ vln\ \text{as}\ med}(support)$ become $\mathbb{f}$. □

When shadowing takes place, beliefs of $\mathbb{B}_{a\ \text{as}\ r_1\ \text{as}\ \ldots\ \text{as}\ r_m}(F)$ may be evaluated in any of structures $\mathbb{S}_a, \mathbb{S}_{r_1}, \ldots, \mathbb{S}_{r_m}$. Indeed, when for $k < i \leq m$, the truth value of $\mathbb{B}_i(F)$ in $\mathbb{S}_i$ is $\mathbb{u}$, the value of $\mathbb{B}_k(F)$ in $\mathbb{S}_k$ is examined. Therefore goal-desires and desires-intentions bridging functions have to be prepared for belief bases no longer understood as sets of world aspects, but as more general servers whose role is to evaluate and return answers to a queries asked to a sequence of structures.

**Table 3** Semantics of belief shadowing, where $F \in \mathcal{F}^I$, $a \in Ag$, $r_1, \ldots, r_m \in Rl$ and $\bar{\mathbb{S}} = \langle \mathbb{S}_i \mid i \in Ag \cup Rl \rangle$

$$\bar{\mathbb{S}}(\mathbb{B}_{a\ \text{as}\ r_1\ \text{as}\ \ldots\ \text{as}\ r_m}(F)) \stackrel{\text{def}}{=} \begin{cases} \bar{\mathbb{S}}(\mathbb{B}_{r_m}(F)) & \text{when } \bar{\mathbb{S}}(\mathbb{B}_{r_m}(F)) \neq \mathbb{u}; \\ \bar{\mathbb{S}}(\mathbb{B}_{a\ \text{as}\ r_1\ \text{as}\ \ldots\ \text{as}\ r_{m-1}}(F)) & \text{otherwise.} \end{cases}$$

**Definition 10** (General belief bases) Let $\mathcal{F}^{bel} \subseteq \mathcal{F}^{ext}$ be a set of extended formulas where extended operators are of the form $\mathbb{B}_e()$, for $e \in \mathcal{E}$. By a *general belief base* we mean any mapping $\mathcal{F}^{bel} \longrightarrow \mathcal{T}$ assigning truth values to extended BDI formulas in $\mathcal{F}^{bel}$. The set of general belief bases is denoted by $\mathbb{B}^*$. □

Though Definition 10 is general, in the sequel we deal with belief bases adequate for shadowing expressions. They can be seen as the sequences $\langle \mathcal{B}_1, \ldots, \mathcal{B}_k \rangle$, reflecting shadowing expressions, where $k \geq 1$ and $\mathcal{B}_1, \ldots, \mathcal{B}_k$ are belief bases. Their semantics follows from Table 3.

The *extended* 4BDI *frames* and *structures* are defined as follows.

**Definition 11** (Extended 4BDI frames and structures) An *extended* 4BDI *frame* is a pair $\mathbb{F}^* = \langle \mathcal{G}^*, \mathcal{A}^* \rangle$, where:

- $\mathcal{G}^* : \mathbb{B}^* \times \mathcal{F}^I \longrightarrow \mathbb{C}$ is a *goal-desires bridging function* (takes an extended belief base and a goal as arguments and returns a desire base);
- $\mathcal{A}^* : \mathbb{B}^* \times \mathbb{C} \longrightarrow \mathbb{C}$ is a *desires-intentions bridging function* (takes an extended belief base and a desire base as arguments and returns an intention base),

such that the requirements formulated in points 1 and 2 of Definition 6 are satisfied for $\mathcal{D}$ and $\mathcal{A}$ substituted respectively by $\mathcal{G}^*$ and $\mathcal{A}^*$.

An *extended* 4BDI *structure for a goal* $G \in \mathcal{F}^I$ *over* $\mathbb{F}^*$ *and* $\mathcal{B}^*$ is a triple $\mathbb{F}^*(G) \stackrel{\text{def}}{=} \langle \mathcal{B}^*, \mathcal{D}^*, \mathcal{I}^* \rangle$, defined as in Definition 7, replacing $\mathcal{B}$ by $\mathcal{B}^*$. □

While belief bases are directly shadowed, desire and intention bases are shadowed indirectly, since:

1. $\mathcal{D}^*$ is constructed using the shadowed belief base $\mathcal{B}^*$;
2. $\mathcal{I}^*$ is constructed using the shadowed belief base $\mathcal{B}^*$ and the desire base $\mathcal{D}^*$.

Extended 4BDI frames and structures behave like 4BDI frames and structures in the sense of Definitions 6, 7. Proposition 1 holds for extended operators, too.

The following example illustrates shadowing of 4BDI structures.

**Example 5** (Example 4 continued) Desire bases for '$eve \,\text{as}\, vln$' and '$eve \,\text{as}\, med$' are computed using a goal and beliefs provided by the belief base querying interface, i.e., using beliefs of the form $\mathbb{B}_{eve \,\text{as}\, vln}()$ and $\mathbb{B}_{eve \,\text{as}\, med}()$, respectively. Intention bases are to be computed from beliefs and desires, so in addition to $\mathbb{B}_{eve \,\text{as} \ldots}()$, one uses desire operators of the form $\mathbb{D}_{eve \,\text{as}\, vln}()$ and $\mathbb{D}_{eve \,\text{as}\, med}()$. After computing intention bases, querying about intentions will use intention operators, like $\mathbb{I}_{eve \,\text{as}\, vln}()$ and $\mathbb{I}_{eve \,\text{as}\, med}()$. Of course, the operators $\mathbb{B}_{eve \,\text{as} \ldots}()$, $\mathbb{D}_{eve \,\text{as} \ldots}()$ and $\mathbb{I}_{eve \,\text{as} \ldots}()$ can occur as a part of a more complex query expressed by 4BDI formulas.

To illustrate Definition 11, consider a goal $support \vee help$. The value of *help* is unknown to Eve as well as to the roles *vln*, *med*. The desire and intention bases may be then computed using $\mathcal{G}^*$ and $\mathcal{A}^*$, e.g., to contain the world aspect $\{support, travel\}$ and perhaps other literals from which one could extract a plan to achieve $support = \mathfrak{t}$ and $travel = \mathfrak{t}$. Given that *help* and a plan to make it true is a part of a background knowledge available to $\mathcal{G}^*$ and $\mathcal{A}^*$, another world aspect in the intention base could also contain *help*, such as being $\{help, travel\}$. □

Observe that the agent's 4BDI loop, discussed in Section 2, is realized both by 4BDI and extended 4BDI frames and structures: a frame delivers goal-desires and desires-intentions bridging functions which are then used to determine a 4BDI structure relevant for a given goal and current beliefs of the agent.

We assume that agents' and roles' mental attitudes are represented by (extended) 4BDI structures. That is, for each 4BDI agent and role $i \in Ag \cup Rl$, there is a goal $G_i$, an associated (extended) 4BDI frame $\mathbb{F}_i$ and structure $\mathbb{S}_i \stackrel{\text{def}}{=} \mathbb{F}_i(G_i) = \left\langle \mathcal{B}_i^{(*)}, \mathcal{D}_i, \mathcal{I}_i \right\rangle$, where the superscript $(*)$ denotes a 4BDI or an extended 4BDI structure, as appropriate.

Finally, the semantics of $\mathbb{D}_e()$ and $\mathbb{I}_e()$, where $e \in \mathcal{E}$, is defined in Table 4, where $\mathcal{D}_{a \text{ as } r_1 \text{ as} \ldots \text{as } r_m}$ and $\mathcal{I}_{a \text{ as } r_1 \text{ as} \ldots \text{as } r_m}$ are a desire base and an intention base constructed as in Definition 11 with $\mathcal{B}^* = \langle \mathcal{B}_a, \mathcal{B}_{r_1}, \ldots \mathcal{B}_{r_m} \rangle$. To broaden the definitions provided in Tables 3 and 4 for all extended 4BDI formulas, successively eliminate 4BDI subformulas by substituting them with truth values they evaluate to (along the lines of Algorithm 1).

## 7 Complexity of the approach

In the rest of this section, by complexity, we mean the data complexity of evaluating queries on the involved structures. That is, we assume that the query is fixed, and the complexity is expressed in terms of the size of the involved databases.

World aspects are sets of literals. Data complexity of first-order queries on sets of literals (thus world aspects, too) is PTIME and LOGSPACE [1] wrt the cardinality of the sets. Let $|w|$ and $|c|$ be the cardinalities of world aspect $w$ and cluster $c$, respectively. By the size of (extended) 4BDI structure $\mathbb{S} = \langle \mathcal{B}^{(*)}, \mathcal{D}, \mathcal{I} \rangle$, denoted by $|\mathbb{S}|$, we mean $|\mathcal{B}^{(*)}| + |\mathcal{D}| + |\mathcal{I}|$. For a tuple of 4BDI structures $\bar{\mathbb{S}}$, we define $|\bar{\mathbb{S}}| \stackrel{\text{def}}{=} \sum_{\mathbb{S} \in \bar{\mathbb{S}}} |\mathbb{S}|$. Note that $\mathcal{G}$ and $\mathcal{A}$ of (extended) 4BDI frames are used to construct 4BDI structures, not being involved in further query evaluation.

Let $\mathbb{O} \in \{\mathbb{B}, \mathbb{D}, \mathbb{I}\}$, $i \in Ag \cup Rl$ and $F \in \mathcal{F}^I$ be a plain formula. Then evaluating a query $\mathbb{O}_i(F)$ on a cluster $c$ is $\mathcal{O}\big(|c| * p(\max_{w \in c} |w|)\big)$, where $p(n)$ is the complexity of evaluating $F$ on a set of literals of cardinality $n$. Using Algorithm 1, we obtain the following results.

**Theorem 2** *(Tractability of querying (extended)* 4BDI *structures) For every (extended)* 4BDI *formula* $A \in \mathcal{F}^{ext}$ *and a tuple of (extended)* 4BDI *structures* $\bar{\mathbb{S}}$*, evaluating the query expressed by* $A$ *on* $\bar{\mathbb{S}}$ *is in* PTIME *and* LOGSPACE *wrt* $|\bar{\mathbb{S}}|$. □

That is, in terms of data complexity, reasoning over (extended) 4BDI structures and their shadowing is tractable wrt the size of the structures involved.

Notice that tractability is achieved wrt the size of the involved 4BDI structures due to the shift from general entailment to the reasoning-by-querying paradigm. Of course, 4BDI structures can be specified using rule-based languages, like variants of DATALOG$^{\neg\neg}$ or 4QL with tractable model computation. Since SQL is a very well tested and efficient technology, and our queries can easily be translated to SQL, we look at the computed models as SQL databases. That way we inherit LOGSPACE complexity of first-order/SQL queries, what does not have to hold for rule languages, where computing models typically requires PSPACE.

**Table 4** Semantics of extended $\mathbb{D}$ and $\mathbb{I}$ operators, where $F \in \mathcal{F}^I$, $a \in Ag$, $r_1, \ldots, r_m \in Rl$ and $\bar{\mathbb{S}} = \langle \mathbb{S}_i \mid i \in Ag \cup Rl \rangle$

| |
|---|
| $\bar{\mathbb{S}}\big(\mathbb{D}_{a \text{ as } r_1 \text{ as} \ldots \text{as } r_m}(F)\big) \stackrel{\text{def}}{=} \text{lub}_{\leq_i} \{w(F) \mid w \in \mathcal{D}_{a \text{ as } r_1 \text{ as} \ldots \text{as } r_m}\}$; |
| $\bar{\mathbb{S}}\big(\mathbb{I}_{a \text{ as } r_1 \text{ as} \ldots \text{as } r_m}(F)\big) \stackrel{\text{def}}{=} \text{lub}_{\leq_i} \{w(F) \mid w \in \mathcal{I}_{a \text{ as } r_1 \text{ as} \ldots \text{as } r_m}\}$. |

This feature is important in applications. For example, given that a 4BDI structure is huge, even of an extremely unrealistic size $\sim 10^{80}$ (being approximately the number of atoms in the observable universe), $\log(10^{80}) = 80 * \log(10) < 270$, so the additional space to compute queries is negligible. Of course, having rule-based specification is desirable and convenient in many scenarios. However, when queries are much more frequent than updates, one can compute the models of the most used parts of rule-based programs after each substantial portion of updates, store models as SQL databases, and expose them to queries enjoying a more efficient evaluation engine.

Complexity of constructing (extended) 4BDI structures from (extended) 4BDI frames and goals depends on the complexity of frame functions $\mathcal{G}$ and $\mathcal{A}$. In general, these functions may be of exponential (or higher) complexity, reducible to acceptable levels in real-world scenarios. For example, instead of generating all desired/intended world aspects, one can focus on a sufficient number or the most important ones.

## 8 Summary of design choices

### 8.1 Querying vs reasoning

One of the primary goals of this research was to develop efficient reasoning machinery by identifying the BDI components that enjoy tractability of reasoning. The natural candidates for such components are belief, intention and desire bases. Of course, even in the most basic instance of classical propositional logic, satisfiability is NPTIME and validity checking is CO-NPTIME. On the other hand, when agents or robots reason about the environment, existing beliefs, desires, and intents after they are established and stored in databases, one does not expect them to draw sophisticated conclusions using proof systems or SAT solvers. Therefore, we decided to chose reasoning-by-querying and use database techniques with tractable model and query evaluation. For the sake of efficient query answering we emphasized the possibility of using SQL database management systems. However, our tractability result also remains true whenever one uses rule-based languages, e.g, belonging to the DATALOG$^{\neg}$ or 4QL families of tractable formalisms. Of course, as discussed in Section 7, LOGSPACE is retained when rule-based specifications are compiled into SQL or other representations of first-order models.

To summarize, the possibility to compile belief, desire, and intention bases to standard database management systems with tractable query evaluation, is a minimal prerequisite for tractability.

### 8.2 Paracompleteness and paraconsistency

In order to express the desired phenomena in real-world applications, we have used a four-valued formalisms with non-classical truth values representing lack of knowledge and inconsistency of the expressed properties.

In real world scenarios inconsistencies may occur due to many factors, including:

- uncertain information that can lead to inconsistencies in agents' beliefs and decisions: agents might interpret the available information differently, leading to conflicts or contradictions;
- communication errors: MAS systems strongly rely on the agents' communication. Errors, such as message loss, duplication, or corruption, can lead to inconsistent beliefs among agents;

- conflicting desires: agents in a multiagent system may have different objectives. That may cause conflicts resulting in inconsistent decision-making;
- timing issues: agents may make decisions based on conflicting information acquired in different time moments, resulting in logical inconsistencies;
- synchronization: lack of proper coordination mechanisms can lead to inconsistencies in actions;
- resource allocation: inconsistent allocation of limited resources among agents can lead to inconsistent outcomes.

The need for dealing with inconsistencies is also addressed in [20, 25, 27, 28, 30].

As indicated in [16], to define shadowing one needs at least three-valued logic of Kleene, $K_3$ [29] with the third value representing unknown. On the other hand, when a part of a belief (desire, intention) base is shadowed by superior beliefs (desires, intentions), inconsistencies are practically unavoidable. One could unify truth values $\mathtt{u}$ and $\mathtt{i}$ and use $K_3$. In fact, $\mathtt{i}$ behaves similarly to $\mathtt{u}$: inconsistency of a formula identifies that one information source claims validity while another one reveals falsity of the formula. That is, at a meta-level it is *unknown* which source is right. Therefore, three-valued logical connectives have the same semantics no matter whether the third truth value is $\mathtt{i}$ or $\mathtt{u}$ (see, e.g., [39]). In order to distinguish those cases, in addition to $\mathtt{u}$, in 4BDI we have employed $\mathtt{i}$ which is additionally useful:

- for modeling convenience;
- to simplify the 4BDI formalism;
- to simplify and strengthen the querying machinery (see, e.g., [17, 34]).

Another important phenomena addressed by paracompleteness and paraconsistency are related to actions and planning with incomplete and/or inconsistent information, where one may need to allow inconsistent beliefs, goals, desires and intentions (see also Remark 3).

All in all, paracomplete formalisms could, in principle, suffice to define shadowing. However, without paraconsistency, several challenges related to expressiveness, modeling convenience and reasoning would arise.

### 8.3 Shadowing vs Revision

As known from the literature, belief revision/update is typically a computationally demanding task [31]. Among others, the sources of complexity are related to inconsistency avoidance and deciding which parts of databases should be replaced/repaired in response to incoming contradictory data. On the other hand, agents often do not need to deeply revise their mental attitudes. Sometimes it suffices to temporarily accept superior attitudes and act accordingly. This happens when an agent joins a group or takes on a particular role. Of course, in such cases one can expect that inconsistencies will occur. However, using a paraconsistent formalism agents can utilize (partially) inconsistent databases without trivializing the reasoning and obtained conclusions. Tools to react on inconsistencies using nonmonotonic/heuristic rules can also be used [17, 33, 34].

In summary, in order to allow agents to live with inconsistencies and react appropriately, in 4BDI we have chosen a four-valued paraconsistent formalism.

### 8.4 Resolving incompleteness and inconsistencies

In the research we focused on SQL-like implementations of 4BDI databases. However, as 4BDI is close to rule-based technologies, we expect that such SQL implementations will routinely store models of beliefs, desires and intentions computed by rule-based engines. During the model computation incomplete/inconsistent information can be resolved using nonmonotonic rules. Adequate tools are available, e.g., in 4QL's module querying mechanism. They can be invoked according to a chosen strategy. As discussed in [19], as regards timing "[…] there are at least three strategies:

- killing inconsistency at the root: to solve them as soon as possible;
- living with inconsistency: to postpone disambiguation to the last possible moment (or even forever);
- intermediate: to solve inconsistency each time new relevant information appears."

In 4BDI the first strategy can be applied when building the model. The others demand model recomputation when necessary or when new information becomes available.

## 9 Related work

Since the pioneering work of Bratman [8], BDI-based models have been intensively studied both from application-oriented and formal point of view. There is a broad literature discussing applied aspects of BDI (see [7, 10] and numerous references there). On the formal side, many logics supporting BDI reasoning have also been developed and investigated [11, 15, 18, 25, 26, 32, 35–37, 41, 46]. They primarily employ multi-modal logics with relatively high reasoning complexity. While the majority of these approaches focus on plans, actions, and temporal aspects, we abstract from them and bring pre-existing beliefs, desires and intentions into focus. Their construction is encapsulated in goal-desires and desires-intentions bridging functions, constituting essential components of 4BDI frames. The shift from general logical reasoning to the database perspective [17, 42] enables us to significantly reduce the high complexity of general logical reasoning to tractable reasoning-by-querying.

Paraconsistent reasoning is a well-established area (see, e.g., [4, 12] and references there). Even though inconsistency in modern AI applications is omnipresent, paraconsistent BDI models have not received sufficient attention. Yet, this aspect is pointed out in [25], where paraconsistent approach to BDI, specifically to desires, is identified as a challenging research area. In [45] the authors discuss paraconsistent logic, where BDI serves as a motivation, however focusing on the classical language. In 4BDI we use paraconsistent databases and reasoning-by-querying at all levels of mental attitudes, where queries are expressed in a BDI specific language. To the best of our knowledge, 4BDI is the first paraconsistent BDI formalism.

As agent systems act in dynamic environments, belief update and revision are in the mainstream of the area. For representative approaches see [21, 38, 40, 43] and references there. However, belief revision/update is typically a computationally demanding task [31]. In contrast to other approaches, the paper [5] provides tractable logical formalism to deal with shallow and possibly transient belief change. Without losing tractability, 4BDI substantially extends this formalism to deal with shadowing desires and intentions, too.

## 10 Conclusions

The research reported in the paper introduces a novel tractable and implementation friendly model of beliefs, desires and intentions, 4BDI, which is developed from the database perspective.

In particular, we have defined 4BDI frames and structures designed for modeling paraconsistent and paracomplete reasoning about BDI agents facing imperfect information. In comparison to other approaches, the formalism is conceptually light. It offers means for BDI revision via shadowing. This provides a level of expressiveness not available in other approaches. Importantly, a shift from general reasoning to reasoning-by-querying, renders the approach tractable. To the best of our knowledge, the reported expressive power while retaining tractability has never been achieved before.

What is also important, 4BDI unifies both theoretical and practical reasoning in a single approach. This is accomplished by a uniform representation of beliefs, desires and intentions, combined with querying and shadowing as closely related logic-based mechanisms.

The 4BDI framework is implementation-oriented. Indeed, belief, desire and intention bases can almost directly be implemented with SQL or rule-based database management systems, or in rule-based languages. Though a single query may refer to many agents, 4BDI structures are meant to be allocated to individual agents and roles. Therefore, 4BDI suits better to distributed rather than centralized architectures.

Future work will include a composition of 4BDI group structures from structures of group members, aiming for a uniform treatment of individual agents and arbitrarily nested groups. This task requires a nontrivial construction of goal-desire and desire-intention bridging functions at the group level.

A separate research may concern intention refinement using 4BDI structures and a suitable action and change theory. The approaches discussed in [25] combined with actions over paraconsistent belief bases presented in [6] could be good starting points in this direction.

**Acknowledgements** This work has been supported by grant 2015/19/B/ST6/02589 of the National Science Centre Poland. We would also like to thank reviewers for constructive comments.

**Data Availibility Statement** Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

## Declarations

**Conflicts of interest** The authors declare that they have no conflict of interest.



## References

1. Abiteboul, S., Hull, R., Vianu, V.: Foundations of Databases. Addison-Wesley, San Francisco, CA (1996)

2. Arieli, O., Avron, A.: The value of the four values. Artif. Intell. **102**(1), 97–141 (1998)
3. Belnap, N.: A useful four-valued logic. In: Epstein, G., Dunn, J. (eds.) Modern Uses of Many Valued Logic, pp. 5–37. Reidel, Dordrecht (1977)
4. Béziau, J.Y., Carnielli, W., Gabbay, D. (eds.): Handbook of Paraconsistency. College Pub, London (2007)
5. Białek Ł, Dunin-Kęplicz B, Szałas A Belief shadowing. In: Mascardi, V., Ricci, A., Weyns, D. (eds) Engineering Multi-Agent Systems. EMAS 2018, LNCS, vol 11375. Springer, Cham, pp 158–180 (2019)
6. Białek, Ł, Dunin-Kęplicz, B., Szałas, A.: A paraconsistent approach to actions in informationally complex environments. Ann. Math. Artif. Intell. **86**(4), 231–255 (2019)
7. Bordini, R., El Fallah, Seghrouchni A., Hindriks, K., et al.: Agent programming in the cognitive era. Auton Agents Multi Agent Syst **34**(2), 37 (2020)
8. Bratman, M.: Intention, Plans, and Practical Reason. Harvard Univ. Press, Cambridge, MA (1987)
9. Bratman, M., Israel, D., Pollack, M.: Plans and resource-bounded practical reasoning. Comput. Intell. **4**, 349–355 (1988)
10. Cardoso, R., Ferrando, A.: A review of agent-based programming for multi-agent systems. Computers **10**(2), 16 (2021)
11. Cohen, P., Levesque, H.: Intention is choice with commitment. Artif Intelligence **42**(3), 213–261 (1990)
12. da Costa N, Béziau J, Bueno OO (2005) On the usefulness of paraconsistent logic. In: Vanderveken D (ed) Logic, Thought and Action, LEUS, vol 2. Springer, Dordrecht, chap 20, p 465 – 478
13. deAmo, S., Pais, M.: A paraconsistent logic programming approach for querying inconsistent databases. Int. J. Approx. Reason. **46**(2), 366–386 (2007)
14. d'Inverno, M., Luck, M.: Understanding Agent Systems. Springer, Berlin, Heidelberg (2001)
15. d'Inverno, M., Luck, M., Georgeff, M., et al.: The dMARS architecture: A specification of the distributed multi-agent reasoning system. Auton Agents Multi Agent Syst **9**(1–2), 5–53 (2004)
16. Dunin-Kęplicz, B., Szałas, A.: Shadowing in many-valued nested structures. In: Proc. 50th IEEE Int Symp. on Multiple-Valued Logics. IEEE, New Jersey, p 230–236 (2020)
17. Dunin-Kęplicz, B., Szałas, A.: Taming complex beliefs. Trans on Comp Collective Intelligence XI LNCS **8065**, 1–21 (2013)
18. Dunin-Kęplicz, B., Verbrugge, R.: Teamwork in Multi-Agent Systems. A Formal Approach. John Wiley & Sons Ltd, Chichester, UK (2010)
19. Dunin-Kęplicz, B., Szałas, A., Verbrugge, R.: Tractable reasoning about group beliefs. In: Dalpiaz, F., Dix, J., van Riemsdijk, M. (eds.) EMAS'2014, LNCS, vol. 8758, pp. 328–350. Springer, Cham (2014)
20. Fabiano, F., Burigana, A., Dovier, A., et al.: Multi-agent epistemic planning with inconsistent beliefs, trust and lies. In: Duc Nghia Pham, D., Theeramunkong, T., Governatori, G., et al. (eds) Proc. PRICAI 2021: Trends in AI 18th Pacific Rim Int. Conf Part I, LNCS, vol 13031. Springer, pp 586–597 (2021)
21. Fermé, E., Hansson, S.O.: AGM 25 years: Twenty-five years of research in belief change. J of Philosophical Logic **40**(2), 295–331 (2011)
22. Gebser, M., Kaminski, R., Kaufmann, B., et al.: Answer Set Solving in Practice. Synthesis Lectures on AI and ML, Morgan and Claypool Pub., San Rafael, CA (2012)
23. Gelfond, M., Kahl, Y.: Knowledge Representation, Reasoning, and the Design of Intelligent Agents - The Answer-Set Programming Approach. Cambridge University Press, Cambridge, UK (2014)
24. Georgeff, M., Lansky, A.: Reactive reasoning and planning. In: Forbus, K., Shrobe, H. (eds) Proc. 6th National Conf. on AI. Morgan Kaufmann, San Francisco, CA, pp 677–682 (1987)
25. Herzig, A., Lorini, E., Perrussel, L., et al.: BDI logics for BDI architectures: Old problems, new perspectives. Künstliche Intell **31**(1), 73–83 (2017)
26. van der Hoek, W., Wooldridge, M.: Logics for multiagent systems. AI Mag. **33**(3), 92–105 (2012)
27. Hunter, A., Parsons, S., Wooldridge, M.: Measuring inconsistency in multi-agent systems. Künstliche Intell **28**(3), 169–178 (2014)
28. Kalech M, Natan A (2022) Model-based diagnosis of multi-agent systems: A survey. In: Proc. 36th AAAI Conference on Artificial Intelligence, AAAI'2022. AAAI Press, pp 12334–12341
29. Kleene, S.: On notation for ordinal numbers. J Symbolic Logic **3**, 150–155 (1938)
30. Kollingbaum, M., Norman, T., Preece, A., et al.: Norm conflicts and inconsistencies in virtual organisations. In: Noriega, P., Vázquez-Salceda, J., Boella, G., et al. (eds) Coordination, Organizations, Institutions, and Norms in Agent Systems II - AAMAS 2006 and ECAI 2006 International Workshops, COIN Revised Selected Papers, LNCS, vol 4386. Springer, pp 245–258 (2006)
31. Liberatore, P.: The complexity of belief update. Artif. Intell. **119**(1), 141–190 (2000)
32. Lorini, E., Herzig, A.: A logic of intention and attempt. Synthese **163**(1), 45–77 (2008)
33. Małuszyński, J., Szałas, A.: Living with inconsistency and taming nonmonotonicity. In: de Moor, O., Gottlob, G., Furche, T., et al. (eds.) Datalog Reloaded, LNCS, vol. 6702, pp. 334–398. Springer, Cham (2011)

34. Małuszyński, J., Szałas, A.: Partiality and inconsistency in agents' belief bases. In: Barbucha, D., Le, M., Howlett, R., et al. (eds.) KES-AMSTA, Frontiers in AI and Applications, vol. 252, pp. 3–17. IOS Press, Amsterdam (2013)
35. Meyer, J., Broersen, J., Herzig, A.: BDI logics. In: van Ditmarsch, H. JHalpern, van der Hoek W, et al (eds) Handbook of Logics of Knowledge and Belief. College Pub., London, chap 10, p 453–498 (2015)
36. Meyer, J.J., de Boer, F., van Eijk, R., et al.: On programming KARO agents. Log. J. IGPL **9**(2), 245–256 (2001)
37. Novák, P., Dix, J.: Modular BDI architecture. In: Nakashima, H., Wellman, M., Weiss, G., et al. (eds.) 5th Int, pp. 1009–1015. Joint Conf. AAMAS, ACM, New York (2006)
38. Peppas, P.: Belief revision. In: van Harmelen, F., Lifschitz, V., Porter, B. (eds.) Handbook of KR, pp. 317–359. Elsevier, Amsterdam (2008)
39. Priest, G.: The logic of paradox. J. Philos. Log. **8**(1), 219–241 (1979)
40. Priest, G.: Paraconsistent belief revision. Theoria **67**(3), 214–228 (2001)
41. Rao, A., Georgeff, M.: Decision procedures for BDI logics. J Logic and Computation **8**(3), 293–344 (1990)
42. Shoham, Y.: Logical theories of intention and the database perspective. J. Philos. Log. **38**(6), 633–647 (2009)
43. Testa, R.R., Coniglio, M.E., Ribeiro, M.M.: AGM-like paraconsistent belief change. Logic Journal of the IGPL **25**(4), 632–672 (2017)
44. Urquhart, A.: Many-Valued Logic. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, vol. 3, pp. 71–116. Reidel, Dordrecht (1986)
45. Villadsen, J.: Paraconsistent assertions. In: Lindemann, G., Denzinger, J., Timm, I., et al. (eds) Proc. MATES: 2nd German Conf. Multiagent System Technologies, LNCS, vol 3187. Springer, Cham, pp 99–113 (2004)
46. Woolridge, M.: Reasoning about Rational Agents. John Wiley & Sons Inc, Chichester, UK (2003)


**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.