

# ACTION, ABDUCTION AND PLAN RECOGNITION

**Aldo Franco Dragoni**  
Istituto di Informatica, Università di Ancona  
via Breccie Bianche, 60131, Ancona (Italy)  
e-mail dragon@inform.unian.it

**Abstract.** In the forthcoming distributed autonomous robotic systems it will be robot to recognize other robots' goals and plans from visual information. The emphasis has been given to plan inference. This paper is about *goal recognition*: when an observer has recognised a plan (may be after the entire plan has been performed), try to recognize the actor's reasons for the plan to be performed. If the actor's planner possesses inferential capabilities, then goal recognition is not a trivial question. This paper is based on simple hypotheses on the nature of the planner that guides an actor's actions. An observer can recognize the actor's goal by means of a simple *clause-based abductive reasoning*. Furthermore, the paper shows how goal recognition can be regarded as a useful tool for plan inference. This results refer to the prototypical state-based STRIPS planner.

## 1. INTRODUCTION

Many people in the Artificial Intelligence community have shown the importance of *goal recognition* as inferring the other agents' plans from their partially performed or planned portions [Carberry 90] [Kautz 90] [Charniak 93]; see [Carberry 93] for a good overview of the subject. In Distributed Artificial Intelligence [Bond 88] it is widely accepted that cooperative interaction depends upon agents reasoning about one another's goals. A plan is a sequence of actions that can be expected to allow an agent to reach a goal. There is no bijective correspondence between plans and goals because, in general, a plan can be performed to achieve different goals and the same goal can be achieved by different plans (Fig 1). Furthermore, the same sequence of actions can be regarded as different plans if performed in different situations; for instance, if you are in a big department store, the same action "go upstairs" can represent a "plan to go to the bank" or a "plan to go to a restaurant" or whatever else plan depending on the floor you are on when you start.

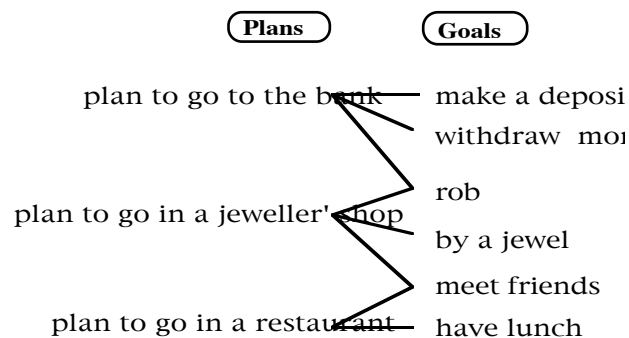


Fig 1. There is no bijective correspondence between plans and goals

This paper is about *goal recognition*: having recognised a plan (may be after the end of the plan has been performed), try to recognise which were the reasons for the plan to be planned. If the planner possesses sufficient inferential abilities, then the plan's goal could be recognised. This is not adding and/or removing the facts explicitly listed in the operators' definitions, but rather, it is some state of affairs that will be implied by these changes; in other words, the goal is a "logical consequence(s)" of the changes made in the world by the plan, particularly the last one (Fig. 2). If this is the case, then goal recognition is not a trivial question and this paper shows that it can be considered an abductive problem. This paper begins defining a class of logic-based abduction problem that fits particular goal recognition task. After a brief sentimental overview on the venerable STRIPS system, we prove the abductive nature of the goal recognition task for a STRIPS-like system. We refer to that old planner as the simplicity of its formal characterisation [Nilsson 81] as means to modify the state of the world by adding and/or removing facts. Its representativeness over a large class of state-based planners. We think that STRIPS plan inference and goal recognition play the same foundational and archetypal role

playing in plan synthesis during the last two decades. Finally, we show how gc can be regarded as a step in plan recognition since the goal of a partially performed plan is to make the rest of the plan performable. This goal centered view of plan synthesis stands side by side with the classical approaches and it's a (small) step towards an ambitious task of "mental state" recognition from visual information or communication acts") [Dragoni 94].

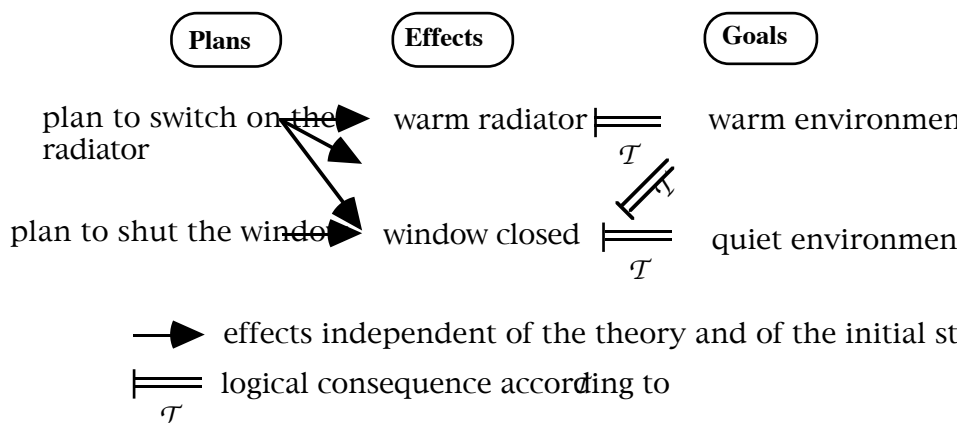


Fig 2. The plan's goal(s) can be some "logical consequence(s)" of the changes produced by

## 2. ABDUCTION

Abduction is generally presented as an abstract hypothetical inferential scheme. It involves a causal theory of the world (a set of formal rules or links between causes and effects), a set of observations (facts that don't follow simply from the causal theory), and a *primitive explanation* of the observed facts. An explanation is a set of hypothetical facts which, with the causal theory, justifies the presence of the observed facts. Recently there have been various formal characterisations of abduction (see [Paul 93] for a complete overview of the subject). The following is a logic-based account for abduction.

Let  $\mathcal{L}$  be a first order language and  $\mathcal{T}$  be a logical theory defined over the language  $\mathcal{L}$  and  $\Omega$  be sets of sentences of  $\mathcal{L}$  respectively called *abducible* and *observable*. A *logic-based abduction problem* (hereafter *logic-based a.p.*) is that of finding a *primitive explanation*  $\varphi \in A$  for an observation  $\omega \in \Omega$  such that:

- $\mathcal{T} \cup \varphi \vdash \omega$  (1)
- $\mathcal{T} \cup \varphi$  is consistent. (2)

Thus, the observation  $\omega$  must be derivable from the logical theory  $\mathcal{T}$  augmented with the explanation  $\varphi$  under the additional condition that  $\varphi$  is consistent with  $\mathcal{T}$ . Any primitive explanation for  $\omega$  still verifies the 1 and 2 even if it doesn't belong to  $A$ . This suggests the following definition.

**Definition 1.** If a disjunction of primitive explanations doesn't belong to the ab then it is said to be a *cautious* explanation for  $\omega$ . If there are a finite number explanations for  $\omega$  then their disjunction is said *most cautious explanation*.

In this paper I'll refer to the following slight general definition of abduction.

**Definition 2.** Let  $\mathcal{L}$  be a first order language and  $\mathcal{T}$  be a logical theory defined in language  $\mathcal{L}$ . Let  $A$ ,  $\Omega$  and  $\Sigma$  be sets of sentences of  $\mathcal{L}$  respectively called *abducible* and *consistency* sets. A *strong logic-based abduction* is that of finding a primitive explanation  $\phi \in A$  of an observation  $\omega \in \Omega$  such that:

- $\mathcal{T} \cup \phi \vdash \omega$
- $\mathcal{T} \cup \Sigma \cup \phi$  is consistent.

When  $\Sigma = \emptyset$ , this strong abduction reduces to a normal logic-based one. From a theoretical view, the introduction of the consistency set  $\Sigma$  renders this definition of abduction more cautious than the classical one in the sense that every explanation for it is also a normal explanation but not vice-versa. In fact, because of the monotonicity of first order logic, if  $\phi$  is consistent with  $\mathcal{T} \cup \Sigma$  then  $\phi$  is consistent with  $\mathcal{T}$  too but not vice-versa. From a pragmatical point of view, the introduction of the consistency set  $\Sigma$  is almost uninfluent. In fact, a procedure for finding abductive explanations can be splitted in two subroutines: the first one takes  $\mathcal{T}$  and  $\omega$  for an explanation  $\phi$ , the second one checks for the consistency of  $\phi$  with  $\mathcal{T}$ , then it checks for the consistency of  $\phi$  with  $\mathcal{T} \cup \Sigma$ .

Let us particularize the preceding definition to the following clause-based abduction. A *literal* is an atomic sentence of  $\mathcal{L}$  or its negation. A *clause* is the disjunction of literals of  $\mathcal{L}$ . Every set of clauses is logically equivalent to a wff and vice-versa.

**Definition 3.** Let  $\mathcal{T}$  and  $\Sigma$  be sets of clauses of  $\mathcal{L}$ . Let  $\Omega$  and  $A$  be sets of *ground* literals of  $\mathcal{L}$ . A *clause-based strong abduction* is that of finding a primitive explanation  $\phi \subseteq A$  of an observation  $\omega \in \Omega$  such that

- $\mathcal{T} \cup \Sigma \not\vdash \neg \phi$ , that is  $\phi$  is consistent with  $\mathcal{T} \cup \Sigma$ ,
- $\mathcal{T} \cup \phi \vdash \omega$ ,
- $\phi$  is subset-minimal.

The last point of the definition embodies a selection criterion for "good" explanations. [Allemang 91] is called "parsimony". It prevents the choice, as explanations, of containing a proper subset that itself constitutes a valid explanation. If  $A$  is restricted to atomic sentences (clauses made of a single, *positive* literal) and if  $\Sigma = \emptyset$ , then this definition collapses to the one defined in [Konolige 92] and in [Reiter 87].

**Definition 4.** A *literal-based strong program* is a clause-based strong program in which the abducible set  $A$  is made of literals.

Although problematic from a pragmatical computational point of view, the conceptualization of clause-based abduction is conceptually stimulating because it provides a case in which, in order to explain some observations, it is not sufficient to hypothesize facts but it is necessary to hypothesize the presence of other rules in the theory ( $\neg\beta \rightarrow \alpha$ ). It is this kind of abduction that (along with induction) is at the base of the development of scientific theories.

There is no intrinsic relationship between the cardinalities of  $\phi$  and  $o$ . Depending on the theory  $\mathcal{T}$ , an explanation  $\phi$  can be a set of more than one clause even if  $o$  is made of a single clause and, vice-versa, a single clause can be an explanation of a set of clauses. So it can be useful to consider only explanations of a certain cardinality. To cope with this, we propose the following definition.

**Definition 5.** An  $n \times m$  *cp.* is a clause-based program in which the cardinality of  $o$  is  $n$  and the cardinality of the explanations  $\phi$  is forced to be  $m$ .

It is well known that the evaluation of alternative explanations turns out to be a non-trivial task. Global criteria as that of the "cardinality comparison" or that of "least presumptive explanation" [Poole 89]. Given a set of alternative explanations  $\mathcal{F}$  for a problem  $\mathcal{T} \cup \phi \vdash o$ , an explanation  $\phi_i \in \mathcal{F}$  is *less presumptive* than an explanation  $\phi_j \in \mathcal{F}$  if  $\mathcal{T} \cup \phi_j \vdash \phi_i$ . This definition can be stressed making it independent of  $\mathcal{T}$ :  $\phi_i \in \mathcal{F}$  is *strongly less presumptive* than an explanation  $\phi_j \in \mathcal{F}$  if  $\phi_j \vdash \phi_i$ . Given a clause-based program, its most cautious explanation is the strongly *least* presumptive one. Obviously, if  $\phi_i$  is (strongly) less presumptive than  $\phi_j$ , then  $\phi_j$  is said (strongly) *more* presumptive than  $\phi_i$ . In [Appelt 92] (p. 6) the authors suggest that some local metric criteria must be adopted. "neither a most specific nor a least specific abduction strategy is completely appropriate". In [Charniak 93] is presented a Bayesian model of plan recognition.

### 3. ACTIONS

I refer to the semantic of STRIPS given in [Lifschitz 87]. A *state theory*  $T$  is a consistent set of sentences of  $\mathcal{L}$ . A *state model*  $M$  is any consistent set of ground atomic sentences. A *world model*  $M^T$  is the consistent union of a state model  $M$  with a state theory  $T$ . A *state* represents the world where the robot is working in at a given instant. The *state theory* represents constraints that must be verified in every instant. To simplify the notation, we will use the same symbol for the relations that must be verified between facts at different instants. An *operator*  $de$

triple  $(P, D, A)$ , where  $P$  (the *preconditions*),  $D$  (the *delete list*) and  $A$  (the *add list*) are *atomic* sentences of  $\mathcal{L}$ . If these atomic sentences contain variables then the operator actually represents a family of operator descriptions, one for each possible instantiation of variables with the constants in  $\mathcal{L}$ .

A STRIPS system consists of an initial world model  $M^T_0$ , a set  $O_p$  of symbolic *operators*, and a family of operator descriptions  $\{(P_{\alpha}, D_{\alpha}, A_{\alpha})\}_{\alpha \in O_p}$ . Given a STRIPS system, a *plan* is any finite sequence of operators. Each plan  $\pi = (\alpha_1, \dots, \alpha_n)$  defines a sequence of world models  $M^T_0, M^T_1, \dots, M^T_n$ , where

$$M^T_i = (M^T_{i-1} \setminus D_{\alpha_i}) \cup A_{\alpha_i} \quad (i=1, \dots, n)$$

We say that a plan  $\pi = (\alpha_1, \dots, \alpha_n)$  is *accepted* by a STRIPS system if

$$M^T_{i-1} \vdash P_{\alpha_i} \quad (i=1, \dots, n)$$

In this case  $M^T_n$  is the result of executing  $\pi$  in  $M^T_0$  and we denote it  $\mathcal{R}(\pi)$ . We assume that the world described by the language  $\mathcal{L}$  as being, at any instant of time, in a certain state  $s$  such that the set of sentences satisfied in state  $s$  is closed under predicate logic.

**Assumption 1.** (The planner has complete and correct knowledge about the environment and the effects of each action in the world, so ..) If the robot has already performed an action/plan then the intent of the action/plan cannot be a disjunction of literals or a conjunction of literals.

**Definition 6.** A goal  $G$  is a literal or a conjunction of literals of  $\mathcal{L}$ .

#### 4. GOALS RECOGNITION

To simplify the matter (without loss of generality) let us begin with plans made of a single operator. Subsequently it is shown how a longer plan can be represented by a single operator. The results of this paper hold upon the following fundamental assumption:

**Assumption 2.** The planner is *correct* and *complete*, that is, if the planner verifies that a sentence  $p$  in a world model  $M^T$  then  $M^T \vdash p$  and if  $M^T \vdash p$  then the planner is able to prove  $p$ . Furthermore, if the planner plans an operator  $\alpha_i$  to pursue a goal  $G$  then:

$$M^T_{i-1} \not\vdash G, G \text{ is not satisfied in } M^T_{i-1} \quad (1)$$

$$M^T_i \vdash G, G \text{ is satisfied in } M^T_i \quad (2)$$

**Theorem 1.** If a planner plans an operator  $\alpha_i = (P_{\alpha_i}, D_{\alpha_i}, A_{\alpha_i})$  to pursue a goal  $G$ , then the problem to recognise the planner's goal in planning  $\alpha_i$  is a 1x1 strip goal base problem that has  $M^T_{i-1} \setminus D_{\alpha_i}$  as "theory",  $\neg A_{\alpha_i}$  as "observation",  $D_{\alpha_i}$  as "consistency set" and  $\neg G$  as explanation.

**Proof.** Let  $M^{T_{i-1}}$  be  $m^{T_{i-1}} \cup D_{\alpha_i}$  and  $M^T_i$  be  $m^{T_{i-1}} \cup A_{\alpha_i}$ . From the 1. and 2. it follows

$$m^{T_{i-1}} \cup D_{\alpha_i} \not\vdash G \quad (1a)$$

$$m^{T_{i-1}} \cup A_{\alpha_i} \vdash G \quad (2a)$$

From the 2a. it follows:

$$m^{T_{i-1}} \cup \neg G \vdash \neg A_{\alpha_i} \quad (2b)$$

If we take  $\phi$  for  $\neg G$ ,  $o$  for  $\neg A_{\alpha_i}$ ,  $\Sigma$  for  $D_{\alpha_i}$ , and  $\mathcal{T}$  for  $m^{T_{i-1}}$  we obtain:

- $\mathcal{T} \cup \Sigma \not\vdash \neg \phi$ ,
- $\mathcal{T} \cup \phi \vdash o$ .

The "observation"  $o$  is made of a single clause because it's the negation of the explanation  $\phi$  is made of a single clause because it's to be intended as the negation of a conjunction of literals, so this is a 1x1 a □

**Example 1.** In a classical block-world domain there are a table, three blocks and a hand. Consider the following instanced STRIPS operator (from [6])

$\alpha$ : **putdown\_b1**  
 $P_\alpha$ : holding\_b1  
 $D_\alpha$ : holding\_b1  
 $A_\alpha$ : ontable\_b1, clear\_b1, handempty

along with the following (piece of) world model  $M^T_0$ :

$M^T_0$  holding\_b1, ontable\_b2, ontable\_b3, clear\_b2, clear\_b3  
 $T$ : holding\_b1  $\vee$  holding\_b2  $\vee$  holding\_b3  $\rightarrow$   $\neg$ handempty  
 on\_b2\_b1  $\vee$  on\_b3\_b1  $\rightarrow$   $\neg$ clear\_b1  
 holding\_b1  $\rightarrow$   $\neg$ on\_b1\_b2  $\wedge$   $\neg$ on\_b1\_b3  $\wedge$   $\neg$ ontable\_b1  
 holding\_b1  $\rightarrow$   $\neg$ holding\_b2  $\wedge$   $\neg$ holding\_b3  
 ontable\_b1  $\wedge$  ontable\_b2  $\wedge$  ontable\_b3  $\rightarrow$  filled\_table  
 handempty  $\rightarrow$  ontable\_b1  $\vee$  on\_b1\_b2  $\vee$  on\_b1\_b3

Let STRIPS plan **putdown\_b1**;  $m^T_0 = \mathcal{T} \cup \{\text{ontable\_b2, ontable\_b3, clear\_b2, clear\_b3}\}$ ,  $D_{\alpha_1} = \{\text{holding\_b1}\}$ ,  $\neg A_{\alpha_1} = \{\neg \text{ontable\_b1} \vee \neg \text{clear\_b1} \vee \neg \text{handempty}\}$ . We obtain the following 1x1 a

- $\mathcal{T} \cup \{\text{ontable\_b2, ontable\_b3, clear\_b2, clear\_b3}\} \cup \{\text{holding\_b1}\} \not\vdash \neg \phi$
- $\mathcal{T} \cup \{\text{ontable\_b2, ontable\_b3, clear\_b2, clear\_b3}\} \cup \phi \vdash \neg \text{ontable\_b1} \vee \neg \text{clear\_b1} \vee \neg \text{handempty}$

here are some single-clause single-literal explanations with their correspondent  $\phi$

	$\phi$	$G$
1	$\neg \text{ontable\_b1}$	<b>ontable_b1</b>
2	$\neg \text{clear\_b1}$	<b>clear_b1</b>
3	$\neg \text{handempty}$	<b>handempty</b>
4	holding_b1	<b><math>\neg</math>holding_b1</b>
5	$\neg \text{filled\_table}$	<b>filled_table</b>

<sup>1</sup> A set of sentences is logically equivalent to their conjunction and the negation of a conjunction of sentences is logically equivalent to the disjunction of their negations.

Goals 1÷3 are trivially the atoms in the add list, the others depend on the state of the initial state model and they couldn't be recognized without abduction.

The abducible set  $A$  is made of the negations of all the possible goals for the planner. We can think of a possible goal as:

1. any conjunction of literals verified in  $M_i^T$  but not in  $M_{i-1}^T$  (1÷5- $G$  in the example)
2. any conjunction of literals verified in  $M_i^T$  but not in  $M_{i-1}^T$  with any other literal both in  $M_{i-1}^T$  and in  $M_i^T$ .

Correspondently, we can think of a possible single-clause explanation as:

1. any disjunction of literals verified in  $M_i^T$  but not in  $M_{i-1}^T$  (1÷5- $\phi$  in the example)
2. any disjunction of literals verified in  $M_i^T$  but not in  $M_{i-1}^T$  with the negation of a literal(s) verified both in  $M_i^T$  and  $M_{i-1}^T$ .

We may admit only goals of type 1. In this case we can define the abducible set of all the ground literals of  $\mathcal{L}$  and refer to the following result.

**Theorem 2.** If a planner plans an operator  $\alpha_i=(P_{\alpha_i},D_{\alpha_i},A_{\alpha_i})$  to pursue a goal  $G$ , the problem to recognise the planner's goal in planning  $\alpha_i$  is a 1×1 strip that has as base  $M_{i-1}^T \setminus D_{\alpha_i}$  as “theory”,  $\neg A_{\alpha_i}$  as “observation”,  $D_{\alpha_i}$  as “consistency set” and  $\neg G$  as *plausible* explanation .

**Proof.** Straightforward from the proof of the theorem 1 and the definitions 1 and 2 .

The set  $\mathcal{G}$  of abducible goals can be sorted according to some domain independent (for instance the number of literals in the goal) or domain dependent criteria (importance of the goal); then a goal can be selected as the abducted one.

Let us move to the significant general case of plans longer than a single action. The reason with STRIPS because its plans are simple sequences of operators and can be carried themselves with operators.

**Definition 7.** For each plan  $\Pi=(\alpha_1,\dots,\alpha_n)$  accepted by a STRIPS system we consider the triple  $(P_\Pi,D_\Pi,A_\Pi)$  its corresponding *composite operator* where:

$$\begin{aligned} P_\Pi &= P_{\alpha_1} \cup P_{\alpha_2} \setminus A_{\alpha_1} \cup P_{\alpha_3} \setminus (A_{\alpha_1} \cup A_{\alpha_2}) \cup \dots \cup P_{\alpha_n} \setminus (A_{\alpha_1} \cup \dots \cup A_{\alpha_{n-1}}) \\ A_\Pi &= A_{\alpha_n} \cup A_{\alpha_{n-1}} \setminus D_{\alpha_n} \cup A_{\alpha_{n-2}} \setminus (D_{\alpha_n} \cup D_{\alpha_{n-1}}) \dots \cup A_{\alpha_1} \setminus (D_{\alpha_n} \cup \dots \cup D_{\alpha_2}) \\ D_\Pi &= D_{\alpha_n} \cup D_{\alpha_{n-1}} \setminus A_{\alpha_n} \cup D_{\alpha_{n-2}} \setminus (A_{\alpha_n} \cup A_{\alpha_{n-1}}) \dots \cup D_{\alpha_1} \setminus (A_{\alpha_n} \cup \dots \cup A_{\alpha_2}) \end{aligned}$$

The generalization of the preceding methods to the present case is straightforward. What we have to do is substitute the triple  $(P_\Pi,D_\Pi,A_\Pi)$  to the triple  $(P_{\alpha_i},D_{\alpha_i},A_{\alpha_i})$ , but something can be said about goals' plausibility. The set  $\mathcal{G}$  of abducible goals of the overall plan can be sorted considering that the most important action in the plan should be the last one. In a soccer game every ball passing is functional to the last kick toward the goal, in each action is almost functional to the last one. However, those intermediate actions



contribute to the final goal by achieving partial goals not vanished by subsequent plan. That's why we have to apply the goal recognition method to the  $\Pi=(\alpha_1,..,\alpha_n)$ , but we can then order the recognized goals from those reached by to those accomplished by the action  $\alpha_1$ .

### 5. PLAN RECOGNITION THROUGH GOALS RECOGNITION

If the executed plan is part of an unknown longer plan, then goal recognition as a useful step in plan recognition. Given a STRIPS system, consider a plan  $\Pi$  the goal  $G$ , and its already executed initial part  $\Pi_0$ .  $\Pi_0$  itself is a plan, let  $\mathcal{G}_0$  be abducible goals,  $G_0 \in \mathcal{G}_0$  its abducted goal and  $(P_{\Pi_0}, D_{\Pi_0}, A_{\Pi_0})$  its corresponding operator. Let  $\Pi_x$  be the portion of the plan  $\Pi$  that has still to be executed,  $G_x$  be respect to the initial world model  $\mathcal{R}(\Pi_0)$  and  $(P_{\Pi_x}, D_{\Pi_x}, A_{\Pi_x})$  its corresponding operator. The plan recognition task is that of recognizing  $\Pi$  (i.e.  $\Pi_x$ ) from  $\Pi_0$ .

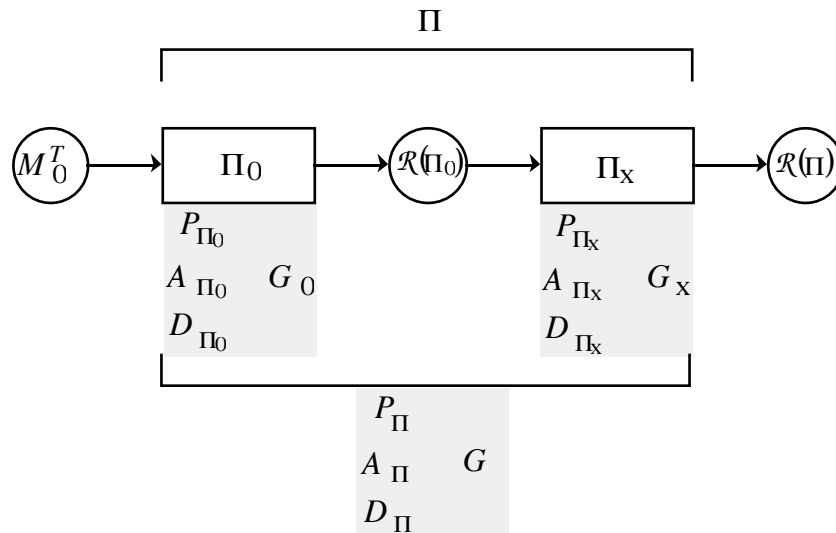


Fig 3. Already executed and still unknown portions of a STRIPS plan.

It can be argued that  $G_0 \neq G$  otherwise the rest of the plan  $\Pi_x$  would be a waste of time. It may be that:

- $G_0 \subset G$ , and/or
- $G_0 = P_{\Pi_x}$ .

If it is only that  $G_0 \subset G$  then  $\Pi_0$  and  $\Pi_x$  are two independent subplans. They may have been performed in parallel or even in the sequence  $\Pi_x, \Pi_0$ . However it is difficult for an observer to expect  $\Pi_x$  after  $\Pi_0$ . The interesting case is that  $G_0 = P_{\Pi_x}$ . This leads to a set of (composite) operators as candidates for  $\Pi_x$ . This set can be reduced if the overall recognized plan  $\Pi$  must be *sensible*. By "sensible" I mean that there must be one reason for the plan to be performed, that is the a

- $m_0^T \cup D_{\Pi} \neq G$
- $m_0^T \cup \neg G \vdash \neg A_{\Pi}$

must have at least one solution  $G \notin \mathcal{G}_0$ .

The reduced set can be sorted by domain independent criteria (for instance, t plan) or domain dependent criteria (for instance, the feasibility of the plan or its

**Example 2.** Consider the propositional STRIPS system of the previous example following operators:

<p><math>\alpha</math>: <b>assembly_layer</b>  <math>P_\alpha</math>: filled_table  <math>D_\alpha</math>: ontable_b1, ontable_b2, ontable_b3  <math>A_\alpha</math>: ontable_layer, handempty</p> <p><math>\alpha</math>: <b>assembly_tower</b>  <math>P_\alpha</math>: ontable_b1, on_b2_b1, on_b3_b2  <math>D_\alpha</math>: ontable_b1, on_b2_b1, on_b3_b2  <math>A_\alpha</math>: ontable_tower, handempty</p> <p><math>\alpha</math>: <b>puton_b3_b1</b>  <math>P_\alpha</math>: holding_b3, clear_b1  <math>D_\alpha</math>: holding_b3, clear_b1  <math>A_\alpha</math>: on_b3_b1, clear_b3, handempty</p> <p><math>\alpha</math>: <b>puton_b2_b1</b>  <math>P_\alpha</math>: holding_b2, clear_b1  <math>D_\alpha</math>: holding_b2, clear_b1  <math>A_\alpha</math>: on_b2_b1, clear_b2, handempty</p> <p><math>\alpha</math>: <b>puton_b3_b2</b>  <math>P_\alpha</math>: holding_b3, clear_b2  <math>D_\alpha</math>: holding_b3, clear_b2  <math>A_\alpha</math>: on_b3_b2, clear_b3, handempty</p>	<p>If the blocks cover all the table then the robot assembles them in a layer.</p> <p>If the block b3 is on b2, b2 is on b1 and b1 is on the table then the robot assembles them in a tower.</p>
<p><math>\alpha</math>: <b>puton_b2_b3</b>  <math>P_\alpha</math>: holding_b2, clear_b3  <math>D_\alpha</math>: holding_b2, clear_b3  <math>A_\alpha</math>: on_b2_b3 clear_b2, handempty</p> <p><math>\alpha</math>: <b>pickup_b2</b>  <math>P_\alpha</math>: handempty, clear_b2  <math>D_\alpha</math>: handempty, clear_b2  <math>A_\alpha</math>: holding_b2</p> <p><math>\alpha</math>: <b>pickup_b3</b>  <math>P_\alpha</math>: handempty, clear_b3  <math>D_\alpha</math>: handempty, clear_b3  <math>A_\alpha</math>: holding_b3</p>	

Let  $\pi_0$  be the single-action plan **putdown\_b1**. Here are three possible  $\pi_x$ :

$\pi_1 = \pi_0$ ,  $\pi_x1 = (\text{putdown\_b1}, )$  **assembly\_layer**  
 $P_{\pi_x1}$ : filled\_table  
 $D_{\pi_x1}$ : ontable\_b1, ontable\_b2, ontable\_b3  
 $A_{\pi_x1}$ : ontable\_layer, handempty

$\pi_2 = \pi_0$ ,  $\pi_x2 = (\text{putdown\_b1}, )$  **pickup\_b2, puton\_b2\_b1, pickup\_b3, puton\_b3\_b2, assembly\_tower**  
 $P_{\pi_x2}$ : handempty, clear\_b2, clear\_b1, clear\_b3, ontable\_b1  
 $D_{\pi_x2}$ : ontable\_b1, on\_b2\_b1, on\_b3\_b2, holding\_b3, clear\_b2, holding\_b2, clear\_b1  
 $A_{\pi_x2}$ : ontable\_tower, clear\_b3, handempty

$\pi_3 = \pi_0$ ,  $\pi_x3 = (\text{putdown\_b1}, )$  **pickup\_b3, puton\_b3\_b1, pickup\_b2, puton\_b2\_b3**  
 $P_{\pi_x3}$ : handempty, clear\_b3, clear\_b1, clear\_b2  
 $D_{\pi_x3}$ : holding\_b2, clear\_b3, holding\_b3, clear\_b1  
 $A_{\pi_x3}$ : on\_b2\_b3 clear\_b2, handempty, on\_b3\_b1

Each left subplan is itself a possible  $\pi_x$ . Consider the preconditions of these operators,  $P_{\pi_x1}$ ,  $P_{\pi_x2}$  and  $P_{\pi_x3}$ . If we limit the goal definition to any conjunctive literal verified in  $M^T_i$  but not in  $M^T_{i-1}$ , then only  $P_{\pi_x1}$  can be considered as possible goal. It should be noted that this result would not be achieved without abduction. If we limit the goal definition to encompass in the conjunction the negation of literal(s) verified before

$M_{i-1}^T$ , then even  $P_{\Pi_{x2}}$  and  $P_{\Pi_{x3}}$  are considerable as possible goals. However we rather than  $\Pi_3$  because the preceding plan  $\Pi_0$  led the world to a more specialized state where  $\Pi_{x2}$  is verified. It should be noted that  $P_{\Pi_{x2}}$  is (strongly) less presumptive than  $\Phi_2 = -P_{\Pi_{x2}}$ . Various examples of this kind suggest that  $P_{\Pi_{x2}}$  is better presumption, while not sufficient to discriminate possible goals, is better presumption.

## 6. CONCLUSIONS

If a STRIPS-like planner possesses sufficient inferential abilities, then the goal can be generalized to be some "logical consequence(s)" of the changes made in the plan. I've shown that, in such a case, under the hypothesis that the planner is complete, goal recognition can be regarded as a clause-based abduction problem. I've given a method to explain making a single clause as hypothesis. Furthermore, I've given a method to sort the abducible goals according to their plausibility and I show how goal recognition can be regarded as a step in plan recognition. However, STRIPS is a very simple plan recognition problem and can be characterized by many other features for hierarchical, parallel, conditional and temporal planners!

## REFERENCES

- [Allemang 91] Tom Bylander, D. Allemang, M. C. Tanner and J. R. Josephson, The computational complexity of abduction, *Artificial Intelligence*, 49,25-60, 1991.
- [Appelt 92] D. A. Appelt and M. E. Pollack, Weighted Abduction for Plan Ascription, *User Mode Dialogues: User-Adapted Interaction*, vol. 2, n 1-2, 1992
- [Bond 88] Readings in Distributed Artificial Intelligence, A. H. Bond and L. Gasser eds, Morgan Kaufmann Publishers, San Mateo, CA, 1988.
- [Carberry 90] Sandra Carberry, Incorporating Default Inferences into Plan Recognition, Proceedings of the American Association for Artificial Intelligence, 471-478, 1990,
- [Carberry 93] S. Carberry and W. A. Pope, Plan Recognition strategies for language understanding, *Man-Machine Studies*, Academic Press London, 39, 529-577, 1993.
- [Charniak 93] J. Charniak and R. P. Goldman, A Bayesian model of Plan Recognition, *Artificial Intelligence* 64(1), pp 53-79, 1993.
- [Dragoni 94] A. Dragoni and P. Puliti, Mental States Recognition from Speech Acts through Plan Recognition, Proceedings of the 11th European Conference on Artificial Intelligence, Amsterdam, 12, 1994.
- [Kautz 90] Henry Kautz, A Circumscriptive Theory of Plan Recognition, in *Intentions in Communication*, P. R. Cohen & J. Morgan & M. E. Pollack eds., MIT Press, Cambridge, Massachusetts, 1990.
- [Konolige 92] K. Konolige, Abduction versus closure in causal theories, *Artificial Intelligence*, 53, pp 1-24, 1992.
- [Lifschitz 87] V. Lifschitz "On the semantics of STRIPS", in *Reasoning About actions and Plans: Proceedings of the 1986 Workshop*, pp 47-82, Michael P. Georgeff and Amy L. Lansky eds., Morgan Kaufmann Publishers, San Mateo, CA, 1987.
- [Nilsson 81] N. J. Nilsson: "Principles of Artificial Intelligence" Springer-Verlag 1981
- [Paul 93] Gabriele Paul, Approaches to abductive reasoning: an overview, *Artificial Intelligence* 7, 109-152, Kluwer Academic Publishers, 1993
- [Poole 89] Poole, D., Explanation and Prediction: An Architecture for Default and Abductive Reasoning, *Computational Intelligence* 5(2), 97-110, 1989.
- [Poole 93] D. Poole, Probabilistic Horn Abduction and Bayesian networks, *Artificial Intelligence*, 6, 81-129, 1993.
- [Reiter 87] Raymond Reiter, Johan De Kleer, Foundations of Assumption-Based Truth Maintenance Systems, Proceedings of 1987 Conference of the American Association for Artificial Intelligence., 183-187, 1987.