

Beliefs in Multiagent Planning: From One Agent to Many

Filippos Kominis

Universitat Pompeu Fabra
08018 Barcelona, Spain
filippos.kominis@upf.edu

Hector Geffner

ICREA & Universitat Pompeu Fabra
08018 Barcelona, Spain
hector.geffner@upf.edu

Abstract

Single-agent planning in partially observable settings is a well understood problem and existing planners can represent and solve a wide variety of meaningful instances. In the most common formulation, the problem is cast as a non-deterministic search problem in belief space where beliefs are sets of states that the agent regards as possible. In this work, we build on the methods developed for representing beliefs in single-agent planning to introduce a simple but expressive formulation for handling beliefs in multi-agent settings. The resulting formulation deals with multiple agents that can act on the world (physical or ontic actions), and can sense either the state of the world (truth of objective formulas) or the mental state of other agents (truth of epistemic formulas). The formulation captures and defines a fragment of dynamic epistemic logics that is simple and expressive but which does not involve event models or product updates, and has the same complexity of belief tracking in the single agent setting and can benefit from the use of similar techniques. We show indeed that the problem of *computing multiagent linear plans* can be actually compiled into a *classical planning problem* using the techniques that have been developed for compiling conformant and contingent problems in the single agent setting and *report experimental results*.

Introduction

Single-agent planning in partially observable settings is a well understood problem and existing planners can represent and solve a wide variety of meaningful instances. In the most common formulation, single-agent planning in partially observable environments is cast as a non-deterministic search problem in belief space where the beliefs are sets of states that the agent regards as possible (Bonet and Geffner 2000). The work in partially observable or contingent planning has been focused on ways for representing beliefs and selecting actions (Bertoli et al. 2001; Brafman and Hoffmann 2004; Albore, Palacios, and Geffner 2009; To, Pontelli, and Son 2011; Brafman and Shani 2012a).

Current approaches for representing beliefs in multiagent dynamic settings, on the other hand, are based on Kripke structures (Fagin et al. 1995). Multiagent Kripke structures are triplets defined by a set of worlds, accessibility relations

among the worlds for each of the agents, and truth valuations that define the propositions that are true in each world. While a truth valuation determines the *objective* formulas that are true in a world, the accessibility relation among worlds provides the truth conditions for *epistemic* formulas.

Dynamic epistemic logics extend epistemic logics with the ability to deal with change (van Ditmarsch, van der Hoek, and Kooi 2007a; van Ditmarsch and Kooi 2008; Van Benthem 2011). The standard approach relies on *event models* and *product updates* by which both the agent beliefs and the events are represented by Kripke structures, and the resulting beliefs are captured by a suitable cross product of the two (Baltag, Moss, and Solecki 1998; Baltag and Moss 2004). Syntactically, axiomatizations have been developed to capture the valid inferences in such a setting, and a number of approaches have been developed to facilitate modeling and inference (Baral et al. 2012; Herzig, Lang, and Marquis 2005). A simple form of planning, however, where an event sequence is sought to achieve a given goal formula, has been shown to be undecidable in dynamic epistemic logic (Aucher and Bolander 2013), while decidable subsets have been identified as well (Löwe, Pacuit, and Witzel 2011).

In this work, we build on the methods developed for representing beliefs in single-agent planning to introduce a simple but expressive formulation for handling beliefs in multi-agent settings. The resulting formulation deals with multiple agents that can act on the world (physical or ontic actions), and can sense either the state of the world (truth of objective formulas) or the mental state of other agents (truth of epistemic formulas). The formulation captures and defines a fragment of dynamic epistemic logics that is simple and expressive, but which does not involve event models or product updates, and has the same complexity of belief tracking in the single agent setting and can benefit from the use of similar techniques. We show indeed that the problem of computing *linear multiagent plans* (Bolander and Andersen 2011) can be actually compiled into a *classical planning problem*, using the techniques that have been developed for compiling conformant and contingent problems in the single agent setting (Palacios and Geffner 2009; Brafman and Shani 2012b).

The proposed formulation exploits certain conventions and restrictions. First, while the agents can have private in-

formation as they have private sensors, they are all assumed to start with a *common initial belief* on the set of worlds that are possible. Second, the effects of physical actions on the world are assumed to be *deterministic*. And third, the *sequence of events* (physical actions, sensing events, and public announcements) that can change the state of the world or the knowledge state of the agents, is *public* to all the agents. In the formulation it is crucial to distinguish between the event of sensing the truth value of an objective or epistemic formula, and the agent coming to know that the formula is true or false. While the sensing event is public, as when all agents know the sensor capabilities of the other agents, the actual information provided by these sensors is private. For example, in the muddy children problem (Fagin et al. 1995), every child i is assumed to be capable of sensing the truth value of the atoms m_j encoding whether child j is muddy for $j \neq i$, and every child knows that. Yet this doesn't mean that children have access to the truth values revealed by the sensors that are not their own. The formulation does imply however that agents know what the other agents may potentially know, as agents start with the same knowledge and then learn about the world or about other agents using sensing events that are public.¹

The rest of the paper is organized as follows. We start with a well known example and introduce the modeling language, the belief representation, and the (linear) planning problem. We then analyze other examples, formulate the compilation of the linear multiagent planning problem into classical planning, present experimental results, and discuss the relation to dynamic epistemic logic.

Example

Before proceeding with the details of the formulation, it will be useful to consider how a familiar example, the Muddy Children Puzzle will be represented (Fagin et al. 1995). We consider three agents $A = \{a, b, c\}$ and atoms m_x , for $x \in A$, each representing that child x is muddy. The states of the problem are the possible truth valuation over these three atoms, and the common initial belief state b_I is given by the set of all such states (8 in total).

Consider then the sequence of events σ given by:

$$\text{update}(m_a \vee m_b \vee m_c), [\text{sense}(a, [m_b, m_c]), \text{sense}(b, [m_a, m_c]), \text{sense}(c, [m_a, m_b])] , \quad (1)$$

that includes the public announcement made by the father, followed by each agent sensing in parallel whether each of the other children is muddy or not. The event $\text{sense}(a, \phi)$ expresses that agent a senses the truth value of formula ϕ . Variations of these events, expressed as $\text{sense}(a, [\phi_1, \dots, \phi_n])$, $\text{sense}([a, b], [\phi_1, \dots, \phi_n])$, and $\text{sense}([\phi_1, \dots, \phi_n])$ represent that agent a senses the truth value of each of the formulas ϕ_i , $i = 1, \dots, n$, in parallel, that both a and b sense such truth values in parallel, and that all agents sense them

¹The assumptions in the model have points in common with the finitary S5 theories (Son et al. 2014) and with the notion of “only knowing” (Levesque 1990; Halpern and Lakemeyer 2001).

in parallel. In addition, groups of sensing events can be enclosed in brackets as in (1), meaning that the events are in parallel.

A possible query after the sequence of events σ may be whether any of the agents will know that he is muddy if the world is such that there is just one muddy child. This query would amount to testing the formula

$$m_a \oplus m_b \oplus m_c \Rightarrow K_a m_a \vee K_b m_b \vee K_c m_c \quad (2)$$

in (the Kripke structure associated with) the situation resulting from the common initial belief state b_I and the event sequence σ . In this formula, ‘ \oplus ’ stands for “exclusive or”; $p \oplus q$ thus being an abbreviation of $(p \vee q) \wedge \neg(p \wedge q)$. The answer to this query will be positive. On the other hand, the answer to the query:

$$\neg(m_a \oplus m_b \oplus m_c) \Rightarrow K_a m_a \vee K_b m_b \vee K_c m_c \quad (3)$$

will be negative, as when there is more than one muddy child, no child will know that he is muddy from the announcement made by the father and the information gathered from his physical sensors alone. It can be shown, however, that if the event sequence σ is extended with the following parallel sensing event:

$$[\text{sense}(K_a m_a), \text{sense}(K_b m_b), \text{sense}(K_c m_c)] \quad (4)$$

where all agents learn whether each of the agents knows that he is muddy, a formula like

$$m_a \wedge m_b \wedge \neg m_c \Rightarrow K_a m_a \quad (5)$$

will become true, as in the world where a and b are muddy and c is not, the sensing captured by (4) would result in a learning that b does not know that b is muddy ($K_a \neg K_b m_b$), while in the other world that is possible to a , where a is not muddy, a would come to learn the opposite; namely that b knows that b is actually muddy ($K_a K_b m_b$).

Language

We consider planning problems $P = \langle A, F, I, O, N, U, G \rangle$ where A is a set of agent names, F is a set of atoms, I is the initial situation, O is a set of *physical actions*, N is a set of *sensing actions*, U is set of *public (action) updates*, and G is the goal. A *plan* for P , as in classical planning, is a *sequence of actions* for achieving the goal G from the initial situation described by I . The main differences to classical planning result from the uncertainty in the initial situation, and the beliefs of the *multiple agents* involved. In addition the actions may come from any of the sets O , N , or U . If we let S stand for the set of all possible truth-valuations s over the atoms in F and call such valuations *states*, we assume that I is an objective formula over F which denotes a non-empty set of possible initial states b_I . A physical action a in O denotes a *deterministic* state-transition function f_a that maps any state s into a state $s' = f_a(s)$. A (parallel) sensing action in N is a set of expressions of the form $\text{sense}[A_k](\phi_k)$, where A_k is a non-empty set of agent names and ϕ_k is an objective or epistemic formula over the atoms F and the knowledge

modalities K_i for $i \in A$. The action updates in U are denoted by expressions of the form **update**(ϕ) where ϕ is a formula. Finally, each action a has a precondition $Pre(a)$, which like the goal G are formulas as well. The grammar of these formulas can be expressed as:

$$\phi = p \mid \neg\phi \mid (\phi \wedge \psi) \mid (\phi \Rightarrow \psi) \mid K_i\phi$$

where p is an atom in F , and i an agent in A .

We regard *plans* as linear sequences of actions (Bolander and Andersen 2011), and call P a *linear multiagent planning problem*. While many problems require *non-linear plans*, as it is the case in contingent planning, *linear plans* suffice for a number of non-trivial contexts and provide the basis for more complex forms of plans. These linear plans involve sensing, however, but like conformant plans, no conditional branching.

Belief Update and Dynamics

In order to define the belief representation and dynamics, let us represent the event sequences or plans σ over a problem P by sequences of the form $e(0), \dots, e(t)$, where $e(k)$ is the event from P that occurs at time k . When convenient, we will assume that the agent names are positive numbers i , $i = 1, \dots, m$ for $m = |A|$, or that they can be enumerated in this way.

The beliefs of all the agents at time step t , called also the joint belief, will be denoted as $B(t)$, and it is represented by a vector of *conditional beliefs* $B(s, t)$, where s is one of the possible initial states, $s \in b_I$; namely,

$$B(t) = \{B(s, t) \mid s \in b_I\}. \quad (6)$$

The conditional beliefs $B(s, t)$ represent the beliefs of all the agents at time t , *under the assumption that the true but hidden initial state is s* . The reason for tagging beliefs with possible initial states is that for a *fixed (hidden) initial state s , the evolution of the beliefs $B(s, t)$ after an arbitrary event sequence is deterministic*. These conditional beliefs $B(s, t)$ are in turn represented by tuples:

$$B(s, t) = \langle v(s, t), r_1(s, t), r_2(s, t), \dots, r_m(s, t) \rangle \quad (7)$$

where $v(s, t)$ is the state of the world that results from the initial state s after the event sequence $e(0), \dots, e(t-1)$, and $r_i(s, t)$ is the set of possible *initial* states $s' \in b_I$ that agent i cannot distinguish at time t from the actual initial state s . Note that s may be the true initial state, and yet the agents may not know about it. Indeed, initially, they only know that if s is the true initial state, it must be part of the initial common belief b_I .

More precisely, the initial beliefs $B(s, t)$ at time $t = 0$ are given by:

$$v(s, 0) = s \text{ and } r_i(s, 0) = b_I \quad (8)$$

for all agents i , meaning that under the assumption that the hidden initial state is s and that no events have yet occurred, the actual state is s and the set of possible initial states is b_I .

The belief $B(t+1)$ at time $t+1$ is a function of the belief $B(t)$ and event $e(t)$ at time t :

$$B(t+1) = \mathbf{F}(B(t), e(t)) \quad (9)$$

We express this function by defining how the *type of event* $e(t)$ at time t affects the state $v(s, t+1)$ and the relations $r_i(s, t+1)$ that define the belief $B(t+1)$ at time $t+1$.

Physical Actions: If $e(t) = \mathbf{do}(a)$ for action a denoting a state-transition function f_a , then the current state $v(s, t)$ associated with the hidden initial state s changes according to f_a , but the sets of initial states $r_i(s, t)$ that agent i regards as possible remain unchanged

$$v(s, t+1) = f_a(v(s, t)) \quad (10)$$

$$r_i(s, t+1) = r_i(s, t) \quad (11)$$

where the index i ranges over all the agents in A .

All the other event types affect instead the sets $r_i(s, t+1)$ but not the state $v(s, t+1)$ that is regarded as current given the assumption that s is the true initial hidden state. That is, for the following event types $v(s, t+1) = v(s, t)$.

Sensing: If $e(t) = [\mathbf{sense}[A_1](\phi_1), \dots, \mathbf{sense}[A_l](\phi_l)]$ is a sensing action denoting the set of sensing expressions $\mathbf{sense}[A_k](\phi_k)$ done in parallel at time t , the current state given s does not change, but the set of possible initial states compatible with the hidden initial state s for agent i given by $r_i(s, t+1)$ becomes:

$$\{s' \mid s' \in r_i(s, t) \text{ and } B(t), s' \models \phi_k \text{ iff } B(t), s \models \phi_k\} \quad (12)$$

where k ranges over all the indices in $[1, l]$ such that A_k includes agent i . If there are no such indices, $r_i(s, t+1) = r_i(s, t)$. The expression $B(t), s \models \phi$ denotes that ϕ is true in the belief at time t conditional on s being the true hidden state. The truth conditions for these expressions are spelled out below.

Updates: If $e(t) = \mathbf{update}(\phi)$, $r_i(s, t+1)$ is

$$\{s' \mid s' \in r_i(s, t) \text{ and } B(t), s' \models \phi\}. \quad (13)$$

The intuition for all these updates is the following. *Physical actions* change the current state of the world according to their state transition function. *Sensing actions* do not change the world but yield information. More specifically, when agent i senses the truth value of formula ϕ at time t , the set of initial states $r_i(s, t+1)$ that he thinks possible under the assumption that the true initial state is s , preserves the states s' in $r_i(s, t)$ that agree with s on the truth value predicted for ϕ at time t . Finally, a *public update* ϕ preserves the possible initial states s' in $r_i(s, t)$ that predict the formula ϕ to be true, and rules out the rest. The conditions under which a possible initial state s predicts that a formula ϕ will be true at time t , and the conditions under which a formula ϕ is true at time t , are made explicit below. Physical, sensing, and update actions are *applicable* at time t only when their preconditions are true at t .

Beliefs and Kripke Structures

A Kripke structure is a tuple $\mathcal{K} = \langle W, R, V \rangle$, where W is the set of worlds, R is a set of binary accessibility relations R_i on W , one for each agent i , and V is a mapping from the worlds w in W into truth valuations $V(w)$. The conditions under which an arbitrary formula ϕ is true in a world w of a Kripke structure $\mathcal{K} = \langle W, R, V \rangle$, written $\mathcal{K}, w \models \phi$, are defined inductively:

- $\mathcal{K}, w \models p$ for an atom p , if p is true in $V(w)$,
- $\mathcal{K}, w \models \phi \vee \psi$ if $\mathcal{K}, w \models \phi$ or $\mathcal{K}, w \models \psi$,
- $\mathcal{K}, w \models (\phi \Rightarrow \psi)$ if $\mathcal{K}, w \models \phi$ implies $\mathcal{K}, w \models \psi$,
- $\mathcal{K}, w \models K_i \phi$ if $\mathcal{K}, w' \models \phi$ for all w' s.t. $R_i(w, w')$, and
- $\mathcal{K}, w \models \neg \phi$ if $\mathcal{K}, w \not\models \phi$

A formula ϕ is valid in the structure \mathcal{K} , written $\mathcal{K} \models \phi$, iff $\mathcal{K}, w \models \phi$ for all worlds w in \mathcal{K} . The conditions under which a possible initial state s predicts the truth of a formula ϕ at time t , written $B(t), s \models \phi$, follow from replacing the belief $B(t)$ by the Kripke structure $\mathcal{K}(t) = \langle W^t, R^t, V^t \rangle$ defined by $B(t)$ where

- $W^t = \{s \mid s \in Poss(t)\}$,
- $R_i^t = \{(s, s') \mid \text{if } s' \in r_i(s, t)\}$,
- $V^t(s) = v(s, t)$

In these expressions, $Poss(t)$ stands for the initial states that remain possible at t ; $Poss(t) = \cup_{s \in b_I} \cup_{i=1, \dots, m} r_i(s, t)$. The worlds w in the structure $\mathcal{K}(t)$ are the possible initial states $s \in b_I$ that have not been ruled out by the updates. The worlds that are accessible from a world s to the agent i are the possible initial states s' that are in $r_i(s, t)$. Last, the valuation associated to a world s in this structure is the state $v(s, t)$ that deterministically follows from the possible initial state s and the event sequence up to $t-1$. $B(t), s \models \phi$ is thus true when $\mathcal{K}(t), s \models \phi$ is true, and $B(t) \models \phi$ iff $\mathcal{K}(t) \models \phi$. It is simple to show that the accessibility relations $R_i(t)$ are reflexive, symmetric, and transitive, meaning that the valid formulas satisfy the axioms of the epistemic logic S5.

Examples

We will show later that a linear multiagent problem P can be translated into a classical planning problem and solved by off-the-shelf planners. Before presenting such a translation, we consider two other examples.

Selective Communication

Let a, b , and c be three agents in a corridor of four rooms (p_1, p_2, p_3 and p_4 from left to right). The agents can move from a room to a contiguous room, and when agent i communicates (tells) some information, all the agents that are in the same room or in a contiguous room, will hear what was communicated. For example, if agent i expresses in room p_3 his knowledge about q , all agents in rooms p_2, p_3 and p_4 will come to know it. We consider the problem where agent a is initially in room p_1 , b in p_2 , c in p_3 , and a has to find out the truth value of a proposition q and let c know without agent b learning it. The planning problem is encoded as the tuple $P = \langle A, F, I, O, N, U, G \rangle$ where

$A = \{a, b, c\}$, $F = \{q\} \cup \{p(x, i)\}$, $x \in A$, $i \in [1, 4]$, $I = \{p(a, 1), p(b, 2), p(c, 3)\} \cup D$, where D contains the formulas expressing that each agent is in a single room, U is empty, and the goal is

$$G = (K_c q \vee K_c \neg q) \wedge (\neg K_b q \wedge \neg K_b \neg q) .$$

The set of physical actions is $O = \{right, left\}$ affecting the location of agent a in the obvious way (the actions have no effects when they'd move the agent away from the four rooms).

The sensing actions in N are two: the first about a learning the value of q when in p_2 , the other, about a expressing his knowledge regarding q , which translates into agents b and c learning this when they are close enough to a . The first sensing action is thus **sense**(a, q) with the precondition $p(a, 2)$, and the second is

$$\mathbf{tell}(a, q) : [\mathbf{sense}(b, \phi_b \Rightarrow K_a q), \mathbf{sense}(b, \phi_b \Rightarrow K_a \neg q), \\ \mathbf{sense}(c, \phi_c \Rightarrow K_a q), \mathbf{sense}(c, \phi_c \Rightarrow K_a \neg q)] ,$$

where **tell**(a, q) is the abbreviation of the action that we will use, and ϕ_b is the formula expressing that agent b is at distance less than 1 from agent a ; namely $\phi_b = \vee_{i,j} [p(a, i) \wedge p(b, j)]$ for i and j in $[1, 4]$ such that $|i - j| \leq 1$. The formula ϕ_c is similar but with c instead of b .

Initially, b_I contains the two states s_1 and s_2 satisfying I , the first where q is true, and the second where it is false. The initial belief at time $t = 0$ is $B(t) = \{B(s_1, t), B(s_2, t)\}$, where $B(s_i, t) = \langle v(s_i, t), r_a(s_i, t), r_b(s_i, t), r_c(s_i, t) \rangle$, $i = 1, 2$, and $r_x(s, t) = b_I$ for $x \in A$ and $s \in b_I$. The shortest plan is

$$\mathbf{do}(right), \mathbf{sense}(a, q), \mathbf{do}(right), \mathbf{do}(right), \mathbf{tell}(a, q) .$$

The first sensing action can be done because its precondition $p(a, 2)$ holds in $B(1)$, and as an effect it removes agent a 's uncertainty regarding q making $r_a(s_1, 2) = \{s_1\}$ and $r_a(s_2, 2) = \{s_2\}$. Agent a then knows whether q is true or false, and in principle, he could communicate this from his current location p_2 by performing the action **tell**(a, q) right away. But since the condition ϕ_b is true, b would come to know whether q is true, making the problem goal G unachievable. The effect of the two *right* actions is to make $p(a, 4)$ true, and all other $p(a, i)$ atoms false, thus making the formula ϕ_b false and the formula ϕ_c true (i.e., agent a is now near c but not near b). The final event in the plan makes the truth value of q known to agent c but not to agent b , thus achieving the goal G . The first part follows because the state $v(s_1, 5)$ where agent a is at p_4 and q is true, makes the formula $\phi_c \Rightarrow K_a q$ sensed by agent c true, while the state $v(s_2, 5)$ makes this formula false, and similarly, the state $v(s_2, 5)$ makes the formula $\phi_c \Rightarrow K_a \neg q$ sensed by agent c true, while the state $v(s_1, 5)$ makes it false. As a result, the state s_2 is not in $r_c(s_1, 5)$, the state s_1 is not in $r_c(s_2, 5)$, both sets become singletons, and hence, the truth value of q becomes known to agent c . The same reasoning does not apply to agent b because the condition ϕ_b is false in the two states $v(s_1, 5)$ and $v(s_2, 5)$, and hence, both states trivially satisfy the formulas $\phi_b \Rightarrow K_a q$ and $\phi_b \Rightarrow K_a \neg q$ that are sensed by agent b , so that $r_b(s_1, 5)$ and $r_b(s_2, 5)$ remain unchanged, and equal to b_I .

Collaboration through Communication

As a third example, we consider a scenario where two agents volunteer information to each other in order to accomplish a task faster that would otherwise be possible without information exchange. It is inspired in the BW4T environment, a proposed testbed for joint activity (Johnson et al. 2009). There is a corridor of four rooms, p_1, p_2, p_3 and p_4 as in the previous example, four blocks b_1, \dots, b_4 that are in some of the rooms, and two agents a and b that can move back and forth along this corridor. Initially, the two agents are in p_2 and do not know where the blocks are (they are not in p_2). When an agent gets into a room, he can see which blocks are in the room if any. The goal of the planning problem is for agent a to know the position of block b_1 , and for agent b to know the position of block b_2 . A shortest plan for the problem involves six steps: one agent, say a , has to move to p_1 , the other agent has to move to p_3 , they both must sense which blocks are in these rooms, and then they must exchange the relevant information. At that point, the goal would be achieved whether or not the information exchanged explicitly conveys the location of the target blocks. Indeed, if agent a does not see block b_1 in p_1 and agent b doesn't see this block either at p_3 , agent a will then know that block b_1 must be in p_4 once b conveys to a the relevant piece of information; in this case $\neg K_b \text{in}(b_1, p_3)$.

The planning problem is $P = \langle A, F, I, O, N, U, G \rangle$, where $A = \{a, b\}$, $F = \{at(x, p_k), \text{in}(b_i, p_k)\}$, $x \in A$, $i, k \in [1, 4]$, $I = \{at(a, p_2), at(b, p_2)\} \cup D$, where D contains the formulas expressing that each block has a unique location. The set of updates U is empty, the goal is $G = (\bigvee_{k=1,4} K_a at(b_1, p_k)) \wedge (\bigvee_{k=1,4} K_b at(b_2, p_k))$, the actions in O are right_x and left_x , for each agent $x \in A$ with the same semantics as in the example above, while the sensing actions are $\text{sense}(x, [\text{in}(b_1, p_k), \dots, \text{in}(b_4, p_k)])$ with precondition $at(x, p_k)$ by which agent $x \in A$ finds out in parallel which blocks b_i , if any, are and are not in p_k , and $\text{sense}(x, [K_y \text{in}(b_i, p_k)])$, by which agent y communicates to agent $x \neq y$, whether he knows $\text{in}(b_i, p_k)$, $i, k \in [1, 4]$. There are thus four physical actions, eight actions that sense the world, and thirty-two communication actions. A shortest plan is:

$\text{do}(\text{left}_a), \text{do}(\text{right}_b), \text{sense}(a, [\text{in}(b_1, p_1), \dots, \text{in}(b_4, p_1)]),$
 $\text{sense}(b, [\text{in}(b_1, p_3), \dots, \text{in}(b_4, p_3)]),$
 $\text{sense}(a, K_b \text{in}(b_1, p_3)), \text{sense}(b, K_a \text{in}(b_2, p_1)).$

This sequential plan achieves the goal in spite of the uncertainty of the agents about the world and about the beliefs of the other agents.

Translation into Classical Planning

We show next how a *linear multiagent planning problem* P can be compiled into a classical planning problem $K(P)$ such that the plans for P are the plans for $K(P)$. The language for $K(P)$ is STRIPS extended with *negation, conditional effects*, and *axioms*. This is a PDDL fragment supported by several classical planners. We will use $\neg L$ for a literal L to stand for the complement of L , so that $\neg\neg L$ is L . A conditional effect is an expression of the form $C \rightarrow E$

associated with an action a that states that the head E becomes true when a is applied and C is true. We write such effects as $a : C \rightarrow E$ when convenient. In addition planners normally assume that C and E are sets (conjunctions) of literals. If $C, C' \rightarrow E$ is one such effect, we take $C, \neg C' \rightarrow E$ as a shorthand for the effects $C, \neg L \rightarrow E$ for each literal L in C' . *Axioms* allow the definition of new, derived atoms in terms of primitive ones, called then the primitive fluents. The derived fluents can be used in action preconditions, goals, and in the body of conditional effects. While it's possible to compile axioms away, there are benefits for dealing with them directly in the computation of heuristics and in state progression (Thiébaux, Hoffmann, and Nebel 2005).

For mapping the multiagent problem $P = \langle A, F, I, O, N, U, G \rangle$ into the classical problem $K(P)$, we will make some simplifying assumptions about the types of formulas that may appear in P . We will assume as in planning, and without loss of generality, that such formulas correspond to conjunctions of literals, where a literal L is an (objective) atom p from F or its negation, or an epistemic literal $K_i L$ or $\neg K_i L$ where L is a literal and i is an agent in A . Other formulas, however, can easily be accommodated by adding extra axioms to $K(P)$. We will denote the set of objective literals in P by $L_F(P)$; i.e., $L_F(P) = \{p, \neg p \mid p \in F\}$, and the set of positive epistemic literals appearing in P by $L_K(P)$; i.e., $L_K(P)$ is the set of $K_i L$ literals that appear as subformula of an action precondition, condition, goal, or sensing or update expression. Indeed, while the set of $K_i L$ literals is infinite, as they can be arbitrarily nested, the set of such literals appearing in P is polynomial in the size of P . As an example, if $\neg K_2 K_1 \neg K_3 p$ is a goal, then $L_K(P)$ will include the (positive epistemic) literals $K_3 p, K_1 \neg K_3 p$ and $K_2 K_1 \neg K_3 p$.

The translation $K(P)$ comprises the fluents L/s for the objective literals L in $L_F(P)$, and possible initial states $s \in b_I$, and fluents $D_i(s, s')$ for agents $i \in A$. The former express that the objective literal L is true given that s is the true initial state, while the latter that agent i can distinguish s from s' and vice versa. The epistemic literals $K_i L$ appearing in P , such as $K_3 p, K_1 \neg K_3 p$ and $K_2 K_1 \neg K_3 p$ above, are mapped into derived atoms in $K(P)$ through the use of axioms. The expression C/s where C is a conjunction of literals L stands for the conjunction of the literals L/s .

Definition 1. Let $P = \langle A, F, I, O, N, U, G \rangle$ be a linear multiagent planning problem. Then the translation $K(P)$ of P is the classical planning problem with axioms $K(P) = \langle F', I', O', G', X' \rangle$ where

- $F' = \{L/s : L \in L_F(P), s \in b_I\} \cup \{D_i(s, s') : i \in A, s, s' \in b_I\}$,
- $I' = \{L/s : L \in L_F(P), s \in b_I, s \models L\}$,
- $G' = G$,
- $O' = O \cup N \cup U$; i.e., same set of actions a with same preconditions $\text{Pre}(a)$, but with
 - effects $a : C/s \rightarrow E/s$ for each $s \in b_I$, in place of the effect $a : C \rightarrow E$ for physical actions $\text{do}(a)$, $a \in O$,
 - effects $a : C/s, \neg C/s' \rightarrow D_i(s, s'), D_i(s', s)$ for each

pair of states s, s' in b_I and (parallel) sensing actions a in N that involve a **sense**(i, C) expression, and

- effects $a : \neg C/s' \rightarrow D_i(s, s')$ for each pair of states s, s' in b_I and $i \in A$, for actions a of the form **update**(C),
- X' is a set of axioms:
 - one for each positive derived fluent $K_i L/s$ where $K_i L \in L_K(P)$ and $s \in b_I$ with (acyclic) definition $L/s \wedge \wedge_{s' \in b_I} [L/s' \vee D_i(s, s')]$,
 - one for each literal L in $L_F(P) \cup L_K(P)$ with definition $\wedge_{s \in b_I} [L/s \vee D_i(s, s)]$

In words, the primitive fluents in $K(P)$ represent the truth of the literals L in P conditioned on each possible hidden initial state s as L/s , and the (in)accessibility relation $D_i(s, s')$ among worlds. Initially, the worlds are all accessible from each other and $D_i(s, s')$ is false for all such pairs. On the other hand, L/s is true initially if L is true in s . The goal G' of $K(P)$ is the same as the (conjunctive) goal G of P , and the actions O' in $K(P)$ are the actions in the sets O , N , and U of P with the same preconditions. However, in the translation, the effect of physical actions is on the L/s literals, while the effect of sensing actions and updates is on the $D_i(s, s')$ literals, with the literals $D_i(s, s)$ for any i being used to denote that the world s is no longer possible. Last, the truth conditions for epistemic literals in the translation is expressed by means of axioms in terms of the primitive literals L/s and $D_i(s, s')$.

The complexity of the translation is quadratic in the number $|b_I|$ of possible initial states. Its soundness and completeness properties can be expressed as follows:

Theorem 1. An action sequence π is a plan that solves the linear multiagent planning problem P iff π is a plan that solves the classical planning problem with axioms $K(P)$.

The translation above follows the pattern of other translations developed for conformant and contingent planning problems in the single agent setting (Palacios and Geffner 2009; Albore, Palacios, and Geffner 2009; Brafman and Shani 2012a) and is closest to the one formulated by Brafman and Shani (2012b). Actually, Brafman, Shani and Zilberstein have recently developed a translation of a class of multiagent contingent planning problems that they refer to as Qualitative Dec-POMDPs (Brafman, Shani, and Zilberstein 2013), as it's a "qualitative" (logical) version of Decentralized POMDPs (Bernstein, Zilberstein, and Immerman 2000). A key difference with our linear multiagent planning problems is that in Q-Dec-POMDPs the agents have beliefs about the world, but not about each other. Hence there are no epistemic modalities or epistemic formulas.

Experimental Results

We have tested the translation above by taking a number of problems P and feeding the translations $K(P)$ into classical planners. The results are shown in Table 1.² As classical planners we used the version of FF known as FF-X

²Software and data at <http://www.dtic.upf.edu/~fkominis/>

(Thiébaux, Hoffmann, and Nebel 2005) that supports axioms and is available from J. Hoffmann, and three configurations of Fast Downward (Helmert 2006) in a version that we obtained from M. Helmert that does less preprocessing. The three configurations differ just on the heuristic that is used to guide an A* search which is *optimal* for the admissible h_{max} heuristic. The results have been obtained on a Linux machine running at 2.93 GHz with 4 GB of RAM and a cutoff of 30 minutes.

A couple of optimizations have been implemented in the translation $K(P)$. In particular, we take advantage of the symmetry of the $D_i(s, s')$ predicates to reduce these atoms in half. In addition, for sensing actions **sense**(i, C) where C is a *static* objective atom, we define the effects *unconditionally* for all pairs $s, s' \in b_I$ such that s and s' disagree on the truth value of C .

About the list of domains in the table, the first three have been discussed already: MuddyChildren(n) with n children, Collab-through-Comm(n) with n blocks, (only two blocks are relevant though), and Selective-Communication. The new domains are discussed below.

Active Muddy Child

MuddyChild(n, m) is a reformulation of MuddyChildren where a particular child must find out whether he is muddy or not. For this he can ask individually each other child i whether i knows that he is muddy, with all other children listening the response. Thus, while in MuddyChildren(n) there is just one epistemic sensing action that lets every child know whether each child knows that he is muddy, in MuddyChild(n, m), there are $n - 1$ epistemic actions depending on the child being asked. In addition, to make things more interesting, the goal in MuddyChild(n, m) is for the selected child k to find out whether he is muddy, *given* that m of the children are *not* muddy in the actual world. For example, in MuddyChild(5, 2), this goal can be encoded by the formula $(\neg m_1 \wedge \neg m_2) \supset (K_3 m_3 \vee K_3 \neg m_3)$. The result of this conditional goal is that in the resulting (shortest) plans, child 3 will not ask questions to children 1 and 2, as there is nothing to achieve in the worlds where either one of them is muddy. While this is not initially known, the child has physical sensors to discover that. Actually, in this domain, in order to represent the initial situation where the children have received the father's announcement and the information from their physical sensors, we force on all plans an initial sequence of actions that contain these $n + 1$ actions. This is easy to do by adding extra fluents. The shortest plans for MuddyChild(n, m) thus will involve these $n + 1$ actions followed by $n - m - 1$ epistemic actions.

Sum

Sum(n) is a domain based on "What is the Sum?" (van Ditmarsch, van der Hoek, and Kooi 2007b), which in turn borrows from the "Sum and Product Riddle" (van Ditmarsch, Ruan, and Verbrugge 2008) and the Muddy Children. There are three agents a, b , and c , each one with a number on his forehead between 1 and n . It is known that one of the numbers must be the sum of the other two. In addition, each agent can see the numbers on the other agent's foreheads, and can

Problems	#Atoms	#Actions	#Axioms	#States	A*(max)	A*(cea)	BFS(add)	FF-X
MuddyChildren(3)	212	5	72	8	(0.02 - 0.01) / 6	(0.02 - 0.02) / 6	(0.02 - 0.02) / 6	0.01 / 6
MuddyChildren(4)	816	6	192	16	(0.16 - 0.06) / 8	(0.1 - 0.01) / 8	(0.15 - 0.02) / 8	0.1 / 8
MuddyChildren(5)	3312	7	480	32	(1.64 - 1.1) / 10	(0.7 - 0.1) / 10	(0.8 - 0.22) / 10	3.6 / 10
MuddyChildren(6)	14080	8	1152	64	(24.5 - 20.1) / 12	(5.4 - 1.1) / 12	(8 - 3.3) / 12	87 / 12
MuddyChildren(7)	61504	9	2688	128	(360 - 311) / 14	(55.1 - 9) / 14	(109.8 - 64) / 14	-
Collab-and-Comm(2)	348	22	132	9	(0.1 - 0.04) / 6	(0.06 - 0.02) / 6	(0.06 - 0.02) / 6	0.05 / 8
Collab-and-Comm(3)	1761	28	546	27	(1.6 - 1.1) / 6	(0.8 - 0.25) / 6	(0.85 - 0.25) / 6	9.3 / 8
Collab-and-Comm(4)	10374	34	2112	81	(48.1 - 33) / 6	(20.3 - 5.3) / 6	(22 - 6.5) / 6	765 / 8
Selective-Comm	59	7	20	2	(0.01 - 0.01) / 9	(0.01 - 0.01) / 9	(0.01 - 0.01) / 9	0.01 / 9
MuddyChild(3,1)	180	6	40	8	(0.01 - 0.01) / 5	(0.01 - 0.01) / 5	(0.01 - 0.01) / 5	0.01 / 5
MuddyChild(4,1)	720	8	96	16	(0.1 - 0.02) / 7	(0.1 - 0.01) / 7	(0.1 - 0.02) / 7	0.05 / 7
MuddyChild(5,2)	3056	10	224	32	(1.3 - 0.06) / 8	(1.14 - 0.02) / 8	(1.2 - 0.06) / 8	1.75 / 8
MuddyChild(5,1)	3056	10	224	32	(1.3 - 0.08) / 9	(1.14 - 0.02) / 9	(1.2 - 0.08) / 9	1.82 / 9
MuddyChild(6,2)	13440	12	512	64	(23 - 0.6) / 10	(22.1 - 0.2) / 10	(22.6 - 0.7) / 10	50 / 10
MuddyChild(6,1)	13440	12	512	64	(23 - 0.6) / 11	(22.1 - 0.25) / 11	(22.7 - 0.7) / 11	51.5 / 11
MuddyChild(7,2)	59968	14	1152	128	(554.5 - 4.5) / 12	(551 - 1.5) / 12	(555 - 5.7) / 12	-
Sum(3)	306	10	90	9	(0.02 - 0.01) / 3	(0.02 - 0.01) / 3	(0.04 - 0.02) / 3	0.02 / 3
Sum(4)	963	13	234	18	(0.32 - 0.2) / 5	(0.2 - 0.02) / 5	(0.2 - 0.06) / 5	0.6 / 5
Sum(5)	2325	16	480	30	(26.5 - 26) / 7	(0.7 - 0.1) / 7	(0.8 - 0.25) / 7	9.1 / 7
Sum(6)	4770	19	855	45	-	(2.4 - 0.7) / 10	(3.2 - 1.5) / 10	53 / 10
Sum(7)	8757	22	1386	63	-	(7.5 - 2.9) / 11	(9.5 - 5.3) / 11	241 / 13
WordRooms(25,8)	935	56	535	8	(9.4 - 9.3) / 9	(0.25 - 0.1) / 11	(0.25 - 0.1) / 11	6.2 / 11
WordRooms(25,10)	1183	56	663	10	(18 - 17.8) / 9	(0.5 - 0.2) / 11	(0.5 - 0.2) / 11	11.9 / 11
WordRooms(25,12)	1439	56	791	12	(60 - 59.6) / 10	(0.6 - 0.26) / 14	(0.6 - 0.3) / 14	20.3 / 10
WordRooms(30,14)	1913	56	1059	14	(134.3 - 133.7) / 10	(1.1 - 0.5) / 15	(1.1 - 0.5) / 15	49.2 / 14
WordRooms(30,16)	2215	56	1207	16	(207 - 206) / 10	(1.5 - 0.7) / 15	(1.5 - 0.6) / 15	73 / 16

Table 1: Experimental results. Problems P shown on the left. The columns indicate number of atoms, actions, and axioms in $K(P)$, the number of possible initial states for P , and the resulting times and plan lengths. FF-X refers to the version of FF that supports axioms. The other columns refer to three different configurations of Fast Downward using the same search algorithm A^* and the heuristics h_{max} , h_{cea} and h_{add} . The first configuration yields provably shortest plans. In the FF-X column, X/Y stands for X seconds and plan length Y . For Fast Downward, X-Y/Z stands for X seconds of total time, Y seconds spent on the search, and plan length Z . Unsolved problems indicated as “-”.

be asked to publicly announce whether he knows that he has a specific number. The goal is for one selected agent or two to learn their numbers. Atoms x_i , for $x \in A = \{a, b, c\}$ and $1 \leq i \leq n$ are used for indicating that agent x has the number i on his forehead. We use one action that lets all agents know the numbers on the forehead of the other agents in parallel. In addition, there are $3n$ actions that let all agents sense whether agent x knows that he has the number i , $x \in A$ and $1 \leq i \leq n$.

The problem is subtle. Consider for example the smallest problem with $n = 3$ where agent a must learn his number, i.e., $G = K_a a_1 \vee K_a a_2 \vee K_a a_3$. Since the largest number must be the sum of the other two, and hence must be larger than the other two, these two other numbers can be 1 and 1, or 1 and 2. There are thus two different tuples of numbers that are possible, 1, 1, 2 and 1, 2, 3, to be distributed among the 3 agents, resulting into 9 possible (initial) states and $|b_I| = 9$.

If agent a sees that a second agent has the number 3, he will know his number from looking at the third agent: if he has number 2, then a must have number 1, and if the third agent has number 1, a must have number 2. On the other hand, if a sees only numbers 1 and 2, he will not know whether he has number 1 or 3. Yet he can ask the agent with the number 1 whether he knows that he has the number 1: if he knows, then a knows that he has number 3, else, he has number 1. These various scenarios can be obtained by setting the goal to an implication like $\neg a_3 \supset K_a a_1 \vee K_a a_2$.

The goals for the instances in the table do not involve conditions on the actual world and thus must work *for all the worlds that are possible*.

In Sum(3), the goal is for one agent, say a , to learn his number and the plan involves all agents sensing the numbers of the others in parallel, and then b and c reporting in sequence whether they each know that his own number is 1. The total number of actions in the plan is thus 3. There are three cases to consider to show that the plan works. If the report from b is $K_b b_1$, a and c must have the numbers 2 and 3, or 3 and 2, but since a can see c , he can figure out his number. Let us thus assume that the report from b is $\neg K_b b_1$ followed by c reporting $K_c c_1$. In such a case, from the first observation, agents a and c cannot have 2 and 3, or 3 and 2, and from the second, a_1 and b_1 cannot be both true either. Thus a and b must have the numbers 2 and 1, 2 and 3, or 3 and 2. Once again, since a can see b , a can figure out his number. Last, if the sensing results in $\neg K_b b_1$ followed by $\neg K_c c_1$, a and b must have the numbers 1 and 1, 1 and 2, or 1 and 3. Therefore a will be able to know that his number is 1.

Interestingly, there is *no* plan for the goal when all agents must learn their numbers. Let us assume that b reports first, and let us focus on two of the possible initial states where the numbers for a , b and c are 2,1,1 and 2,3,1 respectively. In state 2,1,1, a will know his number, and b will express ignorance, from which c will learn that his number is 1. Agent b can predict this, and hence will not learn anything else from

either a or c . Thus, the first agent that speaks up in the *linear* plan, won't be able to figure out his number in all states.

Word Room

WordRoom(m, n) is a variation of the collaboration through communication example. It involves two agents a and b that must find out a hidden word from a list of n possible words. The words can have at most 7 letters with the i -th letter of the word being in room r_i , $i = 1, \dots, 7$. The two agents can move from a corridor to each of the rooms, and from any room back to the corridor. There is no direct connection among rooms, the two agents cannot be in the same room, and they both start in the corridor. The agents have sensors to find out the letter in a room provided that they are in the room, and they can communicate the truth of the literals $K_x l_i$ where x is one of the two agents and l_i expresses that l is the i -th letter of the hidden word. The former amounts to 14 sensing actions of the form **sense**($x, [l_i, l'_i, l''_i, \dots]$) with the precondition that agent x is in room i , and where l, l', \dots are the different letters that may appear at position i of some of the n words. The parameter m in problem WordRoom(m, n) stands for the total number of l_i atoms. There are also 7 actions **sense**($a, [K_b l_i, K_b l'_i, K_b l''_i, \dots]$) where b communicates what he knows about room i to a , and similarly, 7 actions where a communicates to b . If we add the 14 actions for each agent moving from a room to the corridor and back, the total pool of actions is 56. The shortest plan for these problems is interesting when there is a lot of overlap among the n possible words, and in particular, when it may be more efficient for an agent to communicate *not* the letters that he has observed, but the letters that he can derive from what he knows.

Relation to Single Agent Beliefs and DEL

The proposed formulation for handling beliefs in a multiagent setting sits halfway between the standard formulation of beliefs in single agent settings as found in conformant and contingent planning (Geffner and Bonet 2013), and the standard formulation of beliefs in the multiagent settings as found in dynamic epistemic logics (van Ditmarsch, van der Hoek, and Kooi 2007a; van Ditmarsch and Kooi 2008). In the single agent settings, beliefs are represented as the sets of states b that are possible, and physical actions a , whether deterministic or not, affect such beliefs deterministically, mapping a belief b into a belief $b_a = \{s \mid s \in F(a, s') \text{ and } s' \in b\}$ where F represents the system dynamics so that $F(a, s)$ stands for the set of states that may follow action a in state s . If the action a is *deterministic*, $F(a, s)$ contains a single state. The belief resulting from doing action a in the belief b and getting an observation token o is $b_a^o = \{s \mid s \in b_a \text{ such that } o \in O(a, s)\}$ where O represents the *sensor model* so that $O(a, s)$ stands for the set of tokens that can be observed after doing action a , resulting in the (possibly hidden) state s . Sensing is noiseless or *deterministic*, when $O(a, s)$ contains a single token. Interestingly, when both the actions and the sensing are *deterministic*, the set of beliefs $B'(t)$ that may follow from an initial belief b_I and a given action sequence is $B'(t) = \{b(s, t) \mid s \in b_I\}$

where $b(s, t)$ is the unique belief state that results from the action sequence and the initial belief state b_I when s is the hidden state. This expression has indeed close similarities with the beliefs $B(t)$ defined by (6) and (7) above.

While the proposed formulation is an extension of the belief representation used in single-agent planning, it represents also a *fragment of dynamic epistemic logics* where the Kripke structure $\mathcal{K}(t+1)$ that represents the belief at time $t+1$ is obtained from the Kripke structure $\mathcal{K}(t)$ representing the beliefs at time t and the Kripke structure representing the event at time t called the *event model*. The update operation is known as the *product update* as the set of worlds of the new structure is obtained by taking the cross product of the sets of worlds of the two time t structures. In particular, using the framework laid out in (van Ditmarsch and Kooi 2008; Bolander and Andersen 2011) for integrating epistemic and physical actions, the basic actions in our language can be all mapped into simple event models. The event model for **do**(a) is given by a single event whose postcondition in a state s is $f_a(s)$. The event model for **update**(ϕ) has also a single event with precondition ϕ and null postcondition. Finally, the event model for **sense**(A, ϕ) has two events that can be distinguished by the agents in A but not by the other agents, one with precondition ϕ , the other with precondition $\neg\phi$, and both with null postconditions. While the proposed formulation captures only a fragment of dynamic epistemic logics, for this fragment, it provides a *convenient modeling language, a simple semantics, and a computational model*.

Discussion

We have introduced a framework for handling beliefs in the multiagent setting that builds on the methods developed for representing beliefs in single-agent planning. The framework also captures and defines a fragment of dynamic epistemic logics that does not require event models or product updates, and has the same complexity as belief tracking in the single agent setting (exponential in the number of atoms). We have also built on these connections to show how the problem of computing linear multiagent plans can be mapped into a classical planning problem, and have presented a number of examples and experimental results.

A basic assumption is that all uncertainty originates in the set of states that are possible initially and hence that actions are deterministic. Still, non-deterministic physical and sensing actions can be introduced by reducing them to deterministic actions whose effects are made conditional on extra hidden variables. Similarly, while all agents are assumed to start with the same belief state, different initial beliefs that result from a common belief and different public sensing events can be handled easily as well.

In the future, we want to explore more compact translations able to exploit width considerations as in conformant and contingent planning, and, like recent work in single-agent partially observable planning, move to the on-line setting where action selection may depend on the actual observations, and further scalability is achievable.

Acknowledgements: We thank Malte Helmert for the version of the Fast-Downward planner used in the experiments.

References

- Albore, A.; Palacios, H.; and Geffner, H. 2009. A translation-based approach to contingent planning. In *Proc. IJCAI-09*, 1623–1628.
- Aucher, G., and Bolander, T. 2013. Undecidability in epistemic planning. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, 27–33. AAAI Press.
- Baltag, A., and Moss, L. S. 2004. Logics for epistemic programs. *Synthese* 139(2):165–224.
- Baltag, A.; Moss, L. S.; and Solecki, S. 1998. The logic of public announcements, common knowledge, and private suspicions. In *Proc. of the 7th Conf. on Theoretical aspects of rationality and knowledge*, 43–56.
- Baral, C.; Gelfond, G.; Pontelli, E.; and Son, T. C. 2012. An action language for reasoning about beliefs in multi-agent domains. In *Proc. of the 14th International Workshop on Non-Monotonic Reasoning*.
- Bernstein, D.; Zilberstein, S.; and Immerman, N. 2000. The complexity of decentralized control of Markov decision processes. In *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence*, 32–37.
- Bertoli, P.; Cimatti, A.; Roveri, M.; and Traverso, P. 2001. Planning in nondeterministic domains under partial observability via symbolic model checking. In *Proc. IJCAI-01*.
- Bolander, T., and Andersen, M. B. 2011. Epistemic planning for single and multi-agent systems. *Journal of Applied Non-Classical Logics* 21(1):9–34.
- Bonet, B., and Geffner, H. 2000. Planning with incomplete information as heuristic search in belief space. In *Proc. of AIPS-2000*, 52–61.
- Brafman, R., and Hoffmann, J. 2004. Conformant planning via heuristic forward search: A new approach. In *Proc. ICAPS-04*.
- Brafman, R. I., and Shani, G. 2012a. A multi-path compilation approach to contingent planning. In *Proc. AAAI*.
- Brafman, R. I., and Shani, G. 2012b. Replanning in domains with partial information and sensing actions. *Journal of Artificial Intelligence Research* 45(1):565–600.
- Brafman, R. I.; Shani, G.; and Zilberstein, S. 2013. Qualitative planning under partial observability in multi-agent domains. In *Proc. AAAI*.
- Fagin, R.; Halpern, J.; Moses, Y.; and Vardi, M. 1995. *Reasoning about Knowledge*. MIT Press.
- Geffner, H., and Bonet, B. 2013. *A Concise Introduction to Models and Methods for Automated Planning*. Morgan & Claypool Publishers.
- Halpern, J. Y., and Lakemeyer, G. 2001. Multi-agent only knowing. *Journal of Logic and Computation* 11(1):41–70.
- Helmert, M. 2006. The Fast Downward planning system. *Journal of Artificial Intelligence Research* 26:191–246.
- Herzig, A.; Lang, J.; and Marquis, P. 2005. Action progression and revision in multiagent belief structures. In *Proc. 6th Workshop on Nonmonotonic Reasoning, Action, and Change (NRAC 2005)*.
- Johnson, M.; Jonker, C.; van Riemsdijk, B.; Feltovich, P.; and Bradshaw, J. M. 2009. Joint activity testbed: Blocks world for teams (bw4t). In *Engineering Societies in the Agents World X*. Springer. 254–256.
- Levesque, H. 1990. All I know: a study in autoepistemic logic. *Artificial intelligence* 42(2):263–309.
- Löwe, B.; Pacuit, E.; and Witzel, A. 2011. DEL planning and some tractable cases. In *Logic, Rationality, and Interaction*. Springer. 179–192.
- Palacios, H., and Geffner, H. 2009. Compiling Uncertainty Away in Conformant Planning Problems with Bounded Width. *Journal of Artificial Intelligence Research* 35:623–675.
- Son, T. C.; Pontelli, E.; Baral, C.; and Gelfond, G. 2014. Finitary S5-theories. In *Logics in Artificial Intelligence*. Springer. 239–252.
- Thiébaux, S.; Hoffmann, J.; and Nebel, B. 2005. In defense of pddl axioms. *Artif. Intell.* 168(1-2):38–69.
- To, S. T.; Pontelli, E.; and Son, T. C. 2011. On the effectiveness of CNF and DNF representations in contingent planning. In *Proc. IJCAI*, 2033–2038.
- Van Benthem, J. 2011. *Logical dynamics of information and interaction*. Cambridge University Press.
- van Ditmarsch, H., and Kooi, B. 2008. Semantic results for ontic and epistemic change. *Logic and the Foundations of Game and Decision Theory (LOFT 7)* 87–117.
- van Ditmarsch, H.; Ruan, J.; and Verbrugge, R. 2008. Sum and product in dynamic epistemic logic. *Journal of Logic and Computation* 18(4):563–588.
- van Ditmarsch, H.; van der Hoek, W.; and Kooi, B. 2007a. *Dynamic Epistemic Logic*. Springer.
- van Ditmarsch, H.; van der Hoek, W.; and Kooi, B. 2007b. Dynamic epistemic logic and knowledge puzzles. In *Conceptual Structures: Knowledge Architectures for Smart Applications*. Springer. 45–58.