An Experience-Based BDI Logic: Motivating Shared Experiences and Intentionality

Nardine Osman, Marc D'Inverno, Carles Sierra, Leila Amgoud, Henri Prade, Matthew Yee-King, Roberto Confalonieri, Dave De Jonge, Katina Hazelden

To link to this article : DOI :10.1109/IECON.2013.6700233
URL : http://dx.doi.org/10.1109/IECON.2013.6700233

To cite this version : Osman, Nardine and D'Inverno, Marc and Sierra, Carles and Amgoud, Leila and Prade, Henri and Yee-King, Matthew and Confalonieri, Robertoand De Jonge, Dave and Hazelden, Katina *An Experience-Based BDI Logic: Motivating Shared Experiences and Intentionality*. (2013) In: Annual Conference of the IEEE Industrial Electronics Society - IECON 2013, 10 November 2013 - 13 November 2013 (Vienna, Austria).

An Experience-Based BDI Logic: Motivating Shared Experiences and Intentionality

Nardine Osman*, Mark d'Inverno†, Carles Sierra*, Leila Amgoud‡, Henri Prade‡,
Matthew Yee-King†, Roberto Confalonieri†, Dave de Jonge* and Katina Hazelden†
*Artificial Intelligence Research Institute (IIIA-CSIC), Barcelona, Spain
†Computing Department, Goldsmiths College, University of London, London, UK
‡Toulouse Institute of Computer Science Research (IRIT-CNRS), Toulouse, France

***Abstract*—This paper proposes the notion of experience to help situate agents in their environment, providing a link on how the continually evolving environment impacts the evolution of an agent's BDI model and vice versa. Then, using the notion of shared experience as a primitive construct, we develop a novel formal model of shared intention which we believe more adequately describes social behaviour than traditional BDI logics that focus on individual agents. Whilst many philosophers have argued that collective intentionality cannot always be equated to the collection of the individual agents' intentions, there has been no AI model that addresses this issue. We believe this is the first attempt to develop an explicit notion of shared experience from an AI perspective.**

I. Introduction

Existing models of BDI agents typically assume the agent has its belief base already built [1]–[3]. This paper extends such work by presenting an agent model that proposes the notion of experience, which supports situating agents in their environments and closes the BDI circle: an agent's experience in its environment impacts its beliefs, which in turn affect its desires, then intentions, leading to new actions, and back round again to new experiences. In other words, comparing to existing BDI approaches, experience is proposed as part of the missing link between the state of the environment and an agent's mental state.

After presenting the basic agent model, we then introduce a logical formalism (X-BDI) to describe the agent reasoning and whose semantics is grounded in the model. This formalism delivers the main novelly of our proposal: defining the notion of shared experience based upon the notion of individual experience, and using this definition to address the issue of collective intentionality. Following the intuitive assumption that shared experiences and shared intentions provide the basis of social behaviour,[1] which is the foundation of multiagent system research, we say we propose a model that provides the necessary tools for reasoning over such social behaviour.

The remainder of this paper is organised as follows: we start with a specification of the agent model, then we introduce the logic for the experience-based BDI model, and follow it with an example that illustrates the impact of this work. We then close with some literature review and concluding remarks.

[1]We argue that if social behaviour was not based on shared experiences and intentionality, then it would be the result of peers coincidentally acting in a similar way, which is not the type of social behaviour that interests our research work in multiagent systems.

II. The Agent Model

We argue that the physical world is populated with objects (some of which may be labelled as agents) and changes through events (where some subset may be labelled as actions performed by agents). Here, our focus is not on the physical world itself, but on how it is perceived by agents. It is the perception of the agent situated in a given environment that affects the agent's beliefs, desires, and intentions. More precisely, we introduce the notion of experience, which is the result of the agent perceiving its environment and forming some conceptions about it. We argue that experiences shape beliefs,[2] beliefs influence desires, and desires drive intentions. Intentions in turn lead the agent to act in its environment (where the action could be as simple as to observe), which results in the agent having new experiences. In other words, we say that experience is the missing link that completes the BDI cycle as it helps illustrate how the environment impacts the evolution of an agent's BDI and vice versa. This section provides the basic definitions of experiences, beliefs, desires, and intentions. We note that our proposal is based on the assumption that an agent can perceive its environment, resulting in having a set of percepts. Although we note that percepts may sometimes be imagined (following the cognitive neuroscience view that considers imagination as a second order perception [5], [6]). As such, the set of agent α's percepts is referred to as $\mathcal{P}_\alpha = \mathcal{R}_\alpha \cup \mathcal{I}_\alpha$, where percepts are either real ($\mathcal{R}$) or imagined ($\mathcal{I}$). We also note that percepts, both real and imagined, could describe various types of elements, amongst these are percepts of agents ($\mathcal{AG}$) and percepts of actions ($\mathcal{AC}$).

A. Experiences

An experience is simply viewed here as a number of percepts that the agent decides to group together. Experiences may be basic (composed of percepts) or compound (composed of a combination of percepts and other experiences). An example of the former would be the sunset one witnessed. An example of the latter would be going out, which contains nested experiences such as having a drink, seeing a show, etc.

[2]The rich interplay between experiences and beliefs and how experiences impact the formation of beliefs has been discussed in [4]. [4] argues why the classical knowledge representation approaches that equate the set of beliefs to the set of experiences is not expressive enough as it cannot describe what is "not possible" or what is "possible for sure", and it illustrates how experiences may be used to modify beliefs, such as reassessing one's generic beliefs when an experience contradicting this generic belief is observed.

We enrich the definition of an experience (although these may be considered complementary) by adopting Kant's argument that the basic cognitive hard wired relationships are time ordering ($\prec$), causality ($\rightsquigarrow$), and spatial ($\bowtie$) relationships. We also adopt Kant's view that human beings follow the categorical imperative method that makes them classify things as good or bad. As such, we introduce the evaluation function ($\heartsuit$) for evaluating experiences.

We state that the set of all experiences of agent α is $\mathcal{E}_\alpha$ (where the set of all experiences is referred to as $\mathcal{E}$). A single experience of agent α, $E_\alpha \in \mathcal{E}_\alpha$, is defined as the tuple:

$$E_\alpha = \langle \mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha, \prec, \rightsquigarrow, \bowtie, \heartsuit \rangle$$

- $\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha$ describes the experience's content (referenced as $content(E_\alpha)$), which may be composed of percepts ($\mathcal{P}'_\alpha \subseteq \mathcal{P}_\alpha$), sub experiences ($\mathcal{E}'_\alpha \subseteq \mathcal{E}_\alpha$), or a combination of both.
- $\prec \subseteq (\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha) \times (\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha)$ defines a temporal partial order over the elements of $content(E_\alpha)$.
- $\rightsquigarrow \subseteq (\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha) \times (\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha)$ defines a causal relationship over the elements of $content(E_\alpha)$, and it should satisfy the property $\rightsquigarrow \subseteq \prec$.
- $\bowtie \subseteq (\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha) \times (\mathcal{P}'_\alpha \cup \mathcal{E}'_\alpha) \times \mathcal{S}$ defines a spacial relationship between the elements of $content(E_\alpha)$, and it should satisfy the property $(a, b, s) \in \bowtie \Rightarrow (a, b) \notin \prec \wedge (b, a) \notin \prec$. An example of the set of spacial relationships could be $\mathcal{S} = \{above, below, left, right, behind, infront\}$.
- $\heartsuit : \mathcal{L} \rightarrow \mathcal{V}$ defines the experience's evaluation, which maps a label to an evaluation space. An example of an evaluation space is $\mathcal{V} = \{positive, negative\}$. The label is intended to represent the evaluation criteria, such as $\mathcal{L} = \{impact, usefulness, contentment, surprise\}$. The evaluation criteria may either represent rational criteria, such as whether "the experience helps one achieve their goals", or emotional ones.

a) Basic Experience: We distinguish a special type of experience, a basic experience $\overline{E}_\alpha$, in which all the contents of the experience are single percepts: $\overline{E}_\alpha = \langle \mathcal{P}'_\alpha, \prec, \rightsquigarrow, \bowtie, \heartsuit \rangle$. Agent α's set of all basic experiences is referred to as $\overline{\mathcal{E}}_\alpha$.

b) Basic Action: An interesting type of basic experience is the basic action, which we define as a basic experience, whose content contains only the percepts describing the pre-conditions of the action, the action percept describing the action itself ($AC \in \mathcal{AC}$), and the percepts describing the post-conditions of the action. The temporal relation then states that all pre-condition percepts should precede the action percept and that the action percept should precede all the post-condition percepts. Additionally, the causal relation states that the action percept leads to the post-condition percepts. A basic action is then defined as a basic experience $\overline{E}_\alpha$ that satisfies the following properties:

1– $\forall x \in content(\overline{E}_\alpha) \cdot (\exists y \in content(\overline{E}_\alpha) \cdot x \prec y \vee y \prec x)$

2– $\forall x, y, z \in content(\overline{E}_\alpha) \cdot x \prec y \wedge y \prec z \Rightarrow y \in \mathcal{AC}_\alpha$

3– $\mid \mathcal{AC}_\alpha \cap content(\overline{E}_\alpha) \mid = 1$

4– $\forall x, y \in content(\overline{E}_\alpha) \cdot x \rightsquigarrow y \Rightarrow x \in \mathcal{AC}_\alpha \wedge y \notin \mathcal{AC}_\alpha$

B. Beliefs

SImilar to existing BDI models, we say an agent holds a set of beliefs $\mathcal{B}$ about itself, its environment, its feelings (if any), and so on.

C. Desires

We say desires $\mathcal{D}$ are experiences that one desires to be realised in the environment. In other words, the percepts composing these experiences are imagined ones, as they have not yet happened or may never happen. For example, one can desire to win the Nobel prize. A desire is then a basic experience $\overline{E}$ such that $content(\overline{E}) \subset \mathcal{I}$, where $\mathcal{I}$ describes the set of imagined percepts. We note that desires may or may not be feasible. Furthermore, agents may or may not commit to realising their desires; they usually attempt to realise a subset of their desires. Desires that the agent commits to realising through a concrete plan are called intentions, which we introduce next.

D. Intentions

We say intentions are desires with plans: $\mathcal{T} = \mathcal{D} \times P$, where a plan is defined as a partial order of imagined basic experiences: $P \subseteq \overline{\mathcal{E}} \times \overline{\mathcal{E}}$. We note that basic experiences include basic actions, and an imagined basic action is an action that the agent has committed to realise but has not realised yet.

E. Agent Model

An agent model is then defined as the tuple:

$$\langle \mathcal{P}, \mathcal{E}, \mathcal{B}, \mathcal{D}, \mathcal{T} \rangle$$

where $\mathcal{P}$ is the agent's set of all percepts, $\mathcal{E}$ its set of experiences, $\mathcal{B}$ its set of beliefs, $\mathcal{D}$ its set of desires, and $\mathcal{T}$ its set of intentions.

III. Experience-Based BDI Logic

We now define a BDI logic that is grounded on the notion of experience which we will call X-BDI. We first define its syntax, then we give its semantics based on the formal definitions of the previous section, and finally present a sample of the logic's inference rules.

A. Syntax

- If φ is a propositional well formed formula then $\varphi \in$ X-BDI
- If φ is a propositional well formed formula, $\alpha \in A$ is an agent, and P is the set of plans, then $Action(\alpha, \varphi) \in P$
- If $\varphi \in$ X-BDI, $\alpha \in A$ is an agent, and $p \in 2^P$ is a plan then $E(\alpha, \varphi)$, $B(\alpha, \varphi)$, $D(\alpha, \varphi)$, $I(\alpha, \varphi, p) \in$ X-BDI
- If $\varphi \in$ X-BDI, $G = \{\alpha, \beta, \dots\} \subseteq A$, and $p \in 2^P$ then $CE(G, \varphi), JE(G, \varphi), SE(G, \varphi, p)$, $CD(G, \varphi), JD(G, \varphi), SD(G, \varphi, p)$, $CI(G, \varphi, p), JI(G, \varphi, p), SI(G, \varphi, p) \in$ X-BDI
- If $\varphi, \psi \in$ X-BDI then $\neg\varphi, (\varphi \vee \psi) \in$ X-BDI

The meaning of the symbols are as follows: G is a set of agents describing a group (which we also refer to as an agent community), *Action* has the obvious meaning, *B*, *D*, and *I* are the classical BDI symbols, and *E* stands for Experience. We use the letter *C* as a prefix to the *E*, *B*, *D*, and *I* symbols to mean Common (i.e. what is common to all agents in a group); *J* to mean Joint (i.e. what is common to all agents in a group and all those agents are aware that it is common to all agents in the group); and *S* to mean Shared (i.e. what is common to all agents in a group, and all those agents are aware that it is common to all agents in the group, and the shared belief literal is related to the agents willingly acting together a shared plan). Thus, the expression:

$$SI(\{\alpha,\beta\}, seeBCNplay, \{Action(\alpha, tv), Action(\beta, stadium)\})$$

is a literal in X-BDI expressing (according to the following semantics) that α and β have the shared intention to see Barcelona play a football match by α planning to see the match on tv and β going to the stadium.

B. Semantics

Formally, we understand the model of the world (which is populated by agents) as a Kripke structure where the state of the world and the mind state of agents evolve due to the actions of agents. For instance, if φ is perceived by α (and it becomes part of α's experiences), then α may decide to believe φ and we can say that the formula $B(\alpha,\varphi)$ is generated in α's belief base. Similarly, we can define the semantics of formulae like $D(\alpha,\varphi)$. We say agents may have beliefs about other agents, more concretely about the intentions, desires, beliefs, and experiences of other agents. This means that we give semantics to nested expressions. For instance, α may have seen an agent β in the environment watching a football match and jumping with joy when Barcelona scores a goal and thus α deduces from that experience that β desires that Barcelona wins, and that can be represented in the logic as $B(\alpha, D(\beta, winBCN))$.

Thus, given a set of agents $\{\alpha,\beta,\dots\}$, we define an X-BDI semantic model as a pair $\omega = \langle W, A\rangle$ where *W* is a classical logic model (i.e. interpretations for propositions) and *A* is a vector of agent models $A = \langle\alpha,\beta,\dots\rangle$:

a) Propositions: The interpretation of classical formulae is straightforward. We will use the symbol $\models_{PL}$ to refer to classical satisfaction.

$$\langle W,A\rangle \models \varphi \text{ iff } \varphi \in PL \text{ and } W \models_{PL} \varphi$$

b) Actions: We say it is sufficient for one agent in the community to have a percept of the action (even if the percept is not shared by others) and that the world is consistent with the changes in the environment caused by the action, for the action to be considered true, or satisfied (recall that $\mathcal{AC}_\alpha$ is α's set of perceived actions). The satisfaction of a plan consists on the satisfaction of each individual action in the plan.

$$\begin{aligned}
\langle W,A\rangle &\models Action(\alpha,\varphi) \text{ iff } \exists\beta\in A\cdot(\alpha,\varphi)\in\mathcal{AC}_\beta \\
&\qquad \text{and } \langle W,A\rangle \models \varphi \\
\langle W,A\rangle &\models p \text{ iff } p\in 2^P \text{ and } \forall a\in p\cdot\langle W,A\rangle \models a
\end{aligned}$$

c) Experiences: A model satisfies that α had an experience with content φ if it exists in α's set of experiences $\mathcal{E}_\alpha$. A group of agents had a common experience tagged as λ if all had an experience that semantically entails λ. Similarly, a group of agents had a joint experience tagged as λ if all had an experience that semantically entails λ and they recognise each other as part of the experience. Finally, a group of agents had a shared experience tagged as λ if it was a joint experience, they all actively participated in it, and the plan of action is said to be satisfied (in other words, each action has been perceived by at least one agent).[3]

$$\begin{aligned}
\langle W,A\rangle &\models E(\alpha,\varphi) \text{ iff } \varphi\in\mathcal{E}_\alpha \\
\langle W,A\rangle &\models CE(G,\lambda) \text{ iff } \\
&\qquad \forall\alpha\in G\cdot\exists\varphi\in \text{ X-BDI} \\
&\qquad \langle W,A\rangle \models E(\alpha,\varphi)\wedge\varphi\to\lambda \\
\langle W,A\rangle &\models JE(G,\lambda) \text{ iff } \\
&\qquad \forall\alpha,\beta\in G\cdot\exists\varphi\in \text{ X-BDI }\cdot \\
&\qquad \langle W,A\rangle \models E(\alpha,\varphi)\wedge\varphi\to\lambda\wedge B(\beta,E(\alpha,\varphi)) \\
\langle W,A\rangle &\models SE(G,\lambda,p) \text{ iff } \\
&\qquad \langle W,A\rangle \models JE(G,\lambda) \text{ and } \langle W,A\rangle \models p \text{ and } \\
&\qquad \forall\alpha\in G\cdot\exists Action(\alpha,\varphi)\in p\cdot\langle W,A\rangle \models \varphi\to\lambda
\end{aligned}$$

d) Beliefs: The semantics for beliefs is similar to experiences. However, the notion of shared belief is defined to exist when the agents in the group *recognise* one another in a shared experience that led all of them to hold the belief.

$$\begin{aligned}
\langle W,A\rangle &\models B(\alpha,\varphi) \text{ iff } \varphi\in\mathcal{B}_\alpha \\
\langle W,A\rangle &\models CB(G,\varphi) \text{ iff } \\
&\qquad \forall\alpha\in G\cdot\langle W,A\rangle \models B(\alpha,\varphi) \\
\langle W,A\rangle &\models JB(G,\varphi) \text{ iff } \\
&\qquad \forall\alpha,\beta\in G\cdot\langle W,A\rangle \models B(\alpha,\varphi)\wedge B(\beta,B(\alpha,\varphi)) \\
\langle W,A\rangle &\models SB(G,\varphi,p) \text{ iff } \\
&\qquad \langle W,A\rangle \models JB(G,\varphi) \text{ and } \\
&\qquad \exists\lambda\in \text{ X-BDI }\cdot\langle W,A\rangle \models SE(G,\lambda,p)\wedge\lambda\to\varphi
\end{aligned}$$

e) Desires: We use the same approach used for beliefs.

$$\begin{aligned}
\langle W,A\rangle &\models D(\alpha,\varphi) \text{ iff } \varphi\in\mathcal{D}_\alpha \\
\langle W,A\rangle &\models CD(G,\varphi) \text{ iff } \\
&\qquad \forall\alpha\in G\cdot\langle W,A\rangle \models D(\alpha,\varphi) \\
\langle W,A\rangle &\models JD(G,\varphi) \text{ iff } \\
&\qquad \forall\alpha,\beta\in G\cdot\langle W,A\rangle \models D(\alpha,\varphi)\wedge B(\beta,D(\alpha,\varphi)) \\
\langle W,A\rangle &\models SD(G,\varphi,p) \text{ iff } \\
&\qquad \langle W,A\rangle \models JD(G,\varphi) \text{ and } \\
&\qquad \exists\lambda\in \text{ X-BDI }\cdot\langle W,A\rangle \models SE(G,\lambda,p)\wedge\lambda\to\varphi
\end{aligned}$$

f) Intentions: The intention of an agent to reach a goal by following a plan is true if it is internalised in the intentions of the agents. Common intentions and joint intentions follow the same patterns as in the case of beliefs. However, a shared intention requires each agent to intend part of the plan to achieve the goal. That is, there is no shared intention if any agent is not involved in fulfiling (at least) part of the plan for achieving the goal.

$$\begin{aligned}
\langle W,A\rangle &\models I(\alpha,\varphi,p) \text{ iff } (\varphi,p)\in\mathcal{I}_\alpha \\
\langle W,A\rangle &\models CI(G,\varphi,p) \text{ iff } \\
&\qquad \forall\alpha\in G\cdot\langle W,A\rangle \models I(\alpha,\varphi,p) \\
\langle W,A\rangle &\models JI(G,\varphi,p) \text{ iff } \\
&\qquad \forall\alpha,\beta\in G\cdot\langle W,A\rangle \models I(\alpha,\varphi,p)\wedge B(\beta,I(\alpha,\varphi,p)) \\
\langle W,A\rangle &\models SI(G,\varphi,p) \text{ iff } \\
&\qquad \forall\alpha,\beta\in G\cdot\exists p_\alpha\subset p\cdot p_\alpha\neq\varnothing\wedge \\
&\qquad \langle W,A\rangle \models I(\alpha,\varphi,p_\alpha)\wedge B(\beta,I(\alpha,\varphi,p_\alpha))
\end{aligned}$$

[3]For simplification, we abuse notation and write $a \in p$ to mean that action *a* is 'part' of plan p and $p' \subseteq p$ to mean that plan p' is 'part' of plan *p*.

C. Inference

We assume classical propositional deduction for propositional fragments of X-BDI and modus ponens on the X-BDI formulae. X-BDI axioms are those of propositional logic plus specific axioms for X-BDI expressions. Due to space limitations, here we include axioms for Desires only:

$$\vdash \quad D(\alpha,\varphi) \wedge D(\beta,\varphi) \leftrightarrow CD(\{\alpha,\beta\},\varphi) \tag{1}$$
$$\vdash \quad CD(G,\varphi) \wedge D(\beta,\varphi) \leftrightarrow CD(G \cup \{\beta\},\varphi) \tag{2}$$
$$\vdash \quad JD(G,\varphi) \rightarrow CD(G,\varphi) \tag{3}$$
$$\vdash \quad CD(\{\alpha,\beta\},\varphi) \wedge B(\alpha, D(\beta,\varphi)) \wedge B(\beta, D(\alpha,\varphi)) \leftrightarrow JD(\{\alpha,\beta\},\varphi) \tag{4}$$
$$\vdash \quad JD(G \cup G',\varphi) \rightarrow JD(G,\varphi) \wedge JD(G',\varphi) \tag{5}$$
$$\vdash \quad SD(G,\varphi,p) \rightarrow JD(G,\varphi,p) \tag{6}$$
$$\vdash \quad JD(G,\varphi,p) \wedge SE(G,\lambda,p) \wedge (\lambda \rightarrow \varphi) \rightarrow SD(G,\varphi,p) \tag{7}$$

IV. Example

With the basic definitions we can now specify the exact steps needed for intentional agents to achieve shared intentions (and hence, shared experiences).

Motivating the Interest in a Shared Experience.: Agents may have different motives for engaging in shared experiences, for various reasons based on their own personal beliefs and desires. In BDI models, agents' actions are the result of their intentions, which are a subset of desires that the agent has chosen to realise. As such, the action of approaching other agents could in many cases be the result of a desire of sharing an experience with that agent.

Investigating & Realising a Joint Desire: When one agent desires a shared experience with other agents, then it will contact those agents to convince them to share its desire, and eventually fulfil that desire. In multiagent systems, this could be achieved through argumentation [7], or collaborative planning [8]. This step describes the agent's investigation of whether a joint desire for achieving the shared experience may be realised. If an agent believes that all other agents hold the same desire and that they are aware of each one of them holding that desire, then the agent will believe that there is a joint desire to achieve the shared experience.

Investigating & Realising a Shared Intention: After a joint desire is born, agents are then committed to finding the plan of action for fulfilling the desired shared experience. Argumentation (or argumentation-based negotiation) is usually used here to make sure that all parties agree to who does what and under what conditions. If an agent believes that all other agents intend to execute their role in a given plan and that they are all aware of each others' intentions, then the agent will believe that there is a shared intention for executing that given plan.

Realising a Shared Experience: Agents are now committed to carrying out the actions of the plan that they are responsible for. When the plan has been executed, the shared experience is realised.

In what follows, we take the scenario of two agents that are interested in the shared experience of buying a gift together. Let us say the agents adopt the names of their human owners, c for Carla and b for Bill, and they are interested in buying a gift for Mary together (the desire to buy a gift for Mary is referred to as g). In what follows, we describe what actions the agents could perform based on their X-BDI, and the changes resulting in their X-BDI for each action perceived occurring in their environment.

1) *Motivating the Interest in a Shared Experience.* Say Carla is interested in buying Mary a gift, but because she cannot afford a gift on her own, she desires to buy the gift with one of Mary's friends (referred to as *MaryFriend*). As such, her belief base includes:

$$D(c, SE(\{c, MaryFriend\}, g, inanyway)) \tag{i}$$

Note that *inanyway* refers to some plan of action that has not been decided upon yet.
Bill desires to share any experience with Carla (referred to as *anything*), possibly because he is dependent on her, or maybe he is secretly in love with her. As such, Bill's belief base includes:

$$D(b, SE(\{c,b\}, anything, inanyway)) \tag{ii}$$

2) *Investigating & Realising a Joint Desire.* In order for Carla to fulfil her desire of finding a partner for buying Mary a gift with (literal (i) in Carla's belief base), she contacts Mary's friends, asking each "Would you like to buy a gift for Mary with me?" Bill replies with a "Yes" to fulfil his desire of sharing any experience with Carla (literal (ii) in Bill's belief base). This exchange of information leads to Carla adding the following to her belief base:

$$B(c, D(b, SE(\{c,b\}, g, inanyway))) \tag{iii}$$
$$B(c, B(b, D(c, SE(\{c,b\}, g, inanyway))))$$

And deducing that:

$$B(c, JD(\{c,b\}, SE(\{c,b\}, g, inanyway)))$$

Similarly, Bill modifies his belief base by adding:

$$B(b, D(c, SE(\{c,b\}, g, inanyway))) \tag{iv}$$
$$B(b, B(c, D(b, SE(\{c,b\}, g, inanyway))))$$

And deducing that:

$$B(b, JD(\{c,b\}, SE(\{c,b\}, g, inanyway)))$$

As such, the following now holds:

$$JD(\{c,b\}, SE(\{c,b\}, g, inanyway)) \tag{v}$$

Note that the joint desire for a shared experience may be the result of differing individual desires. In our example, Carla just needs someone to share her expenses, whereas Bill just wants to do anything jointly with Carla.
Also note that (v) is deduced from axiom (4), the belief literals (iii) and (iv), and formula $CD(\{c,b\}, SE(\{c,b\}, g, inanyway))$, which in turn is deduced from axiom (1) and belief literals (i) and (ii).
All the deduced formulae that we mention in this section are deduced in a similar straightforward manner.

3) *Investigating & Realising a Shared Intention.* After both agents agree that there is a joint desire for buying the gift together, they go on to argue on the details of the plan they will follow. For example, where to buy the gift? What to buy? In this paper, we do not dwell on the details of how a plan is agreed upon. Assuming both agents agree on a plan p, and they both communicate their agreement to each other, then Carla's belief base is updated to contain the following:

$$I(c, SE(\{c,b\}, g, p), p_c)$$
$$B(c, I(b, SE(\{c,b\}, g, p), p_b))$$
$$B(c, p_c \subset p \wedge p_c \neq \varnothing \wedge p_b \subset p \wedge p_b \neq \varnothing)$$

And Carla deduces:

$$B(c, SI(\{c,b\}, SE(\{c,b\}, g, p), p))$$

Note that p_c is Carla's part of the plan and p_b is Bill's part of the plan.
Similarly, Bill's belief base is updated by adding:

$$I(b, SE(\{c,b\}, g, p), p_b)$$
$$B(b, I(c, SE(\{c,b\}, g, p), p_c))$$
$$B(b, p_c \subset p \wedge p_c \neq \varnothing \wedge p_b \subset p \wedge p_b \neq \varnothing)$$

And Bill deduces:

$$B(b, SI(\{c,b\}, SE(\{c,b\}, g, p), p))$$

As such, the following now holds:

$$SI(\{c,b\}, SE(\{c,b\}, g, p), p)$$

Note that each agent commits to performing its own part of the plan (in this case, p_c and p_b) to make sure the plan is realised, and both agents believe that this plan will realise their shared experience.

4) *Realising a Shared Experience.* After executing their actions and the plan is fulfilled, if the agents can perceive that the plan has been fulfilled, then Carla's belief base is updated to contain the following:

$$B(c, SE(\{c,b\}, g, p))$$

Similarly, Bill's belief base is updated by adding:

$$B(b, SE(\{c,b\}, g, p))$$

And the shared experience is said to have been realised:

$$SE(\{c,b\}, g, p)$$

V. Background

The literature on BDI is vast. However, since the original proposal by Bratman [9], most work has concentrated on providing agent-oriented programming languages based on BDI concepts such as Jason [2], Jack [3], AgentSpeak [1], [10], 3APL [11] or 2APL [12]; and on logical approaches to the BDI model such as modal logic [13], first-order logic [10], or degrees [14], [15]. Differently from existing work, we extend the BDI model with the notion of experience, and use that to introduce the notions of shared experiences and collective intentionality.

Another related line of work with respect to the notion of experience is Case-Based Reasoning (CBR) [16]. In CBR certain types of (problem-solving) experiences are kept in memory and recalled in order to be reused in the solution of new problems. More importantly, the notion of a shared experience is also not central in CBR; although federated CBR [17] discusses a remotely related issue, which is the communicability of knowledge and the coordination amongst agents for acquiring new problem solving capabilities and improving performance.

The main contribution of our work is on experiences and shared intentionality. Individual actions are dictated by individual intentions, which are the result of individual beliefs, and desires. But how do collective actions come about? Philosophers, especially those interested in action theory, have been more and more interested in the notion of collective actions, intentionality, and belief. Since these issues have not been addressed by current BDI approaches, we now discuss existing philosophical views that have addressed these issues.

Some argue that collective attitude simply refers to the fact that the majority of the group's members share that attitude [18]. These accounts were labelled as *summative* accounts by Gilbert [19]. He argued the insufficiency of simple summative accounts (SSA) and presented the complex summative accounts (CSA) as one approach to address *some* of their problems by introducing the notion of common knowledge. For example, a group G intends X if and only if most of the members of G intend X and it is common knowledge in G that most of the members of G intend X.

In our model, simple summative accounts are described by *common* modalities. For example, a common experience is an experience that has been experienced by the individuals of the group. Complex summative accounts are described by *joint* modalities. For example, a joint experience is an experience that is common to a group, and the members of the group are aware that they have all experienced the same experience.

Gilbert further argues that both the SSA and CSA accounts are not the right approaches since a group's attitude cannot always be described in terms of the individual ones, even if it was enhanced with common knowledge. For instance, a group of people who are jogging in the park are aware of each other's intentions to jog; yet, there is something different between this group and a group of friends who decide to jog together.

While our model does define common and joint modalities, we note that these collective modalities are simply used to describe the collection of individual modalities. They are not used to describe the groups' intentions as a whole. For that, the *shared* modality is introduced. A shared experience is a joint experience in which the agents are actively involved. A shared intention is an intention shared by the group members to carry out a predefined plan for fulfilling a joint desire.

Searle [20] states that the we-intend cannot be reduced to a set of I-intend, even if it was supplemented with mutual beliefs. He argues that collective intentions should combine the sense of acting with the sense of willing something together. In our model, the notion of shared experience does not necessarily imply that the different agents sharing an experience were willing to perform the actions they did together. However, the shared intention is defined as having a shared plan, where each agent intends to fulfil its part of the plan. The very definition of an agent's intention is its willingness and commitment to execute the corresponding plan. As such, a shared intention is realised only if the agents are willing and committing to execute their actions as detailed by the plan agreed upon.

Bratman [21], like Searle, does not believe in a plural agent (representing a group as a whole) that could hold shared intentions, nor does he believe that shared intentions can be reduced to individual intentional states. In his definition, an intention is shared if and only if our intentional actions are coordinated by making sure our personal plans of action meld together. As such, shared intentions give rise to argumentation and negotiation for agreeing on the coordinated plans of actions. Bratman further argues, like Searle, that a single agent can have a we-intention, which has attracted a great deal of criticism [22], [23], since one cannot intend what he/she cannot fulfil (in other words, one cannot intend actions that may only be executed by others).

Similar to Bratman, we say an intention is shared only if there is a shared plan of action that meshes and coordinates individual actions, although the level of coordination remains loose. For example, the plan may state who does what, without going into the order or the pre and post conditions of actions. Unlike Bratman and Searle, we say a shared intention cannot be held by a single agent (although single agents may form beliefs about shared intentions), but by the group of agents whom the plan of action cannot be carried out without. Agents can only intend to perform their part of a shared plan.

VI. Conclusion

This paper presented a formal agent model that introduces the notion of experience, which situates agents in their environment, providing a clear and tangible link that was missing in past BDI approaches and that help describe how the environment (and its changes) impact the evolution of an agent's BDI and vice versa.

We also built a logical formalism, the X-BDI, to describe the agent's reasoning and whose semantics are based on our proposed agent model. The main novel contribution of this work is defining, through the X-BDI logic, the notion of shared experience and using it as a fundamental construct to underpin collective intentionality, which we believe are the basis of social behaviour and the foundation of multiagent systems in general. While BDI was used as a model for individual agent reasoning, X-BDI can be used (as our example illustrates) as a model for collective agents' reasoning by illustrating when, why, and how do social interactions take place.

Acknowledgment

This work is supported by: the ACE ERA-Net project; the CBIT project (funded by the Spanish Ministry of Science & Innovation TIN2010-16306); and the Agreement Technologies project (CONSOLIDER CSD 2007-0022, INGENIO 2010).

References

[1] M. d'Inverno and M. Luck, "Engineering AgentSpeak(L): A formal computational model," *Logic and Computation*, vol. 8, no. 3, pp. 233–260, 1998.

[2] R. H. Bordini, J. F. Hübner, and M. Wooldridge, *Programming Multi-Agent Systems in AgentSpeak using Jason*. John Wiley & Sons, Ltd, 2007.

[3] M. Winikoff, "Jack intelligent agents: An industrial strength platform," in *Multi-Agent Programming*, ser. Multiagent Systems, Artificial Societies, and Simulated Organizations, R. H. Bordini, M. Dastani, J. Dix, and A. E. Fallah-Seghrouchni, Eds. Springer, 2005, vol. 15, pp. 175–193.

[4] L. Amgoud, M. d'Inverno, N. Osman, H. Prade, and C. Sierra, "Experiences - a forgotten component of epistemic states," in *Proceedings of the First International Conference on Agreement Technologies*, ser. CEUR Workshop Proceedings, S. Ossowski, F. Toni, and G. A. Vouros, Eds., vol. 918. CEUR-WS.org, 2012, pp. 220–230.

[5] S. Schütz-Bosbach and W. Prinz, "Prospective coding in event representation," *Cognitive Processing*, vol. 8, no. 2, pp. 93–102, 2007.

[6] D. I. Perrett, D. Xiao, N. E. Barraclough, C. Keysers, and M. W. Oram, "Seeing the future: Natural image sequences produce "anticipatory" neuronal activity and bias perceptual report," *The Quarterly Journal of Experimental Psychology*, vol. 62, no. 11, pp. 2081–2104, 2009.

[7] N. Maudet, S. Parsons, and I. Rahwan, "Argumentation in multi-agent systems: Context and recent developments," in *Argumentation in Multi-Agent Systems*, ser. Lecture Notes in Computer Science, N. Maudet, S. Parsons, and I. Rahwan, Eds. Springer Berlin Heidelberg, 2007, vol. 4766, pp. 1–16.

[8] B. J. Grosz, L. Hunsberger, and S. Kraus, "Planning and acting together," *AI Magazine*, vol. 20, no. 4, pp. 23–34, 1999.

[9] M. Bratman, *Intentions, Plans, and Practical Reason*. Cambridge, MA, 1987.

[10] A. S. Rao, "AgentSpeak(L): BDI agents speak out in a logical computable language," in *MAAMAW*, 1996, pp. 42–55.

[11] K. Hindriks, M. d'Inverno, and M. Luck, "Architecture for agent programming languages," in *Proceedings of the 14th European Conference on Artificial Intelligence*, 2000, pp. 363–367.

[12] M. Dastani, "2APL: a practical agent programming language," *Autonomous Agents and Multi-Agent Systems*, vol. 16, no. 3, pp. 214–248, 2008.

[13] A. S. Rao and M. P. Georgeff, "Modeling rational agents within a BDI-architecture," in *KR*, 1991, pp. 473–484.

[14] S. Parsons and P. Giorgini, *An approach to using degrees of belief in BDI agents*. Kluwer, 1999, pp. 81–92.

[15] A. Casali, L. Godo, and C. Sierra, "A graded BDI agent model to represent and reason about preferences," *Artif. Intell.*, vol. 175, no. 7-8, pp. 1468–1478, 2011.

[16] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *AI Commun.*, vol. 7, no. 1, pp. 39–59, 1994.

[17] S. Ontañón and E. Plaza, "Learning when to collaborate among learning agents," in *Proceedings of the 12th European Conference on Machine Learning*. London, UK: Springer-Verlag, 2001, pp. 394–405.

[18] A. Quinton, "The presidential address: Social objects," *Proceedings of the Aristotelian Society*, vol. 76, pp. 1–27+viii, 1975.

[19] M. P. Gilbert, *On social facts*, ser. International library of philosophy. New York, NY: Routledge, 1989.

[20] J. R. Searle, *The construction of social reality*. Free Press, 1995.

[21] M. E. Bratman, *Faces of Intention*, ser. Cambridge Studies in Philosophy. Cambridge University Press, 1999.

[22] F. Stoutland, "Why are philosophers of action so anti-social?" in *Commonality and particularity in ethics*, ser. Swansea studies in philosophy, L. Alanen, S. Heinämaa, and T. Wallgren, Eds. New York, NY: St. Martin's Press, 1997.

[23] J. D. Velleman, "How to share an intention," *Philosophy and Phenomenological Research*, vol. 57, no. 1, pp. 29–50, 1997.