Knowledge-Based Programs as Plans - The Complexity of Plan Verification - (Extended Abstract)

Jérôme Lang, Bruno Zanuttini

To cite this version:
Knowledge-Based Programs as Plans
– The Complexity of Plan Verification –
(Extended Abstract)

Jérôme Lang
LAMSADE, Université Paris-Dauphine & CNRS
lang@lamsade.dauphine.fr

Bruno Zanuttini
GREYC, Université de Caen Basse-Normandie,
CNRS UMR 6072, ENSICAEN
bruno.zanuttini@unicaen.fr

June 7, 2012

Abstract
Knowledge-based programs (KBPs) are high-level protocols describing the course of action an agent should perform as a function of its knowledge. The use of KBPs for expressing action policies in AI planning has been surprisingly underlooked. Given that to each KBP corresponds an equivalent plan and vice versa, KBPs are typically more succinct than standard plans, but imply more on-line computation time. Here we compare KBPs and standard plans according to succinctness and to the complexity of plan verification.

1 INTRODUCTION
Knowledge-based programs (KBPs) [4] are high-level protocols which describe the actions an agent should perform as a function of its knowledge, such as, typically, if $K \varphi$ then $\pi$ else $\pi'$, where $K$ is an epistemic modality and $\pi, \pi'$ are subprograms.

Thus, in a KBP, branching conditions are epistemically interpretable, and deduction tasks are involved at execution time (on-line). KBPs can be seen as a powerful language for expressing policies or plans, in the sense that epistemic branching conditions allow for exponentially more compact representations. In contrast, standard policies (as in POMDPs) or plans (as in contingent planning) either are sequential or branch on objective formulas (on environment and internal variables), and hence can be executed efficiently, but they can be exponentially larger (see for instance [1]).

Having said this, KBPs have surprisingly been underlooked in the perspective of planning. Initially developed for distributed computing, they have been considered in
AI for agent design [14] and game theory [7]. For planning, the only works we know of are by Reiter [13], who gives an implementation of KBPs in Golog; Classen and Lakemeyer [3] who implement KBPs in a decidable fragment of the situation calculus; Herzig et al. [6], who discuss KBPs for propositional planning problems, and Laverny and Lang [9, 10], who generalize KBPs to belief-based programs allowing for uncertain action effects and noisy observations.

None of these papers really addresses computational issues. Our aim is to contribute to filling this gap. After some background on epistemic logic (Section 2), we define KBPs (Section 3). Then we address expressivity and succinctness issues (Section 4): we show that, as expected, KBPs can be exponentially more compact than standard policies/plans. Then we give our main contributions, about the complexity of verifying that a KBP is a valid plan for a planning problem: we show $\Pi_2^p$-completeness for while-free KBPs (Section 5) and EXPSPACE-completeness in the general case (Section 6).

2 KNOWLEDGE

A KBP is executed by an agent in an environment. We model what the agent knows about the current state (of the environment and internal variables) in the propositional epistemic logic $S_5$. Let $X = \{x_1, \ldots, x_n\}$ be a set of propositional symbols. A state is a valuation of $X$. For instance, $\overline{x_1}x_2$ is the state where $x_1$ is false and $x_2$ is true. A knowledge state $M$ for $S_5$ is a nonempty set of states, representing those which the agent considers as possible: at any point in time, the agent has a knowledge state $M \subseteq 2^X$ and the current state is some $s \in M$. For instance, $M = \{x_1 \overline{x_2}, \overline{x_1} x_2\}$ means that the agent knows $x_1$ and $x_2$ to have different values in the current state.

Formulas of $S_5$ are built up from $X$, the usual connectives, and the knowledge modality $K$. An $S_5$ formula is objective if it does not contain any occurrence of $K$. Objective formulas are denoted by $\varphi, \psi$, etc. whereas general $S_5$ formulas are denoted by $\Phi, \Psi$ etc. For an objective formula $\varphi$, we denote by $\text{Mod}(\varphi)$ the set of all states which satisfy $\varphi$ (i.e., $\text{Mod}(\varphi) = \{s \in 2^X, s \models \varphi\}$). The size $|\Phi|$ of an $S_5$ formula $\Phi$ is the total number of occurrences of propositional symbols, connectives and modality $K$ in $\Phi$.

It is well-known (see, e.g., [4]) that any $S_5$ formula is equivalent to a formula without nested $K$ modalities; therefore we disallow them. An $S_5$ formula $\Phi$ is purely subjective if objective formulas occur only in the scope of $K$. In the whole paper we only need purely subjective formulas, because we are only interested in what the agent knows, not on the actual state of the environment. A purely subjective $S_5$ formula is in knowledge negative normal form (KNNF) if the negation symbol $\neg$ occurs only in objective formulas (in the scope of $K$) or directly before a $K$ modality. Any purely subjective $S_5$ formula $\Phi$ can be rewritten into an equivalently KNNF of polynomial size, by pushing all occurrences of $\neg$ that are out of the scope of $K$ as far as possible with de Morgan’s laws. For instance, $K\neg(p \land q) \lor \neg(Kr \lor K\neg r)$ is not in KNNF, but

\[ K\neg(p \land q) \lor \neg(Kr \lor K\neg r) \]
is equivalent to $K\neg(p \land q) \lor (\neg K r \land \neg K \neg r)$. Summarizing, a subjective $S_5$ formula $\Phi$ in KNNF (for short, $\Phi \in SKNNF$) is either a positive (resp. negative) epistemic atom $K \varphi$ (resp. $\neg K \varphi$), where $\varphi$ is objective, or a combination of such atoms using $\land, \lor$.

The satisfaction of a purely subjective formulas depends only on a knowledge state $M$, not on the actual current state (see, e.g., [4]):

- $M \models K \varphi$ if for all $s' \in M$, $s' \models \varphi$,
- $M \models \neg K \varphi$ if $M \not\models \varphi$,
- $M \models \Phi \land \Psi$ if $M \models \Phi$ and $M \models \Psi$,
- $M \models \Phi \lor \Psi$ if $M \models \Phi$ or $M \models \Psi$.

An $S_5$ formula is valid (resp. satisfiable) if it is satisfied by all (resp. at least one) knowledge states $M \subseteq 2^X$. Given two $S_5$ formulas $\Phi$ and $\Psi$, $\Phi$ entails $\Psi$, written $\Phi \models \Psi$, if every knowledge state $M \subseteq 2^X$ which satisfies $\Phi$ also satisfies $\Psi$, and $\Phi$ is equivalent to $\Psi$ if $\Phi$ and $\Psi$ entail each other. Note that $K \varphi \land K \psi$ is equivalent to $K(\varphi \land \psi)$, but that $K \varphi \lor K \psi$ is not equivalent to $K(\varphi \lor \psi)$: for instance, $K(\varphi \lor \neg \varphi)$ is valid whereas $K \varphi \lor K \neg \varphi$ is true only when the agent knows the value of $\varphi$. Also note that $K \varphi$ entails $\neg K \neg \varphi$, and that $M \models \neg K \varphi$ is weaker than $M \models K \neg \varphi$.

We use a syntactical representation of the current knowledge state $M \subseteq 2^X$ (at some timestep) of the agent executing a KBP. For this we observe that $M$ can be identified with any objective formula $\varphi$ which satisfies $M = Mods(\varphi)$, and we represent $M$ by the epistemic atom $O \varphi$. Intuitively, $O \varphi$ means “I know $\varphi$ and nothing else”. Without loss of generality, we disallow occurrences of $O$ or $K$ in the scope of $O$ or $K$. Formally, $O$ is an epistemic modality whose semantics is given in the logic of all I know [11] by

- $M \models O \varphi$ if $M = Mods(\varphi)$.

Hence any atom $\Phi$ of the form $O \varphi$ (with $\varphi$ satisfiable) has exactly one model, written $M(\Phi) = Mods(\varphi) \subseteq 2^X$. With this in mind, we use the term “knowledge state” to refer either to some $M \subseteq 2^X$ or to some atom $O \varphi$. For instance, $M = \{x_1 x_2, \bar{x}_1 \bar{x}_2, \bar{x}_1 x_2\}$ will also be denoted by $\Phi = O(\neg x_1 \lor x_2)$.

Satisfiability in the logic of all I know is $\Sigma^P_3$-complete [16]. However, we only need restricted entailment tests, the complexity of which we show to lie at the first level of the polynomial hierarchy only. We recall that $\Delta^P_2 = \text{NP}$ is the class of all decision problems that can be solved in deterministic polynomial time using NP-oracles, $\Pi^P_2 = \text{coNP}$ is the class of all decision problems whose complement can be solved in nondeterministic polynomial time using NP-oracles, and $\text{EXPSPACE}$ is the class of all decision problems that can be solved using exponential space.

**Proposition 1** Deciding $O \varphi \models \Phi$, where $\varphi$ is objective and $\Phi$ is purely epistemic (without any occurrence of $O$), is in $\Delta^P_2$.

**Proof** Let $\varphi$ and $\Phi$ be as in the claim. Hence $\Phi$ is a Boolean combination of atoms $K \psi$. We give a polynomial time algorithm for deciding $O \varphi \models \Phi$ with a linear number of calls to an oracle for propositional satisfiability, reasoning by induction on the structure of $\Phi$. 

3
First let $\Phi = K\psi$. Then $O\varphi \models \Phi$ reads $\forall M \subseteq 2^X, (M \models O\varphi \Rightarrow \forall s \in M, s \models \psi)$. Since $O\varphi$ has exactly one model $M = Mod(\varphi)$, this is equivalent to $\varphi \models \psi$, i.e., $\varphi \land \neg \psi$ is not satisfiable.

Now let $\Phi = \Phi_1 \lor \Phi_2$ ($\land$, $\neg$ are similar). Then $O\varphi$ entails $\Phi$ iff it entails $\Phi_1$ or it entails $\Phi_2$. Indeed, $O\varphi$ has only one model $M = Mod(\varphi)$, hence $O\varphi$ entails $\Phi$ iff $M(O\varphi)$ satisfies $\Phi_1$ or $\Phi_2$, that is, iff $O\varphi \models \Phi_1$ or $O\varphi \models \Phi_2$ holds. Hence, deciding $O\varphi \models \Phi$ involves a linear number of calls to the oracle by the induction hypothesis. $\square$

3 KBPS AS PLANS

Our definitions specialize those in [4] to our propositional framework and to a single-agent version. Given a set $A$ of primitive actions, a knowledge-based program (KBP) is defined inductively as follows:

- the empty plan is a KBP,
- any action $\alpha \in A$ is a KBP,
- if $\pi$ and $\pi'$ are KBPs, then $\pi; \pi'$ is a KBP;
- for KBPs $\pi, \pi'$ and $\Phi \in SKNNF$, if $\Phi$ then $\pi$ else $\pi'$ is a KBP;
- for a KBP $\pi$ and $\Phi \in SKNNF$, while $\Phi$ do $\pi$ is a KBP.

The class of while-free KBPs is obtained by omitting the while construct. The size $|\pi|$ of a KBP $\pi$ is defined to be the number of occurrences of actions, plus the size of branching conditions, in $\pi$.

3.1 Representation of Actions

Following [6], we assume without loss of generality that the set of actions is partitioned into purely ontic and purely epistemic actions.

An ontic action $\alpha$ modifies the current state of the environment but gives no feedback. Ontic actions may be nondeterministic. For the sake of simplicity we assume them to be fully executable\(^2\).

Each ontic action is represented by a propositional theory expressing constraints on the transitions between the states of the environment before and after $\alpha$ is taken. Let $X' = \{x' \mid x \in X\}$, denoting the values of variables after the action was taken. The theory of $\alpha$ is a propositional formula $\Sigma_\alpha$ over $X \cup X'$ such that for all states $s \in 2^X$, the set $\{s' \in 2^{X'} \mid ss' \models \Sigma_\alpha\}$ is nonempty, and is exactly the set of possible states after $\alpha$ is performed in $s$. For instance, with $X = \{x_1, x_2\}$, the action $\alpha$ which nondeterministically reinitializes the value of $x_1$ has the theory $\Sigma_\alpha = (x_2' \leftrightarrow x_2)$.

In the paper we will use the following actions:

- reinit($Y$) (for some $Y \subseteq X$) with theory $\bigwedge_{x \notin Y} x' \leftrightarrow x$,
- $x_i := \varphi$ (for $\varphi$ objective) with theory $x_i' \leftrightarrow \varphi \land \bigwedge_{j \neq i} x_j' \leftrightarrow x_j$.

\(^2\)This, in theory, induces a loss of generality, but in practice if $\alpha$ is not executable in $s$, this can be expressed by letting $\alpha$ lead to a sink state incompatible with the goal.
• switch($x_i$) with theory $x'_i \leftrightarrow \neg x_i \land \bigwedge_{j \neq i} x'_j \leftrightarrow x_j$,

• the void action $\lambda$ with theory $\bigwedge_{x \in X} x' \leftrightarrow x$.

Now, an epistemic action has no effect on the current state, but gives some feedback about it, that is, it modifies only the knowledge state of the agent (typically, a sensing action). We represent such an action by the list of possible feedbacks. Formally, the feedback theory of $\alpha$ is a list of positive epistemic atoms, of the form $\Omega_\alpha = (K\varphi_1, \ldots, K\varphi_n)$. For instance, the epistemic action which senses the value of an objective formula $\varphi$ is

• $\text{test}(\varphi)$ with feedback theory $\Omega_{\text{test}(\varphi)} = (K\varphi, K\neg\varphi)$.

Finally, for an objective formula $\varphi$ over $X$, we write $\varphi^t$ for the formula obtained from $\varphi$ by replacing each occurrence of $x \in X$ with $x^t$. We also write $\Sigma_\alpha^{t+1}$ for the formula obtained from $\Sigma_\alpha$ by replacing each unprimed variable $x \in X$ with $x^t$, and each primed variable $x'$ with $x'^{t+1}$. For instance, for $\varphi_1 = (x_1 \lor x_2)$, $\varphi^3_1$ is $(x_3^1 \lor x_2^1)$, and for $\Sigma_\alpha = (x_2^t \leftrightarrow x_2)$, $\Sigma_\alpha^{t+4}$ is $(x_2^t \leftrightarrow x_2^3)$.

### 3.2 Semantics

The agent executing a KBP starts in some knowledge state $M^0$, and at any timestep $t$ until the execution terminates, it has a current knowledge state $M^t$. When execution comes to a branching condition $\Phi$, $\Phi$ is evaluated in the current knowledge state (the agent decides $M^t \models \Phi$).

The knowledge state $M^t$ is defined inductively as the progression of $M^{t-1}$ by the action executed between $t - 1$ and $t$. Formally, given a knowledge state $M \subseteq 2^X$ and an ontic action $\alpha$, the progression of $M$ by $\alpha$ is defined to be the knowledge state $\text{Prog}(M, \alpha) = M' \subseteq 2^X$ defined by $M' = \{ s' \in 2^X \mid s \in M, ss' \models \Sigma_\alpha \}$. Intuitively, after taking $\alpha$ in a state which it knows to be one in $M$, the agent knows that the resulting state is one of those $s'$ which are reachable from any $s \in M$ through $\alpha$. Note that the agent knows that some outcome of the action has occurred (it knows $\Sigma_\alpha$), but not which one.

Now given an epistemic action $\alpha$, a knowledge state $M$, and a feedback $K\varphi_i \in \Omega_\alpha$ with $M \not\models K\neg\varphi_i$, the progression of $M$ by $K\varphi_i$ is defined to be $\text{Prog}(M, K\varphi_i) = M_i = \{ s \in M \mid s \models \varphi_i \}$. The progression is undefined when $M \models K\neg\varphi_i$. Intuitively, a state is considered to be possible after obtaining feedback $\varphi_i$ if it was considered to be possible before taking the epistemic action, and it is consistent with the feedback obtained. Here, observe that though an epistemic action can yield different feedbacks, at execution time the agent knows which one it gets.

**Example 1 (from [6]). Consider the following KBP**

$\pi$: $\text{test}(x_1 \leftrightarrow x_2)$; **If** $K(x_1 \leftrightarrow x_2)$ **then** $\text{test}(x_1 \land x_2)$ **else** $\text{switch}(x_1)$; $\text{test}(x_1 \land x_2)$

With $M^0 = O \top$ (nothing known), $\text{Prog}(M^0, K(\neg(x_1 \leftrightarrow x_2)))$ is $M^1 = O(x_1 \leftrightarrow \neg x_2)$, $\text{Prog}(M^1, \text{switch}(x_1))$ is $M^2 = O(x_1 \leftrightarrow x_2)$, and $\text{Prog}(M^2, K(\neg(x_1 \land x_2)))$ is $M^3 = O(\neg x_1 \land \neg x_2)$.

We are now ready to give an operational semantics for KBPs. Given a knowledge state $M'$ involving only primed variables of the form $x' (x \in X)$, we write $\text{plain}(M')$ for the knowledge state obtained by replacing $x'$ with $x$ for all $x \in X$. 

5
An execution trace (or trace) $\tau$ of a KBP $\pi$ in $M^0$ is a sequence of knowledge states, either infinite, i.e. $\tau = (M^i)_{i \geq 0}$, or finite, i.e. $\tau = (M^0, M^1, \ldots, M^T)$, and satisfying:

- if $\pi$ is the empty plan, then $\tau = (M^0)$;
- if $\pi$ is an ontic action $\alpha$, then $\tau = (M^0, \text{plain}(\text{Prog}(M^0, \alpha)))$;
- if $\pi$ is an epistemic action $\alpha$, then $\tau = (M^0, \text{Prog}(M^0, K\varphi_i))$ for some $K\varphi_i \in \Omega_\alpha$ with $M \not\models K\neg\varphi_i$;
- for $\pi = \pi_1; \pi_2$, either $\tau = \pi_1$ with $\pi_1$ an infinite trace of $\pi_1$, or $\tau = \pi_1\pi_2$ with $\pi_1$ a finite trace of $\pi_1$ and $\pi_2$ a trace of $\pi_2$;
- if $\pi$ is if $\Phi$ then $\pi_1$ else $\pi_2$, then either $M^0 \models \Phi$ and $\tau$ is a trace of $\pi_1$, or $M^0 \not\models \Phi$ and $\tau$ is a trace of $\pi_2$;
- if $\pi$ is while $\Phi$ do $\pi_1$, then either $M^0 \models \Phi$ and $\tau$ is a trace of $\pi_1; \pi$, or $M^0 \not\models \Phi$ and $\tau = (M^0)$.

We say that $\pi$ terminates in $M^0$ if every trace of $\pi$ in $M^0$ is finite.

**Example 2** Let $\pi, M^0, \ldots, M^3$ as in Ex. 1, and $M^4 = O(x_1 \wedge x_2)$. The traces of $\pi$ in $M^0$ (with the corresponding feedbacks) are:

<table>
<thead>
<tr>
<th>$(M^0, M^1, M^2, M^4)$</th>
<th>$K(\neg(x_1 \leftrightarrow x_2)), K(\neg(x_1 \wedge x_2))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(M^0, M^1, M^2, M^4)$</td>
<td>$K(\neg(x_1 \leftrightarrow x_2)), K(x_1 \wedge x_2)$</td>
</tr>
<tr>
<td>$(M^0, M^2, M^4)$</td>
<td>$K(x_1 \leftrightarrow x_2), K(\neg(x_1 \wedge x_2))$</td>
</tr>
<tr>
<td>$(M^0, M^2, M^4)$</td>
<td>$K(x_1 \leftrightarrow x_2), K(x_1 \wedge x_2)$</td>
</tr>
</tbody>
</table>

## 4 KBPS VS. STANDARD POLICIES

We now briefly compare KBPs with standard policies (or plans) with respect to succinctness and expressiveness\(^3\). As opposed to a KBP, define a standard policy to be a program with objective branching conditions. This encompasses plans for classical planning, which are simply sequences of actions $a_1; a_2; \ldots; a_n$, but also POMDP policies, which branch on observations, and other types of policies, such as controllers with finite memory [2].

Clearly, every KBP $\pi$ can be translated into an equivalent standard policy (a “protocol” in [4]), by simulating all possible executions of $\pi$ and, for all possible executions of the program, evaluating all (epistemic) branching conditions. *Vice versa*, it is clear that any standard policy can be translated to an equivalent KBP.

Such translations are of course not guaranteed to be polynomial. In particular, it is easily seen that a standard policy $\pi$ described in space $O(n)$ can manipulate at most $n$ variables (through actions or branching conditions). It follows that it can be in at most $|\pi|^2^n$ different configurations (value of each variable plus control point in the policy), hence any finite trace of such a policy can have length at most $|\pi|^2^n$ (being twice in the

\(^3\)For space reasons, our discussion is informal. Proofs and details are omitted.
same configuration would imply an infinite loop). In contrast, we will give in Section 6 a KBP described in space polynomial in \( n \) but with a finite trace of length \( 2^n \).

However, what is gained on succinctness is lost on the complexity of execution. When executing a KBP, the problem of evaluating a branching condition is in \( \Delta^P_2 \), but is both \( \text{NP} \)- and \( \text{coNP} \)-hard. Indeed, it is \( \text{coNP} \)-hard because \( O^+ \models K\varphi \) corresponds to \( \varphi \) being valid, and \( \text{NP} \)-hard because \( O^+ \models \neg K\neg \varphi \) corresponds to \( \varphi \) being satisfiable. On the other hand, when executing a standard policy, evaluating an (objective) condition can be done in linear time by reading the values of the (internal and environment) variables involved.

Interestingly, even the restriction to while-free KBPs does not imply a loss of expressivity. Indeed, if a loop terminates, then it is guaranteed to be executed less than \( 2^n \) times (see Section 6), and hence it can be unrolled, yielding an equivalent while-free KBP. However, this obviously comes also with a loss of succinctness.

When KBPs are seen as plans which achieve goals, as we consider in this article, the translations outlined above preserve the property that the KBP/policy indeed achieves the goal. Therefore, from the point of view of plan existence, considering KBPs or standard plans makes no difference: there is a plan for a given problem if and only if there is a KBP for it. Moreover, since the input is the same in both cases, the complexity of plan existence is independent of whether we look for policies of KBPs. Things are different for the problem of verifying that a KBP/policy is a plan for some goal, because the KBP or policy is part of the input. For example, we will see in Section 6 that verifying while-free KBPs is \( \Pi^p_2 \)-complete. In contrast, it can easily be shown that verifying a while-free policy is in \( \text{coNP} \) (with essentially the same proof as Proposition 2).

5 VERIFYING WHILE-FREE KBPS

We now investigate the computational problem of verifying that a KBP \( \pi \) is valid for a planning problem. Precisely, we define a knowledge-base planning problem \( P \) to be a tuple \( (\Phi^0, A_O, A_E, G) \), where \( \Phi^0 = O^0 \) is the initial knowledge state, \( G \) is a SKNNF \( S_5 \) formula called the goal, and \( A_O \) (resp. \( A_E \)) is a set of ontic (resp. epistemic) actions together with their theories. Then a KBP \( \pi \) (using actions in \( A_O \cup A_E \)) is said to be a (valid) plan for \( P \) if its execution in \( \Phi^0 \) terminates, and for all traces \( (M^0, \ldots, M^T) \) of \( \pi \) with \( M^0 = M(\Phi^0) \), \( M^T \models G \) holds. Intuitively, this means that executing \( \pi \) always leads to a knowledge state where the agent is sure that \( G \) holds. For instance, in Example 1, \( \pi \) is a plan for \( \Phi^0 = O^+ \) and the goal \( G = (Kx_1 \lor K\neg x_1) \land (Kx_2 \lor K\neg x_2) \).

**Definition 1 (verifying)** The plan verification problem takes as input a knowledge-based planning problem \( P = (\Phi^0, A_O, A_E, G) \) and a KBP \( \pi \), and asks whether \( \pi \) is a plan for \( P \).

In this section we show that verification is \( \Pi^p_2 \)-complete for while-free KBPs, even under several further restrictions. Observe first that a while-free KBP always terminates.

We start with membership in \( \Pi^p_2 \). In the broad lines, the argument is that \( \pi \) is *not* a plan for \( P \) if there exists a trace \( \tau \) of \( \pi \) (or, equivalently, a sequence of feedbacks
for the epistemic actions executed) in which the last knowledge state does not satisfy $G$. Hence $\pi$ can be verified not to be a plan for $P$ by guessing such a sequence of feedbacks and simulating the corresponding execution.

Nevertheless, we must perform such simulation in polynomial space. Unfortunately, in general the progression of a knowledge state $M$ represented as $O\varphi$ cannot be performed in polynomial space.

**Example 3** The progression of $O\varphi$, for $\varphi = \bigwedge_{i=1}^{n}(x_i \lor y_i \rightarrow z_i) \land (\bigwedge_{i=1}^{n} z_i \rightarrow z)$, by reinit($\{z_1, \ldots, z_n\}$), is equivalent to $O(\exists z_1, \ldots, z_n, \varphi) \equiv \bigwedge_{i_1} \in \{x_1, y_1\}, \ldots, \ell_n \in \{x_n, y_n\} (\ell_1 \land \cdots \land \ell_n \rightarrow z)$, which has no polynomial representation, while $|\varphi|$ is linear.

Hence we introduce another form of progression, called memoryful progression, which explicitly keeps track of successive knowledge states instead of projecting to the current instant. Namely, we define a memoryful knowledge state for a timestep $t$ to be a formula of the form $O\varphi^t$, where $\varphi^t$ is an objective formula over the set of variables $\bigcup_{t=0}^{\infty} \{x^t \mid x \in X\}$. Intuitively, $O\varphi^t$ represents the past and present knowledge of the agent at timestep $t$. Formally:

- for ontic $\alpha$, MemProg($O\varphi^t, \alpha$) = $O(\varphi^t \land \Sigma^t\alpha^{t+1})$,
- for a feedback $K\varphi_i$, MemProg($O\varphi^t, K\varphi_i$) = $O(\varphi^t \land \varphi_i^t)$.

Observe that ontic actions increment the current timestep, while epistemic actions do not (they do not modify the current state).

**Example 4 (Example 1, continued)** The memoryful progression of $O\top$ by $K(\neg(x_1 \leftrightarrow x_2)$, then switch($x_1$), then $K(\neg(x_1 \land x_2)$ is

$$O\left(\top \land (x_1^0 \leftrightarrow \neg x_2^0) \land (x_1^1 \leftrightarrow \neg x_1^0) \land (x_2^1 \leftrightarrow x_2^0) \land \neg (x_1^1 \land x_2^1)\right)$$

Clearly, the memoryful progression of $O\varphi^t$ by $\alpha$ (resp. $K\varphi_i$) has a size linear in $|\varphi^t|$ and $|\Sigma_\alpha|$ (resp. $|K\varphi_i|$). Hence iterating the memoryful progression of $\Phi^0$ a polynomial number of times poly$(|\pi|)$ yields a (memoryful) knowledge state of polynomial size.

**Lemma 1** Let $O\varphi^0$ be an initial knowledge state, and let $(\ell_1, \ldots, \ell_T)$ be a sequence containing $U$ ontic actions and $T - U$ feedbacks. Then the iterated progression $M$ of $M^0$ by $\ell_1, \ldots, \ell_T$ satisfies an epistemic formula $\Phi$ if and only if the iterated memoryful progression $O\varphi_t^U$ of $O\varphi^0$ by $\ell_1, \ldots, \ell_T$ entails $\Phi^U$.

**Proof Sketch** Renaming variables, $M \models \Phi$ is equivalent to $M^U \models \Phi^U$. On the other hand, it is easily seen from the definitions that $M^U$ is exactly Mod($\exists X^0, \ldots, \exists X^{U-1} \varphi^t_U$).

Because $\Phi^U$ contains only variables in $X^U$, it follows that $M^U \models \Phi^U$ is equivalent to $\text{Mod}(\varphi^U) \models \Phi^U$ ([8], Corollary 7), i.e., to $O\varphi_t^U \models \Phi^U$.

Recall that a problem is in $\Pi^0_2$ if its complement can be solved by a polytime non-deterministic algorithm which uses an NP-oracle.

**Proposition 2** Plan verification is in $\Pi^0_2$ for while-free KBPs.
Algorithm 1: Deciding whether a while-free $\pi$ is not valid

begin
  $t := 0$, $O^{\top}_{\varphi} = \Phi^0$;
  while $\pi$ is not the empty KBP do
    if $\pi = \alpha; \pi'$ and $\alpha$ is ontic then
      $O^{t+1}_{\varphi} := $ MemProg($O^t_{\varphi}, \alpha$);
      $\pi := \pi'$, $t := t + 1$;
    else if $\pi = \alpha; \pi'$ and $\alpha$ is epistemic then
      guess a feedback $K\varphi_i$ in $\Omega_\alpha$;
      check that $\varphi^t \land \varphi_i$ is satisfiable;
      $O^{t+1}_{\varphi} := $ MemProg($O^t_{\varphi}, K\varphi_i$);
    else if $\pi$ is of the form if $\Phi$ then $\pi_1$ else $\pi_2$ then
      if $O^t_{\varphi} \models \varphi_i$ then $\pi := \pi_1; \pi'$;
    else $\pi := \pi_2; \pi'$;
  end
  check $O^{t+1}_{\varphi} \not\models G^t$;
end

Proof We use Algorithm 1, which decides whether $\pi$ is not valid in nondeterministic polynomial time with access to an oracle for propositional satisfiability. Intuitively, the algorithm simulates an execution of $\pi$, guessing a sequence of feedbacks witnessing that $\pi$ is not valid. Clearly, it runs in nondeterministic polynomial time, and it uses a polynomial number of calls to the oracle: one per check that $\varphi^t \land \varphi_i$ is satisfiable, a linear number per check $O^{t+1}_{\varphi} \models \Phi^t$ (Proposition 1), and a linear number for the final check.

Proposition 3 Plan verification is $\Pi^P_2$-hard. Hardness holds even if the KBPs $\pi$ are restricted to be while-free and either to $A_O = \emptyset$ (no ontic action), or to $A_E = \emptyset$ (no epistemic action).

Proof We give two reductions from the ($\Pi^P_2$-complete) problem of deciding the validity of a QBF of the form $\forall x_1, \ldots, x_p \exists y_1, \ldots, y_q \varphi$, where $\varphi$ is a propositional formula over $\{x_1, \ldots, x_p\} \cup \{y_1, \ldots, y_q\}$. In both cases we build a planning problem $P = (\Phi^0, A_O, A_E, G)$ and a KBP $\pi$, with $\Phi^0 = \top$ and $G = \neg K \neg \varphi$.

Given only epistemic actions, we let $\pi = test(x_1); \ldots; test(x_p)$. Then $\pi$ is not a valid plan if and only if there is a sequence of feedbacks for $test(x_1), \ldots, test(x_p)$ such that $K \neg \varphi$ holds, i.e., the agent knows that $\varphi$ is false. This is equivalent to there being values for $x_1, \ldots, x_p$ such that whatever the value of $y_1, \ldots, y_q$, $\varphi$ is false, that is, to $\forall x_1, \ldots, x_p \exists y_1, \ldots, y_q \varphi$ not being valid.

Similarly, given only ontic actions we let $\pi = reinit(\{x_1, \ldots, x_p\})$. Again, $\pi$ is not a valid plan if and only if there is an outcome for $\alpha$, that is, values for $x_1, \ldots, x_p$, such that $K \neg \varphi$ holds, i.e., $\forall x_1, \ldots, x_p \exists y_1, \ldots, y_q \varphi$ is not valid.

\end{document}
6 VERIFYING KBPS WITH LOOPS

For general KBPs, we now show verification to be EXPSPACE-complete (EXPSPACE is the class of decision problems with an exponential space algorithm). On the way, we build a polysize KBP with a doubly exponentially long trace, which we use as a clock. Since the construction is of independent interest, we present it first.

6.1 A Very Slow KBP

We write $>$ for the lexicographic order on states. For instance, $2^X$ is ordered by $x_1x_2x_3 > x_1x_2x_3 > x_1x_2x_3 > \cdots > \bar{x}_1\bar{x}_2\bar{x}_3$ for $n = 3$ variables. Given $X$ and a knowledge state $M$ over a superset of $X$, we write $M_X$ for $\{s_X \mid s \in M\}$, where $s_X$ denotes the restriction of $s$ to the variables in $X$. This allows us to use auxiliary variables and still talk about the knowledge state about $X$.

We build a compact KBP (of size polynomial in $n$) with exactly one trace, of size $2^{2n} - 1$. As discussed in Section 4, this is impossible with standard policies, but possible for KBPs because their configurations include a knowledge state, and there are $2^{2n} - 1$ of them (every nonempty subset of $2^X$). Hence there can be a program $\pi$ which passes through $2^{2n}$ different configurations while being specified with only $O(n)$ variables and in space $|\pi|$ polynomial in $n$.

Routines and Actions We build our KBP so that its execution passes through each possible knowledge state exactly once. To do so, we need some specific actions and routines which allow to go from a knowledge state to the next one.

The first routine determines the state $s$ in $M$ with the greatest restriction $M_X$ (wrt $>$), and stores it over some auxiliary variables $g_1, \ldots, g_n$. For instance, if the current knowledge state satisfies $M_X = \{\bar{x}_1\bar{x}_2x_3, \bar{x}_1x_2x_3, x_1\bar{x}_2\bar{x}_3\}$, then after executing the routine, the agent knows $g_1 \land \neg g_2 \land \neg g_3$ (and $M_X$ is unchanged).

We define $\pi_g$ to perform a dichotomic search in $M$. For instance, if $K((x_1 \leftrightarrow g_1) \rightarrow \neg x_2)$ is true, then no assignment in $M$ which satisfies $x_1 \leftrightarrow g_1$ (i.e., by construction, none of the assignments with greatest $x_1$) satisfies $x_2$, hence the greatest one satisfies $\neg x_2$. Precisely, $\pi_g$ is the following KBP:

\[
\begin{align*}
\text{If } K(\neg x_1) & \text{ then } g_1 := 0 \text{ else } g_1 := 1; \\
\text{If } K((x_1 \leftrightarrow g_1) \rightarrow \neg x_2) & \text{ then } g_2 := 0 \text{ else } g_2 := 1; \\
\vdots \\
\text{If } K(\bigwedge_{i=1}^{n-1} x_i \leftrightarrow g_i) & \rightarrow \neg x_n) \text{ then } g_n := 0 \text{ else } g_n := 1;
\end{align*}
\]

We now introduce an ontic action which adds a given state $s_a \in 2^X$ to $M_X$. We assume $s_a$ is encoded over some auxiliary variables $a_1, \ldots, a_n$. The action $a_{\text{add}}^n$ is a simple nondeterministic one, which either does nothing or sets $x_1, \ldots, x_n$ to the values of $a_1, \ldots, a_n$. Formally, its action theory is $(\bigwedge_{i=1}^n x_i' \leftrightarrow x_i) \lor (\bigwedge_{i=1}^n x_i' \leftrightarrow a_i)$ (and no effect on auxiliary variables). Hence after taking this action, the agent exactly knows that either the environment is in the same state as before, or it is in the state $a_1 \ldots a_n$.\[\]
We finally introduce a routine $\pi_n$, which removes a given state $s_r \in 2^X$ (encoded over auxiliary variables $r_1, \ldots, r_n$) from $M_X$. We assume that the agent knows the state to be removed, that is, $M$ satisfies $K r_i \lor K \neg r_i$ for all $i = 1, \ldots, n$.

Recall that by definition, knowledge states are nonempty. Hence we allow removal of $s_r$ only if there is another state $s_g \in M_X$, ensuring $M_X \setminus \{s_r\} \neq \emptyset$. Then $\pi_n$ removes $s_r$ from $M_X$ by identifying a distinguished state $s_g \neq s_r$ in $M_X$, then executing an action $a^n_g$ which maps any state to itself except for $s_r$, which it maps to $s_g$.

We identify $s_g$ by running $\pi_n$. If it turns out that $s_g$ is precisely $s_r$, as can be decided since the agent knows (i) the value of $s_r$ by assumption, and (ii) that of $s_g$ by construction of $\pi_n$, then $\pi_n$ replaces $s_g$ with the least assignment in $M_X$, using the dual of $\pi_n$. Finally, $a^n_g$ is defined to be the deterministic ontic action with theory $\bigwedge_{i=1}^n x'_i \leftrightarrow (x_i \oplus ((\lor_{i=1}^n x_i) \leftrightarrow (x_i \oplus g_i)))$ (and no effect on auxiliary variables). A case analysis shows that $a^n_g$ maps $s$ to itself except for $s_r$, which it maps to $s_g$, as desired.

Proposition 4 Let $\pi_n^g$ run in knowledge state $M$. Then the progressed knowledge state $M'$ satisfies $\bigwedge_{i=1}^n \neg \neg g_i$, where $g_i$ is $g_1$ (resp. $\neg g_i$) if the greatest state in $M_X$ satisfies $x_i$ (resp. $\neg x_i$). Now if $a^n_{\text{add}}$ (resp. $\pi^n_g$) is run, then $M'_X = M_X \cup \{s_a\}$ (resp. $M'_X = M_X \setminus \{s_g\}$) holds.

Importantly, observe that $\pi^n_g, a^n_{\text{add}}, \pi^n_g$ all have a description of size at most quadratic in $n$. Finally, we use a routine, written $\pi_d$ ("decrement"), which replaces the state encoded by $g_1, \ldots, g_n$ by its predecessor wrt $>$ (the definition of $\pi_d$ is straightforward, and omitted for space reasons).

A Slow KBP We see a knowledge state $M$ as a vector $\vec{m} = m_1 m_2 \ldots m_{2^n-1} m_{2^n}$, with $m_i = 1$ if and only if the $i$th state $s_i$ (wrt $>$) is in $M$. Then our KBP starts with $\vec{m}^0 = 00 \ldots 01$ (i.e., $M^0 = \{1 \ldots 1\} \equiv K \neg x_1 \land \cdots \land \neg x_n$), and loops until $\vec{m}^t = 10 \ldots 00 (M^t = \{00 \ldots 0\} \equiv K \neg x_1 \land \cdots \land \neg x_n)$. The loop changes the current $\vec{m}^t$ to $\vec{m}^{t+1}$ using the Gray code, which is a way to enumerate all Boolean vectors by changing exactly one bit at a time.

Definition 2 (Gray Code) The successor of $\vec{m}$ according to the Gray Code is obtained from $\vec{m}$ as follows:

1. if $\vec{m}$ has an even number of 1's, flip $m_{2^n}$.
2. otherwise, let $g = \max\{i \mid m_i = 1\}$ and flip $m_{g-1}$.

For instance, the enumeration is 0001, 0011, 0010, 0110, 0111, 0101, 0100, 1100, 1101, 1111, 1110, 1010, 1011, 1001, 1000, 0000, 0001, 0011, 0010, 0110, 0111, 0101, 0100, 1100, 1101, 1111, 1110, 1010, 1011, 1001, 1000. In terms of knowledge states, this is $\{x_1 x_2\}, \{x_1 \bar{x}_2, x_1 x_2\}, \{x_1 \bar{x}_2\}, \{\bar{x}_1 \bar{x}_2\} \ldots \{\bar{x}_1 x_2\}$, which indeed passes through all knowledge states.

By definition of $\vec{m}^t$, the greatest $i$ with $m_i^t = 1$ identifies the greatest state in $M^t$, and flipping $m_i^t$ amounts to add/remove $s_i$ to $M^t$. With this in hand, our KBP clock

Proposition 5 The unique trace for clock in $M^0$ has size $2^{2^n} - 1$. 

11
Algorithm 2: The KBP \( \text{clock}^n \)

begin

  \( \text{odd} \leftarrow 1 \) \{number of 1’s in knowledge state \( M^0 \)\};

  while \( \neg K(\neg x_1 \land \cdots \land \neg x_n) \) do

  \( \text{if } K \neg o \) then \{even number of 1’s, flip \( m_{2n} \)\}

  \( \text{if } K(\neg x_1 \lor \cdots \lor \neg x_n) \) then

  \( \{11 \ldots 1 \notin M^t, \text{add it} \} a_1 := 1; a_2 := 1; \ldots; a_n := 1; a_{\text{add}}; \)

  \( \text{else } r_1 := 1; r_2 := 1; \ldots; r_n := 1; \pi^n \)

  \( \text{else} \) \{odd number of 1’s, flip \( m_{g-1} \)\}

  \( \pi_g; \pi_d; \)

  \( \text{if } K(x_1 \not\leftrightarrow g_1 \lor \cdots \lor x_n \not\leftrightarrow g_n) \) then

  \( \{s_{g-1} \notin M^t, \text{add it} \} a_1 := g_1; a_2 := g_2; \ldots; a_n := g_n; a_{\text{add}}; \)

  \( \text{else } r_1 := g_1; r_2 := g_2; \ldots; r_n := g_n; \pi^n; \)

  \( \text{odd} := \neg \text{odd}; \)

end

6.2 EXSPACE-hardness

We now show that verifying general KBPs is \text{EXSPACE}-complete. We prove harness with a reduction from nondeterministic unobservable planning (NUP) [5]. An instance of NUP is a tuple \((\phi^0, A_{HJ}, \phi^G)\) where \(\phi^0, \phi^G\) are propositional formulas and \(A_{HJ}\) is a set of ontic, nondeterministic actions (see below). The question is whether there is a plan, i.e., a sequence of actions, which reaches a state satisfying \(\phi^G\) from any state satisfying \(\phi^0\) (in our terms, whose traces in \(O_{\phi^0}\) all end in a knowledge state satisfying \(K\phi^G\)).

The actions considered by Haslum and Jonsson (\textit{HJ-actions} for short) are different from ours. They are defined inductively as follows (we adapt their notation for consistency):

- \(x_i := 0\) and \(x_i := 1\) are HJ-actions for any \(x_i \in X\),
- if \(a_1, a_2\) are HJ-actions, then \(a_1; a_2\) is an HJ-action,
- if \(\varphi\) is a propositional formula and \(a_1, a_2\) are HJ-actions, then \(\text{if } \varphi \text{ then } a_1 \text{ else } a_2\) is an HJ-action\(^4\),
- if \(a_1, a_2\) are HJ-actions, then \(a_1|a_2\) is an HJ-action.

The semantics of executing such an action is the same as ours for the three first constructs, given that \(\varphi^0\) defines the initial knowledge state \(O\varphi^0\). As for nondeterminism, the progression of \(M^t\) by \(\neg |a_2\) is simply defined to be \(\text{Prog}(M^t, a_1) \cup \text{Prog}(M^t, a_2)\) (at execution time exactly one of \(a_1, a_2\) occurs, but we do not know which one).

The idea of our reduction is to build a KBP, written \textit{simulate}, which explores all possible plans (up to size \(2^n\), see below) for an NUP problem, and which is valid if and

\(^4\)In [5] this operator is \(n\)-ary, but as the conditions are mutually inconsistent, their \(\varphi_1 \land a_1; \ldots; \varphi_k \land a_k\) can be rewritten as if \(\varphi_1 \text{ then } a_1 \text{ else } (\varphi_2 \text{ then } a_2 \text{ else } \ldots)\), which has the same size.
only if none achieves the goal. For this, we first associate a routine \((\text{KBP})\) \(\pi(a)\) to any HJ-action \(a\), so as to be able to use \(a\) in \(\text{simulate}\). Indeed, conditional HJ-actions are not allowed in KBPs because they branch on objective formulas, and nondeterministic HJ-actions are not directly allowed.

For \(a\) of the form \(\textbf{if } \varphi \textbf{ then } a_1 \textbf{ else } a_2\), we define \(\pi(a)\) by “pushing in” the objective test to the assignments. Precisely, we define \(\pi(a)\) to be the \(c := \varphi; \pi_c(a_1); \pi_{\neg \varphi}(a_2)\) with \(\pi_c\) defined inductively by:

- \(\pi_c(x := \psi) = (x := (x \oplus (c \land (x \oplus \psi))))\),
- \(\pi_c(a_1; a_2) = (\pi_c(a_1); \pi_c(a_2))\),
- \(\pi_c(\textbf{if } \varphi \textbf{ then } a_1 \textbf{ else } a_2) = (c' := c \land \varphi; \pi_{c'}(a_1); \pi_{\neg \varphi}(a_2))\),
- \(\pi_c(a_1|a_2) = (\pi_c(a_1)|\pi_c(a_2))\).

Intuitively, executing \(a\) or \(\pi_1(a)\) in \(M^t\) leads to the same knowledge state \(M^{t+1}\), while \(\pi_0(a)\) leaves \(M^t\) unchanged (the construction of \(\pi_c(x := \psi)\) is the same as for the action \(a^n\) in Section 6.1). Interestingly, this construction shows that the restriction to purely subjective branching conditions in KBPs is without loss of generality.

Finally, for nondeterminism we use the action \(\text{reinit}(h)\), where \(h\) is an auxiliary variable, for simulating a coin flip (\(h\) stands for “heads”), and we define \(\pi(a_1|a_2)\) to be \((\text{reinit}(h); \pi_h(a_1); \pi_{\neg h}(a_2))\).

**Lemma 2** For any HJ-action \(a\), the KBP \(\pi(a)\) can be built efficiently, and for any \(M\), the progressions \(M_{HJ}^t\) and \(M^t\) of \(M\) by a (resp. \(\pi(a)\)) satisfy \((M_{HJ}^t)_X = M^t_X\) (ignoring auxiliary variables).

**Proposition 6** The verification problem for KBPs is \(\text{EXSPACE}\)-hard. Hardness holds even if only one while-loop is allowed and the KBPs to be verified are known to terminate.

**Proof** We use both results that deciding whether an NUP instance has a plan is \(\text{EXPSpace}\)-complete, and that an instance has a plan if and only if it has one of size at most \(2^2^n\) [5].

Given an instance \((\varphi^0, A_{HJ}, \varphi_G)\) of NUP, we build the knowledge-based planning problem \((O\varphi^0, A_O, A_E, G, A_G)\) and the KBP \(\text{simulate}\). This KBP uses the set of \(n\) variables \(X\) of the NUP instance, together with auxiliary sets of variables of size \(n\) for use by \(\text{clock}^n\) (using disjoint sets of variables makes \(\text{clock}^n\) run in parallel of the simulation itself). Then it loops over a guess of an action in \(A_{HJ} = \{a_1, \ldots, a_k\}\); this is done by flipping \(k\) coins (using \(\text{reinit}([h_1, \ldots, h_k])\)) and executing the first action whose coin turned heads (determined by taking the epistemic actions \(\text{test}(h_i)\))\(^5\).

The KBP \(\text{simulate}\) is depicted on Figure 3. We let \(\text{clock}^n+\) be a KB which counts up to \(2^n\) (obtained, say, by adding a dummy action to \(\text{clock}^n\)). Clearly, \(\text{simulate}\) can be built in polynomial time. Let \(p\) be a plan of length at most \(2^2^n\) for the NUP instance. Then by definition, the trace of \(\text{simulate}\) in which precisely the actions in \(p\) are chosen by \(\text{reinit}([h_1, \ldots, h_k])\) ends up with \(K\varphi_G\) being true, i.e., the goal \(\neg K\varphi_G\)

\(^5\) Clearly, \(\log k\) coins would be enough, but we keep the presentation simple.
Algorithm 3: The KBP simulate

begin
  initialize clock\(^{n+}\);
  {Loop until NUP goal reached for sure or clock beeps};
  while \(\neg K\varphi_G \land \neg K(\neg x_1 \land \cdots \land \neg x_n)\) do
    run one step of clock\(^{n+}\);
    reinit(\(\{h_1, \ldots, h_k\}\)); test(\(h_1\)); \ldots; test(\(h_k\));
    if \(Kh_1\) then \(\pi(a_1)\);
    else if \(Kh_2\) then \(\pi(a_2)\);
      \ldots;
    else if \(Kh_k\) then \(\pi(a_k)\);
end

being false. Hence simulate is not valid. Conversely, only a choice of actions which
achieve \(K\varphi_G\) can witness that simulate is not valid, hence if simulate is not valid
then there is a plan for the NUP instance. Hence NUP reduces to the complement of
KBP verification, hence the latter is coEXPSPACE-hard, that is, EXPSPACE-hard.
\(\square\)

Proposition 7 The verification problem for KBPs is in EXPSPACE.

Proof The proof mimicks Proposition 2 and Algorithm 1. Because a while loop
being executed more than \(2^n\) times would necessarily start at least twice in the same
knowledge state and hence run forever, such loops are unrolled \(2^n\) times. These can be
counted over \(2^n\) bits, hence in exponential space. As for the current knowledge state,
instead of using memoryful progression, we maintain \(M^t\) in extension (as its explicit
list of states), again in exponential space. Hence the problem is in coNEXPSPACE,
hence in coEXPSPACE=EXPSPACE by Savitch’s theorem. \(\square\)

7 CONCLUSION

We investigated the use of knowledge-based programs as compact representations of
policies or plans. It turns out that they are an interesting representation, dual to stan-
dard once in the sense that they can be exponentially more compact, but execution is
computationally more difficult, as well as verification.

This work is a first step towards bridging knowledge-based programming and AI
planning. Plan generation (that is, synthesis of KBPs) is the next issue to be addressed.

Acknowledgements

This research has been supported by the ANR-10-BLAN-0215 project Learning And
Reasoning for Deciding Optimally using Numerical and Symbolic information (LARDONS).
References


