Transition Logic Revisited

Wolfgang Bibel
Darmstadt University of Technology, Germany
Bibel@Informatik.TU-Darmstadt.De
also affiliated with
University of British Columbia

22nd April 2005

Abstract

A new version of transition logic is presented. It integrates (dynamic) transitions, which change world states, and classical (static) reasoning, restricted in the paper to Horn logic. This is achieved by defining a deductive relationship $\vdash_{(T,\leq)}$ among formulas for a partially ordered set $(T, \leq)$ of transitions. This novel integration might form the core for a unified framework for practical reasoning with the potential of a full exploitation of the maturing techniques from classical planning and deduction. For the chosen formula type the logic at the same time offers one possible clarification of the deductive formalism envisioned for original STRIPS but never made precise before.

1 Introduction

In the classical approach to Artificial Intelligence (AI), also called GOFAI (good old-fashioned AI), the basis of intelligent behavior consists in a rich knowledge base (KB) possibly containing millions of items. An intelligent agent, equipped with such a KB, would model the current state of the world and deduce from it, on the basis of KB, a sequence of actions to reach its goals. Thereby the deduction is carried out by some general mechanism intended to model human reasoning to a certain degree.

Deductive planning is a special setting of this general scenario. Hereby we assume that the KB is represented as a set (or conjunction) W of logical formulas, some of which formally model the possible actions. Similarly, the current (or initial) state I and the goal state G both are formulas. The general mechanism is provided by classical first-order deduction which attempts to deduce G from W and I, formally $W \cup I \vdash G$. From a successful deduction (or proof) a plan can easily be extracted by noting the formulas representing actions, which are used in the proof, along with their relative ordering.

While the idea underlying deductive planning has always been appealing, there has also been an inherent problem with it. This problem arises from the intrinsic conflict between the *static* formalism of first-order logic (FOL) and the *dynamics* of a world which changes by actions and for other reasons. Technically, we speak of *transitions* to refer to such changes. The problem thus amounts to the fundamental question how to marry the static FOL with the dynamic transitions.

Numerous attempts have been made to solve this problem (and involved ones like the famous frame problem). We briefly review several of them in Section 4, pointing out their respective weaknesses. But none of them has become a winning candidate for a solution, thus leaving deductive planning – and with it the more general problem of modelling intelligent behavior – at an impasse. In terms of practical applications the most successful approach among these is *classical planning* à la STRIPS [FN71]. Systems based on this approach demonstrate a remarkable performance indeed. But most of them have dispensed with the deductive power of FOL and focus exclusively on the transitions (also called actions or operators), resulting in a formalism which excels on the dynamic side but lacks the static one altogether and thus seems to be too poor for more general planning and problem solving. As a special instance of our fundamental question one might therefore also consider the question how to marry FOL with STRIPS.

STRIPS in its original form pretended to have achieved the marriage. Upon closer inspection it turned out however that the underlying formalism was not clear at all. One of the two main problems was its semantics. This problem was later solved in [Lif86]. The other was how to integrate derivations which extend across transitions in a precisely defined way. This second problem is solved in the present paper. In other words we integrate classical planning and deduction techniques resulting in what might become the core of a unified framework for practical reasoning. Thereby we restrict the underlying logic in our discussions to (function-free and stratified) Horn clause logic.

The author has attacked the fundamental question just described in a series of papers including [Bib86, Bib98]. The present paper in a certain sense completes these attempts with a simple formalism overcoming several gaps and weaknesses in the earlier versions. The formalism is based on the following view the intelligent agent takes of the world.

First of all, there is in principle only a single world under consideration characterized by I, W in the setting above. Reasoning within this world is done in classical FOL (without the need for situational parameters or such). In addition, the formalism accommodates the transitions as local changes in I, W, quasi within an orthogonal dimension and without more interference with the "deductive" dimension than absolutely necessary. These transitions may be thought of as executions of STRIPS-like operators. Of course, the world resulting from I, W by a transition is different from I, W, so that we might speak of a "different" world (thus in a strict sense resulting in many worlds). But since everything remains the same in I, W except for the local change caused by the transition, we rather regard it as the same world – exactly as we humans regard the real world as one and the same regardless of the myriads of changes taking place at any moment. The difficulty with the fundamental question was how to formally connect reasoning chains before and after such a change caused by a transition. Similarly as people do this at ease, the formalism just takes formal note of the local change in I, W and otherwise rests on first-order reasoning, a conceptually truly simple solution. Formally, we obtain a deductive relation of the kind described above, namely $W \cup I \vdash_T G$ whereby T denotes the set of transitions (possibly along with some partial order on T).

In such a setting deductive reasoning can to some extent be eliminated but at the cost of increasing the number of different transitions (or actions). Classical planning tends to take advantage of this possibility and to dispense with deductive reasoning altogether, putting the burden on the shoulders of the planning engineers who need to anticipate any possible deductive reasoning and to consider it offline in the modelling of the actions which leads to an overall increase of the complexity of the task to be solved by the planning system as we will point out in Section 4. Human reasoning in contrast amounts to a well-balanced trade-off between handling a minimal number of actions along with a rich reasoning capability. With transition logic we could model this trade-off in our systems.

In order to provide as easy an access to the new formalism as possible we introduce it in the next section under rather restricted assumptions, namely with a definite (ie. Prolog-like) formalism within propositional logic for the world description. Since already this restricted version of a transition logic demonstrates all characteristic points, we can be rather short in describing various possible extensions (generalization to the first-order level, addition of negation and constraints) within Section 3. In Section 4 we compare the transition logic in its new version with its ancestors as well as with the most important alternative formalisms for reasoning about actions and change. Section 5 then concludes the paper with a summary and with the prospects for applications and future research.

2 The propositional case

In the present section we provide a solution for the case of propositional logic to the fundamental question stated in the Introduction. For this purpose we assume a propositional language, ie. an alphabet consisting of propositional variables $P, Q, \ldots$ and the propositional junctors $\neg, \wedge, \vee, \rightarrow, \ldots$ along with the usual rules for defining (propositional) formulas. The well-known concepts of an interpretation (defined eg. as a subset of the variables), of satisfiability, validity, semantic inference $\models$, and of some consistent and complete calculus with a syntactic inference relation $\vdash$ all are assumed to be familiar (eg. see [Bib93]).

On this basis let us consider a very simple blocksworld example with blocks (of which we focus only on a single one), a table and a robot hand. The world description W_1 captures several facts including the following ones. If the block under consideration is uncovered (U) then it is also free (F), ie. $U \rightarrow F$. If the hand holds the block (H) then the block is airborne (A), ie. $H \rightarrow A$. If the block is put down (D) then it is on the table (O), ie. $D \rightarrow O$. Initially U ($= I_1$) holds and the goal G_1 is O which can be achieved by the set T_1 of two possible transitions, namely a free block can be picked up, ie. $F \Rightarrow H$, and an airborne block can be put down, ie. $A \Rightarrow D$. To apply any of these two transitions, their preconditions must be satisfied, eg. F must be deduced from the initial state U by classical deduction, viz. $U, U \rightarrow F \vdash F$. This justifies the transit from F to H via the first transition rule, so that we might now try to solve the possibly reduced planning problem which is to reach the goal O from the new initial state H given the world description W_1 and the transitions T_1. Pursuing in this way the resulting "plan" might be represented in the form of the following "deduction".

$$U, U \to F \vdash F/H, H \to A \vdash A/D, D \to O \vdash O$$

The plan consists in applying the two transitions (or actions) $F \Rightarrow H$ and $A \Rightarrow D$ which in the deduction are shown by way of the replacements F/H and A/D. Prior to the application of the second transition another deductive step is required like for the first one. The goal is then reached by the third deductive step.

We may think of two orthogonal dimensions of progression, the horizontal one for deductive reasoning and the vertical one for transitions. While the deductive dimension is well-known and needs no further explanations, it is the transitional dimension which in the deductive context requires additional elaboration and some terminology. To begin with the concepts of a transition and a transitional problem are introduced.

Definition 1. A *transition* is a pair (F_1, F_2) whereby F_i, $i = 1, 2$, is a conjunction of atoms. A *transitional problem* is a quadruple (W, T, I, G) whereby the world description W is a (finite) set of definite formulas, T is a (finite) set of transitions, and the initial and goal states I and G are conjunctions (or sets) of atoms. A pair like (I, W) of a state together with a world description is called a *generalized state*. For ease of presentation we assume here that all atoms of a generalized state are in I so that W consists only of proper rules.

To keep things in this first setting as simple as possible we have restricted here the formulas in W, T, I, G to their simplest possible form. For possible extensions see the subsequent section. In terms of STRIPS operators the first component F_1 of a transition is the operator's precondition, ie. a conjunction of atoms under our restrictions; the second component F_2 is the conjunction of the atoms in the add list A along with the ones occurring in the precondition except for those in the delete list D, ie. $F_2 = P \cup A$ with $P = F_1 \backslash D$. Those precondition atoms P of a STRIPS operator not in the delete list (quasi the "proper" precondition atoms) occur in both components of a transition. Intuitively one might think that they are deleted along with D and then again added (along with A). Under this view we may require that $D \cap A = \emptyset$ which we do in this paper although one might argue that this is a matter of discretion for the user of the formalism. Altogether, the slight – and, in terms of its semantic effect, neutral – deviation of the concept of transition from the STRIPS formalism presents itself under deductive considerations (cf. [Bib98]) and also simplifies the definition of transitions as only two instead of three lists are needed, although the finer distinction made in STRIPS will still show up in the Definition 3 below. Note that we blur the distinction between conjunctions, sets and lists of atoms as long as no confusion may arise.

The next step now consists in a formal clarification of what exactly changes in the state of the world (or "world model" in STRIPS terminology) by executing a transition. Recall the first transition in our example which replaces a free block (F) by a held one (H). Before this transition, the state of the world is characterized by the generalized state (I_1, W_1). What does the world look like after the transition? Intuitively, along with the derived F also its deductive "source", viz. U, should disappear in the changed world while the "carrier" of the deduction, $U \to F$, as general knowledge (along with everything else in I_1, W_1) should of course stay true. In other words, we need to define a function σ which enables to determine such a deductive source for a given precondition like F. This function will be defined below such that $\sigma(F) = \{\{U\}\}$ in our example (whereby the reason for the braces will become clear shortly).

Modifying the example slightly assume that the rule has two premises, ie. $U_1 \wedge U_2 \to F$. There would be two alternative sources to be deleted from the (appropriately modified) initial state in this case, since by the deletion already of one of the two premises the rule would become inapplicable. In consequence $\sigma(F) = \{\{U_1\}, \{U_2\}\}$ specifies two alternatives for removal upon the transition.

To illustrate also the second possible generic case of complication let us assume that instead of the single rule there are two rules with the same conclusion, ie. $U_1 \to F$ and $U_2 \to F$ in the simplest case. Then either premise is a source to be deleted, ie. $\sigma(F) = \{\{U_1, U_2\}\}$. This altogether illustrates why the source is a set of alternative sets of atoms from the initial state to be deleted. It also illustrates that each of these sets comprises a minimal set of literals whose deletion from the state repeals the entailment of F.

These ideas lead to the subsequent semantic and syntactic definition of this set which is prepared as follows. In order to keep its details as simple as possible we make two assumptions without loss of generality. First we regard the literals L in a state as rules $\top \to L$ so that any generalized state has nothing but rules. Second we abbreviate all rules with the same conclusion by a single rule with alternative premises. For instance, the two rules in the last modification would be abbreviated as $U_1 \vee U_2 \to F$. Third we assume that the set of literals in the delete list of any transition is just a singleton, ie. a single atom, which may be achieved by introducing for each transition with a non-atomic deletion component D in F_1 a rule $D' \to B$, whereby D' is the disjunction of the literals in D, and substituting the transition (F_1, F_2) by $(P \cup B, F_2)$, where B is an atom not mentioned anywhere else in the transitional problem. This atom may be used as the name of the transition. Note again that these changes are just for the sake of enabling a compact definition and need not actually be carried out in practice.

Since we want to define the source as a set on the basis of logical formulas, matters are harmonized if we use the set-theoretic representation of formulas which is widely used in the deduction community (eg. see [Bib93]). One of its advantages is that it abstracts from multiple occurrences of literals and thereby simplifies the formalism. In this representation a formula (of the simple kind considered here) is a *set of sets of ... of atoms* which in our context logically is interpreted as a *disjunction of conjunctions of disjunctions of ... of atoms*. For instance, the set $\{\{A, \{B, C\}\}, \{D\}\}$ logically reads $A \wedge (B \vee C) \vee D$ (and vice versa) whereby we use the standard assumption that $\wedge$ binds stronger than $\vee$ to spare parentheses.

With this correspondence between nested sets of atoms and logical formulas it is straightforward to introduce set-theoretic notions corresponding to well-known logical ones. For instance a formula in disjunctive normal form set-theoretically is a set of sets of atoms, ie. the nesting structure is limited to depth 2. Since we know from any standard text in logic that any formula may be transformed into disjunctive normal form, there is also a corresponding function, which we call nf, which transforms the corresponding set-theoretic version of the formula to its normal form (in the set-theoretic sense). The transformation process mimics the one known from logic. Thereby $\perp$ (or undefined) plays the usual Boolean role of *false*, so that eg. $\mathrm{nf}\{\{A,B\},\{C,\perp\}\} = \{\{A,B\}\}$; set-theoretically $\perp$ plays the role of the empty set $\emptyset$. In the subsequent definition we also use a standard function, here denoted by subs, which eliminates all subsumed alternatives (ie. sets of atoms) in its normal form argument. Recall that $\mathrm{subs}\{S_1,S_2\} = \{S_i\}$ if $S_i \subseteq S_j$, and $\mathrm{subs}\{S_1,S_2\} = \{S_1,S_2\}$ else, so that for instance $\mathrm{subs}\{\{A\},\{A,B\}\} = \{\{A\}\}$.

Above we have illustrated the source of an atom if this atom is the conclusion of a rule the premise of which consists of atoms of the initial state. If some of these atoms are not in the initial state but again in the conclusion of a rule then the same process has to be iterated which is realized in the following inductive definition by a recursive formula for σ (and illustrated thereafter). The definition begins with a formalization of the semantic characterization of the source illustrated above for which we will show thereafter that it is satisfied by the syntactic definition.

Definition 2. For a (satisfiable) world description W containing only definite rules, for a state S and an atom A the *semantic source of* A in the generalized state (S,W) is a set of minimal subsets M of S such that $S \setminus M, W \not\models A$.
The function $\sigma_{(S,W)}$ determining the *source* $\sigma_{(S,W)}(A)$ of an atom wrt. the generalized state (S,W) is inductively defined as follows. (Wherever the generalized state is clear from the context it will be deleted.)
If $B_{11} \wedge \ldots \wedge B_{1m_1} \vee \ldots \vee B_{n1} \wedge \ldots \wedge B_{nm_n} \rightarrow A \in W$ then
$\sigma'(A) = \{\{\sigma'(B_{1j_1}),\ldots,\sigma'(B_{nj_n})\} \mid 1 \leq j_1 \leq m_1,\ldots,1 \leq j_n \leq m_n\}$, whereby $\sigma'(B_{ij_i}) = A$ for $B_{ij_i} = \top$, else $\sigma'(A) = \perp$; finally, $\sigma(A) = \mathrm{subs}\,(\mathrm{nf}\,(\sigma'(A)))$.

For illustration let us apply the syntactic part of this definition to the two generic examples above. For the rule $R_1 = U_1 \wedge U_2 \rightarrow F$ we get $\sigma(F) = \mathrm{subs}\,(\mathrm{nf}\{\{\sigma'(U_1)\},\{\sigma'(U_2)\}\}) = \mathrm{subs}\,(\mathrm{nf}\{\{\{\{U_1\}\}\},\{\{\{U_2\}\}\}\}) = \mathrm{subs}\{\{U_1\},\{U_2\}\} = \{\{U_1\},\{U_2\}\}$, assuming the state is $S' = \{U_1,U_2\}$. Obviously, the definition does nothing else than recursively tracing back from F the deductive chain beginning with S' and ending in F, like any Prolog interpreter does. Starting with F the recursion has to consider the rule for F with two independent subgoals U_1, U_2, giving rise to two independent deductive chains (which may formally be defined as partially ordered sets of connections as in [Bib93], or as partially ordered sets of subgoals). Since $U_1 \in S'$, the second call of σ' has to consider the rule $\top \rightarrow U_1$ so that the third call results in $\sigma'(\top) = U_1$ according to the terminating case in the definition. The case for U_2 is analogue. While S' along with the rule entails F this obviously is no more the case if either of the sets in $\sigma(F)$ is removed from S'. In other words, $\sigma(F)$ is the semantic source of F in $(S',\{R_1\})$. The proposition below states that this is the case in general.

Similarly for the rule $U_1 \vee U_2 \rightarrow F$ we get $\sigma(F) = \mathrm{subs}\,(\mathrm{nf}\{\{\sigma'(U_1),\sigma'(U_2)\}\}) = \mathrm{subs}\,(\mathrm{nf}\{\{\{\{U_1\}\},\{\{U_2\}\}\}\}) = \mathrm{subs}\{\{U_1,U_2\}\} = \{\{U_1,U_2\}\}$, assuming again that the state is $\{U_1, U_2\}$. Here we only have a single deductive chain since the two subgoals are equal-ranking alternatives (in deductive terms they are both contained in a path through the formula rather than in different ones as in the previous example, cf. [Bib93]) so that both need to be removed from the state to make F false.

This altogether illustrates that in the general case σ does nothing else than marking in S for each deductive chain the subset of literals which are deductively connected to A. Algorithmically, in the present case of definite rules on the propositional level, this marking procedure can be solved in linear time for each single set in $\sigma(A)$. In fact our recursive definition obviously defines such a linear marking algorithm. The number of sets in $\sigma(A)$ is bounded by 2^S, ie. in complex scenarios there may be exponentially many alternatives. This then is a complexity intrinsic in the particular scenario and not a particular weakness of our proposed method for treating transitions. In many applications the number of literals in S is rather small so that the number of possible alternatives is rather restricted.

Note that the set of rules in W may be recursive (which will become relevant for the first-order case). While this may in principle lead to deductive chains of infinite length, the stabilized marking effect is always achieved in a finite number of steps. For instance, if $S = \{G, F\}$ and $W = \{F \wedge G \rightarrow F\}$, then $\sigma(F) = \{\{F\}, \{G\}\}$ while there are infinitely many deductive chains of arbitrary length for each of the two subsets. In such a case we may restrict the consideration to the shortest one for each subset, ie. the chain with a single connection (or subgoal) for $\{F\}$ and the chain with two connections (or subgoals) for $\{G\}$ in our present example. If there is no such shortest chain as in the case of $S = \emptyset$ and $W = \{F \rightarrow F\}$ then $\sigma(F) = \perp$ according to Definition 2.

The example with $S = \{D, E\}$ and $W = \{C \wedge D \rightarrow F, D \vee E \rightarrow C\}$ illustrates why subsumption is needed to determine the semantic source, leading to $\sigma(F) = \{\{D\}\}$ in this case. We are now stating and proving the proposition announced earlier.

Proposition. For a generalized state as in Definition 2 and for an atom A, $\sigma(A)$ is the semantic source of A.

Proof. Let $M \in \sigma(A)$ which by definition of σ implies that there is a deductive chain relating M with A. We have to show that M is a minimal set such that $S \setminus M, W \not\models A$. The proof is by induction on the length of the deductive chain involved. If $A \in M$ then its deletion from S obviously allows A to become false. Otherwise there is a rule with head A in W of the form shown in Definition 2 and some set $\{B_{1j_1}, \ldots, B_{nj_n}\}$ as part of the particular deductive chain. By the induction hypothesis we can assume that there is an interpretation under which $S \setminus M$ and W are true and all literals in this set are false and that M is minimal in this respect. Since this minimally falsifies each disjunct in the rule, we see that logically A may be false as well while no smaller set would do so, given that $\sigma(A)$ is the result of a subsumption operation, qed.

With the concept of the source of a set of literals we can now provide a precise answer to the above question about the state of the world after a transition. Since the source in general leads to different alternatives we have to consider thereby a set of alternative states rather than a single one.

Definition 3. For a set $s = \{S_1, \ldots, S_m\}$ of states, $m \geq 1$, a world description W and a transition $t = (P \cup D, P \cup A)$, the state after performing the transition is $s' = \bigcup_{i=1}^{m} \{(S_i \setminus D') \cup A \,|\, D' \in \sigma(D), S_i \setminus D', W \models P\}$. s' is called the state set *resulting from s by t*. The formula S' *corresponding to* s', also called the state set formula, is the disjunction $S'_1 \vee \ldots \vee S'_m$ whereby S'_i is the conjunction of the atoms in the ith component of the union for s', $i = 1, \ldots, m$.

Examples for illustration of this definition will follow below. Here we only want to point out the fact that the definition restricts each alternative succeeding state such that the precondition P does persist under all circumstances. To illustrate this point consider the example $S = \{B\}$, $W = \{B \rightarrow C\}$ and $t = (\{B, C\}, \{B, D\})$ for which Definition 3 produces $s' = \emptyset$ because the condition $\emptyset, W \models P$ in the formula for s' does not hold. In other words, because the source of the delete literal C at the same time is a source of the proper precondition B it must not be deleted so that the intended transition cannot be performed (as realized in the subsequent definition). We feel that this models exactly the meaning of a proper precondition.

With all this terminology we are now in a position to define a deductive relationship in the presence of transitions.

Definition 4. For a transitional problem (W, T, I, G) we define the deductive relation $\vdash_T$ as $\vdash_T = \vdash_{T^n}$ for some $n \geq 0$, which in turn is defined as $\vdash_{T^n} = \vdash_{(t_1, \ldots, t_n)}$ for some $t_1, \ldots, t_n$, $t_i \in T$ for $i = 1, \ldots, n$, whereby $W, I \vdash_{(t_1, \ldots, t_n)} G$ and the state set resulting from I by the sequence $(t_1, \ldots, t_n)$ of transitions is defined inductively as follows.

- If $n = 0$ then $\vdash_{()} = \vdash_{T^0} = \vdash$, whereby $\vdash$ denotes some classical deductive relation, and the state set is $\{I\}$.
- By the induction hypothesis we may assume that $\vdash_{(t_1, \ldots, t_{n-1})}$ is defined and that the state set s resulting from I by the sequence $(t_1, \ldots, t_{n-1})$ is $s = \{S_1, \ldots, S_m\}$ for some states S_i, $i = 1, \ldots, m$, $m \geq 0$. If $t_n = (P \wedge D, P \wedge A)$ and $W, I \vdash_{(t_1, \ldots, t_{n-1})} P \wedge D$, then the state set resulting from I by the sequence $(t_1, \ldots, t_n)$ is s' (as in Def. 3), and $W, I \vdash_{(t_1, \ldots, t_n)} F$ iff $W, S' \vdash F$, whereby S' is the formula corresponding to s'.

Let us illustrate these definitions with some variants of the initial example $E_1 = (W_1, T_1, I_1, G_1)$. For E_1 itself with its two transitions t_1, t_2 we obviously have $W_1, U \vdash_{(t_1, t_2)} O$; the sequence of states is U, H, D (ie. the state sets each consists of a singleton which in turn is a singleton).

Let E_2 be like E_1 except that T_2 has $t_3 = U \Rightarrow H$ as an additional transition which renders the deductive chain resulting in F along with t_1 as redundant since there is now also the simpler deduction $W_1, U \vdash_{(t_3, t_2)} O$. This illustrates that additional transitions may substitute deductive reasoning steps as mentioned already in the Introduction. STRIPS planning usually takes advantage of this fact by eliminating deductive steps altogether at the cost of requiring many more transitions than actually necessary (see Section 4 for notable exceptions).

For our third example think of a room with two closed windows, $C1, C2$, so that $I_3 = \{C1, C2\}$. Assume that W_3 contains a rule $C1 \wedge C2 \rightarrow S$ which might be interpreted as "the room is stuffy if the two windows are closed". Finally, assume T_3 to contain the transition $S \Rightarrow F$ which makes the stuffy room fresh. Of course fresh air comes in by opening at least one of the windows, but the transition does not specify which of the two. So the state set after executing the transition is $\{\{F, C1\}, \{F, C2\}\}$ according to the two minimal possibilities. The corresponding formula is $F \wedge C1 \vee F \wedge C2$. So we get $W_3, I_3 \vdash_{T_3} F$ since $F \wedge C1 \vee F \wedge C2 \rightarrow F$ is a classically valid formula. Had the problem been to derive only one of the two alternatives, say $F \wedge C1$ then only one alternative in the source would have to be determined in a goal-oriented way which is an important way of cutting down in the number of alternatives. This example also illustrates why as the source only a minimally possible subset of atoms (and not, for instance, all literals of the premise or even the entire I) is replaced by the transition.

Consider a variant of this example with $I_3' = \{C, F, G\}$, $W_3' = \{C \wedge G \rightarrow S, F \rightarrow S\}$ and $T_3' = \{G \wedge S \Rightarrow G\}$. C, F, G, S may be read as "window closed", "fire in room", "window in good condition", and "stuffy", respectively. We obtain $W_3', I_3' \vdash_{T_3'} G$. Note that no other alternative state after the transition is possible since the source F, G is discarded by the constraint on the precondition in Definition 3.

As a further example of the same kind but back in the blocks world let E_4 be like E_1 except that $I_4 = \{U, V\}$, $G_4 = \{O, V\}$, and W_4 is obtained from W_1 by replacing the rule $U \rightarrow F$ by $U \wedge V \rightarrow F$ yielding the following deduction: $W_4, U, V \vdash_{(t_1,t_2)} O \wedge V \vee O \wedge U \vdash O$.

Our formalism is powerful enough to allow the designer to produce surprising effects as the following example demonstrates. Let $I_5 = \{B, D\}$, $W_5 = \{A \wedge B \rightarrow D\}$ and $T_5 = \{D \Rightarrow A\}$. We get $\sigma(D) = \{\{D\}\}$, but have also $W_5, I_5 \vdash_{T_5} D$. That is D is both deleted and reintroduced upon the transition, but for two different reasons. Namely, it is deleted on the basis that the delete literal is in I_5 while the rule (because of the occurrence of $\perp$) by subsumption does not contribute to the source. On the other hand, by introducing A instead of D the rule is activated after the transition thus deductively yielding D. In fact, a second instance of the transition from there will yield A (and nothing else anymore).

So far we have considered only linearly ordered plans which is, of course, an unnecessary restriction. Consider example E_6 which again is like E_1 except that $I_6 = \{U, V\}$, $G_6 = \{A, B\}$, and T_6 has $t_6 = V \Rightarrow B$ as an additional transition. Then we obtain $W_1, U, V \vdash_{\{t_1,t_6\}} A \wedge B$ which leaves open an ordering among the two engaged transitions since these may be activated completely independent of each other. In this vein, we consider in general a partial ordering $\leq$ among the transitions T involved in the final plan, like in classical planning. In other words, the deductive relationship from Definition 4 more generally is defined as $\vdash_{(T,\leq)}$.

The distinction between the world knowledge W and the initial state I is somewhat arbitrary, but stays in line with common terminology in the planning community which explicitly refers to "domain axioms" in addition to a planning problem.

Due to the author's preferences we have chosen to introduce the formalism mainly in a syntactic way. The underlying semantics is straightforward, namely a classical semantics within each of the considered states of the world and a change of the semantics by each transition in the way precisely defined in [Lif86]. This semantics is as close to our intuitions as it possibly could be (admittedly a subjective statement, but one shared apparently by all classical planning researchers). We omit here the straightforward formal definitions on which then a consistency and completeness theorem could be based in the standard way.

3 Generalization to the first-order level

Naturally we wish to generalize the formalism introduced in the previous section from the ground to the first-order level. Since function symbols are rarely needed in planning we will not consider them here. Let us first stay within Horn clause logic. With regard to the deductive part this amounts to allowing (standard) first-order (instead of propositional) atoms, leaving everything else as before. The transitions in principle remain propositional although we allow variables as parameters for ground terms (like in default rules). In terms of the transition notation the move-operator $m(x, y, z)$ from the blocks world thus reads:

$$CLEAR(x) \wedge ON(x, y) \wedge CLEAR(z) \Rightarrow CLEAR(x) \wedge ON(x, z) \wedge CLEAR(y)$$

Thereby as before $CLEAR(x)$ is the "proper" precondition, ie. the one not on the delete list, which appears twice in the transition. So if the initial state I is characterized by $U(b_1)$, ie. block b_1 is uncovered, $\exists y ON(b_1, y)$ (more precisely $\exists y! ON(b_1, y)$), and $CLEAR(b_2)$, and W contains the rule $U(x) \rightarrow CLEAR(x)$ then the source of the two literals on the delete list, $ON(x, y), CLEAR(z)$, in m, determined exactly as in the previous section, is $\{\{\exists y ON(b_1, y), CLEAR(b_2)\}\}$ which upon an appropriate instantiation of $m(x, y, z)$, determined by unification, is substituted in the world after the transition by the instantiated add-list $ON(b_1, b_2), CLEAR(c)$. Altogether this leads to the deduction $W, I \vdash_{(m)} CLEAR(b_1) \wedge ON(b_1, b_2) \wedge CLEAR(c)$ whereby c is the Skolem term for the y in I. $CLEAR(b_1)$ is called a *derived* predicate in the planning community. As we see our solution from the previous section carries over to the general level in all details except for the unificational aspects to be added when lifting ground phenomena to the first-order level in the standard way.

To put these observations into more general terms, the source atoms are determined as on the ground level (according to the definitions of the previous section). The corresponding deduction gives rise to a substitution σ which instantiates the atoms involved, especially those in the transition rule, but possibly also those in the source. If the source atoms are originally ground (as is the case in most applications) then their replacement is carried out as on the ground level.

If an atom in the source is not ground then only those instantiations of the atom are replaced which are determined by σ. There are two further generic cases corresponding to the two types of quantifiers. One is illustrated by the previous example where $ON(b_1, y)$ is fully deleted since y is substitutable by a single Skolem constant c only. For the other generic case, imagine an initial state $\forall x A(x)$ and a precondition $A(c)$ in a transition, both deductively related to each other. Then the resulting substitution demands the replacement of only those instantiations of $A(x)$ which unify with $A(c)$ while all remaining ones may stay in the world after the transition. Formally this requires the source $\forall x A(x)$

to be replaced by $\forall x(x \neq c \rightarrow A(x))$, ie. by an atom whose instantiation domain is restricted in the world after carrying out the transition. So in the general case we are actually dealing with substitution domain restricted atoms in the sense just illustrated.

What is still lacking for practical applications is negation. Following the Prolog tradition negation may be introduced and deductively handled as finite failure. In order to avoid complications due to unrestricted negation it has become standard in applications like planning to restrict the rules to a so-called *stratified* rule set [Llo93, THN03] which, for instance, features unambiguous semantics and guarantees the well-foundedness of the iteration in determining the source of a literal even in the case of recursive rules. How this generalization is to be carried out in detail and in terms of the definitions from the previous section, is left for future research. At present we believe that the generalization of our definitions to stratified rule sets with negation does not affect our mechanism. Namely, if there is a source of a negated atom then it is handled as before; if there is no explicit source and the truth of the literal is established by the failure mechanism then there is obviously no need to remove a (syntactically non-existent) source. Similarly the addition of domain constraints is believed not to affect the definition of the source function in any way, so that in summary one would have a formalism with an expressivenes of the kind used for instance in [THN03]. But at present we are not quite there yet.

4 Comparison with existing techniques

We want to compare the transition logic (TL) introduced in this paper with existing techniques in planning and common-sense reasoning and begin this review with its closest relative, the STRIPS formalism.

The STRIPS approach as originally specified in [FN71] in spirit as in its details has much in common with TL. In Section 2 we have already pointed out that STRIPS operators are the same as our transitions (except for irrelevant syntactic sugar). Original STRIPS even had no restrictions on the formulas like those assumed in this paper for TL and therefore seems to be even more general than TL. But from the point of view of TL original STRIPS suffers from two main problems.

First, its semantics lacked precision. This problem has been overcome with the clarifications in [Lif86] to which we referred in view of the semantics of TL as well. Second, although STRIPS dealt already with a full integration of deduction as envisaged in TL, it again was rather vague about the details in this respect. So despite the informal description in [FN71, p.198] it remained unclear exactly which derived clauses had to be removed with "clauses on which the derived clause depends" and how the deductive relationship should look like. [Lif86] did not mention this second problem at all, presumably assuming this to be taken care through an appropriate operator definition by the user. Also [FN93] does not re-address the problem. TL has now clarified also this second problem for the class of formulas under consideration.

Most modern STRIPS-based planning systems lack the generality of original STRIPS in that they dispense with any deductive mechanism altogether. In the terminology of the present paper these systems deal only with the special case of a transitional problem where $W = \emptyset$. So the question arises whether the more general case $W \neq \emptyset$ considered here offers any advantages. Recall in this context from the previous two sections (example E_2 in Section 2 and the move example in Section 3), that deductive rules may be replaced by additional transitions.

The same question was recently addressed in [THN03] in the context of PDDL, namely whether PDDL axioms (like our W) could (and should) be "compiled away" in this sense of replacing them by additional transitions. "As it turns out, axioms are an essential feature because it is impossible to compile them away – provided we require the domain descriptions to grow only polynomially and the plans to grow only polynomially in the size of the original plans and domain descriptions." [THN03, p.961] The paper establishes this conclusion by theorems proving for particular classes of planning problems that their plan sizes grow exponentially with such a compilation. They also provide experimental evidence using their FF planner that handling the axioms inside the planner is beneficial for its performance. Other planners handling domain axioms and deduction in certain ways are mentioned in the paper. These experimental experiences confirm the following rather informal meta-consideration. Namely, it is well-known that (on the ground level) deductive reasoning is coNP-complete while planning is even PSPACE-complete (even if only transitions without domain axioms are considered). So reducing the "more costly" transitional part by shifting work into the "simpler" deductive area could perhaps improve the overall performance.

All these arguments established within the PDDL context apply to TL in exactly the same way, since PDDL can obviously be simulated in TL even in a stepwise fashion. But then the question arises what is offered by TL in comparison with the integration of PDDL axioms the way described in [THN03]. The obvious difference is that no deductive relation like our $\vdash_{(T,\leq)}$ is introduced in that paper, relating formulas from *different* states of the world. The deductive relationship is used only *within* a given state thereby reducing the problem to usual deduction. This is made possible because the paper assumes a partition of the predicates into basic and defined ones such that defined predicates must not appear in the initial state nor in any effect part of a transition; "they may only be used in preconditions, effect contexts and goals". Only because of this partition may the deductive relationship be confined to stay within the states which are characterized only by the basic predicates. While this simplifies matters, so that the concept of a source like in the present paper is not needed, it seems to amount to an unrealistic restriction, especially in cases with large sets of rules and transitions needed in common sense reasoning. For instance, in [Sin02] arguments are reported that we have to envisage not only millions but perhaps a hundred millions of rules and transitions for achieving common sense intelligence. It seems illusionary to be able to sort out the deductive relationships of the occurring predicates in such (or even in much smaller) quantities in the neat way required in that paper. Since TL does not feature such a restriction it seems to offer a notable advantage over [THN03].

Of course there is no free lunch also in this case in the sense that determining the source of literals may be costly. Namely, we have seen that the need of handling state sets with a number of alternative states is an intrinsic feature of integrating deduction and transitions (noted also in [Win88]). In realistic applications some of these alternatives might be less likely than others or might be irrelevant under goal-oriented aspects like in Example E_3 in Section 2. But considerations of this kind are beyond the present note.

As the title of the paper indicates TL in the new version presented here has a direct predecessor [Bib98]. Since this is based on the connection method in Automated Deduction, a short introduction to this method seems appropriate.

Consider the following simple deductive example:

$$A(b), A(x) \to B(x), B(x) \to C(x) \vdash C(b)$$

All deductive methods establish such a deductive relationship by generating intermediate results. For instance, resolution would derive $B(b)$ before completion although not requested by the problem. The connection method is unique insofar it completely avoids producing such superfluous intermediate results. Rather it establishes the deductive correctness by a structural analysis of the stated problem in terms of a set of connections and their position in relation with the structure of the formula as illustrated in the following picture.

$$A(b) \wedge [A(x) \to B(x)] \wedge [B(x) \to C(x)] \to C(b)$$

We refer to the literature (eg. [Bib93]) for the details of the structural requirements and here only point out the fact that each connection relates atoms on different sides of an implication and do this in a certain exhaustive way (in technical terms the set is "spanning"). Although this structural analysis provides a deeper insight into the nature of deductive problems and has led to more efficient provers (eg. [OB03]), people seem to prefer intuitively more accessible methods such as resolution.

In [Bib86] the observation was made that in the absense of world knowledge W the transitions can deductively be treated as logical rules but with resources consuming effects: any instance of a literal can be connected only once (so-called strictly linear proofs). With the terminology of the present paper we now understand that any such connection includes a source in the initial state or in the state generated by a transition, which therefore has to be substituted by the results of the transitions. Further, as already pointed out in [Bib98, pp.194], it became clear after the discovery of linear logic [Gir87] that our formalism from that earlier paper, also denoted as linear connection method, coincides with the multiplicative part of linear logic.

Under the impression of linear logic the paper [Bib98] then attempted to integrate transitions within a classical deductive environment by treating them as a multiplicative implication in the sense of linear logic. Technically we extended the linearity restriction to accommodate for the deductively classical parts of a problem (leading to the technical concept of r-compatibility). So we were able to prove formulas like $(P' \to P) \wedge (P \Rightarrow Q) \to (P' \Rightarrow Q)$ which in the terminology of the present paper is the planning problem with $W = \{P' \to P\}$, $T = \{P \Rightarrow Q\}$, initial state P', and goal Q. The technical restrictions, however, were rather unintuitive and left doubts whether they covered the general case correctly; no proof was provided in this regard. The present solution substitutes those attempts, for the case of definite formulas it is provably complete and correct (with straightforward proofs), and it is intuitively simple and convincing. The previous version of transition logic also used multiplicative conjunction (&) and disjunction (|) to account for a resource-oriented treatment of quantities (such as two, rather than one, pieces of an Euro). But such quantities can as well be formalized within classical logic the way presented in textbooks such as [RN03]. Apparently the new version of TL has thus dispensed with the previous more detailed relationship with linear logic.

The fluent calculus [Thi99] (FC) is a derivative of the linear connection method discussed above. But FC inherited some features also from the situation calculus (Sitcalc) to be discussed shortly. It is the first purely logical calculus that has overcome anomalies and deficiences such as those discussed in [Win88] and solved the (technical) frame problem including the inferential one and features many other of the central issues in reasoning about actions and change. Only the treatment of domain axioms as in our paper so far has not been a topic studied in the FC context. One can regard FC as a formalism which models the mechanisms of the dynamic part of TL on the meta-level. This way transitions become represented in logic on the functional term level rather than, as in TL, on the sentential level. For AI applications the distinction is irrelevant, but if we try to stay close enough to human problem solving as a guide towards further improvements, TL (like STRIPS) seems to be a more adequate approach towards reasoning about actions and change.

This last argument is certainly vague. I still regard it as an important one which is supported by a vast number of "experiments" (carried out by daily human use) which are engraved into the structure of natural language over the centuries. Namely, transitions are first-class citizens in natural language where they are treated similarly as logical rules by talking about predicates and their logical relations. An authority in natural language theory describes this similarity as follows [Hob90, p.97].

> Implication can be viewed as a kind of bloodless causality; it plays the role in informational systems that causality plays in physical systems, and it seems likely to me that we understand implication by analogy with causality. That is, the inference S_1 *imply* S_0 is a variety of S_1 *cause* S_0.

Because of the undisputed success of natural language I take the structural similarity as a cognitive argument for TL (which one may or may not accept). Experimental studies in cognitive science, which have demonstrated the importance of predicates (on the sentential level) in human reasoning [Mar99] as well, to some extent support this viewpoint.

Sitcalc is the most widely used formalism for reasoning about actions and change and "has been a part of the artificial intelligence Zeitgeist almost from the very beginning of the field" [Rei01, p.44]. But, in contrast to STRIPS-type planning, it has remained "a theoretical tool without much practical importance" [Rei01, p.xvii]. There is a concrete reason for this contradictory situation. While Sitcalc offers all the power of first-order representation and reasoning and the representational frame problem was solved for it in [Rei91], experiments, written in the FC-based language FLUX and reported in [Thi04, Thi05], demonstrated that it seriously suffers from the inferential frame problem. Namely, many Sitcalc systems compute plans by regression to the initial situation (by way of its successor state axioms) causing a heavy overload in computation while FC as well as TL and STRIPS among others are calculi designed to perform computations in a progressive way, naturally updating the course of changes caused by transitions. It is therefore not clear at all how this kind of Sitcalc systems could handle the agents' reasoning about their "life-long" histories. Of course, Sitcalc can be rephrased to compute plans in a progressive way which indeed has been worked out (see [Rei01]). Also we are of course dealing here with formalisms which can be implemented in a variety of ways with differing successes. So again these arguments are vague and only point to the motivations behind TL rather than disqualifying any of the competing calculi.

Sitcalc also departs from natural language (in the sense discussed just before) because it relates all world knowledge to the situational context by way of an extra parameter. Like the one in relation with FC this cognitive argument also is a vague and computationally irrelevant one. But it still seems awkward to reformalize, say, mathematical laws needed for reasoning about some dynamical system by adding to each predicate this extra parameter. Certainly a mathematician would not be delighted with this unnecessary complication (although the system could of course hide it from him/her).

There is an extensive literature on a related but not identical issue which is ramification and the handling of indirect effects of actions. As an example recall our "stuffy room" example from Section 2 in which the action of opening one window indirectly causes the room to become filled with fresh air. In the interpretation given by this wording in the previous sentence the example does not really fall under the issue discussed in the present paper since we are concerned here with the integration of *static* knowledge, not with causal laws which feature a dynamic nature (we gave the example a different reading in Section 2 though). The handling of causality and ramification within TL is discussed in [Bib98] with a technique applicable also in the new version of TL which keeps static and dynamic knowledge separate. In contrast the causal theories in the literature in some way or another compile static knowledge into causal rules and/or constraints [Thi99, MT97, Lin03]. Given that causal theories should cover engineering which heavily involves mathematical knowledge the present author doubts whether the compilation of lots of mathematics into causal theories is a viable perspective.

There are several more formalisms dealing with dynamical systems. Since they are not directly related with TL we just mention a few of them: production systems [Nil98], disjunctive logic programming formalisms like those in [Bar03, SBD$^+$00], the event calculus [Sha97], propositional dynamic logic [DL95], $\mathcal{C}+$ and the Causal Calculator [GLL$^+$04], modal logic [DHV03], active logics [EDP90, EDKM$^+$99]. A detailed comparison with TL of these and others would go far beyond the scope of this note.

5 Conclusions

In this paper a revised version of transition logic has been presented which overcomes weaknesses of its previous version. TL integrates in a classical logic environment – restricted in our discussions to a function-free definite and stratified clause logic – transitions which change the state of the world. In short, TL can be regarded as classical logic with integrated actions à la STRIPS. Because of this intimate relationship with STRIPS the paper produces at the same time a clarification of the deductive part of original STRIPS which had never been made formally precise.

Given the rich knowledge on classical logic as well as on classical planning including the remarkable performance of the systems based on it, the new formalism provides the core of a potentially broad basis for a variety of applications. Even the ample work on the combination of these two areas within the earlier or related versions of TL (such as the fluent calculus) as well as within the classical planning community (PDDL axioms etc.) can be adapted as described in the previous section. In summary this note implicitly might offer a broad formal platform for reasoning about actions and change in general.

Of course this particular area in AI has many more aspects than addressed in this short note. The most important ones were already discussed in [Bib98, pp.197ff] in the context of the predecessor of the new version of TL; their adaptation has to be worked out but should not be too difficult. Apart from the important issue of causal relationships (already mentioned at the end of the last section) these aspects include (among several others) the issue of qualification. We proposed there to engage transitions in treating nonmonotonic reasoning. That is, if tweety is a bird our world knowledge leads to the conclusion that it flies, unless we hear that it is a penguin in which case a transition (or "belief revision" [BKLW04]) is causally triggered and removes the flying ability. This general idea has recently been developed into a full-fledged theory of nonmonotonic reasoning [Kha02] which fits nicely to the solution presented in this paper.

Altogether, the work presented here leaves ample room for further research. In addition to the extensions just discussed we mention first the generalizations such as partial instead of linear orders in Section 2 or first-order instead of propositional logic in Section 3 which were only outlined and need to be worked out in detail. An obvious further topic would be research towards the extension of the expressiveness of the language of TL. While the language presently envisaged seems expressive enough to cope with many practical problems in a comfortable way, it would still be worthwhile to know how far one could go beyond it in the same vein. For instance, in a common sense reasoning context the need for "elaboration tolerance" [McC98] would probably also require to retract (by some transition) say whole theorems which for whatever reason ceased to hold in the course of events. This short note is also not meant to study the properties of the inference relation introduced here, a task left to future research as well.

Determining the source of an atom is obviously related – but not identical – with abductively determining an explanation of the atom. It would therefore be interesting to see whether the area of abduction would have tools in store for use in this new context. For instance, one would of course like to have preferences among the possible alternatives in the source instead of considering all at the same time. The theoretical basis for this aspect is studied in [PPU03] (where the major references on abduction may be found).

While there are obvious ways to integrate the deductive dimension into the current planning technology, TL opens new algorithmic opportunities for a more involved and optimized integration of the two dimensions. We believe that the work in [Bib86, Bib98] contains a number of relevant indications in this direction which need to be studied under the new viewpoint.

The area of reasoning about action and change has reached a maturity which would allow to organize concerted efforts towards the integration of large knowledge bases (like CYC [Len95]) with first-rate proof and planning systems. A new category of competitions would additionally be required in such a broader context which goes beyond the relatively specialized competitions like the AIPS planning or the automated theorem proving competitions (CASC), in order to foster integrative approaches. Only with such larger scale experiments would we get a better feel how relevant the deductive features would be for the overall performance of such systems. In this sense it is time for AI to start attacking challenging problems of our societies with its technological means, an appeal elaborated by the author in a recent general (ie. non-technical) book [Bib03].

Acknowledgements. I am grateful to Gunter Grieser, Ken Hennacy, Jack Minker, Erik T. Mueller, Don Perlis, Eric Sandewall, Ulrich Scholz, Michael Thielscher, three anonymous referees and the members of the Intellectics group at the TU Darmstadt for their useful comments on earlier versions of the paper which led to a substantial improvement. Especially two of the referees have constructively guided the author from a rather informal presentation to a satisfactory text with precise and correct definitions in embarrassingly many rounds of refereeing. Because of their substantial contributions (including many of the examples now treated in the text) they would well deserve the status of co-authors whereby the text and its possible flaws remain of course the responsibility of the author. I am deeply indebted to them and thank the editor for his patient and fair handling of the case.

References

[Bar03] Chitta Baral. *Knowledge Representation, Reasoning and Declarative Problem Solving*. Cambridge University Press, Cambridge, 2003.

[Bib86] Wolfgang Bibel. A deductive solution for plan generation. *New Generation Computing*, 4:115–132, 1986.

[Bib93] Wolfgang Bibel. *Deduction: Automated Logic*. Academic Press, London, 1993.

[Bib98] Wolfgang Bibel. Let's plan it deductively! *Artificial Intelligence*, 103(1–2):183–208, 1998.

[Bib03] Wolfgang Bibel. *Lehren vom Leben – Essays über Mensch und Gesellschaft*. Sozialwissenschaft. Deutscher Universitäts-Verlag, Wiesbaden, 2003.

[BKLW04] Salem Benferhat, Souhila Kaci, Daniel Le Berre and Mary-Anne Williams. Weakening conflicting information for iterated revision and knowledge integration. *Artificial Intelligence Journal*, 153(1-2):339–371, 2004.

[DHV03] Robert Demolombe, Andreas Herzig and Ivan Varzinczak. Regression in modal logic. Technical Report, IRIT, Université Paul Sabatier, Toulouse, 2003.

[DL95] Giuseppe De Giacomo and Maurizio Lenzerini. PDL-based framework for reasoning about actions. In *Proc. of AI*IA'95*, volume 992 from *LNAI*, pp. 103–114, Berlin, 1995. Springer.

[EDKM+99] Jennifer Elgot-Drapkin, Sarit Kraus, Michael Miller, Madhura Nirkhe and Donald Perlis. Active Logics: A Unified Formal Approach to Episodic Reasoning. Technical Report, University of Maryland, College Park MD, 1999.

[EDP90] J. Elgot-Drapkin and Don Perlis. Reasoning Situated in Time I: Basic Concepts. *J. of Experimental and Theoretical Artificial Intelligence*, 2(1):75–98, 1990.

[FN71] Richard E. Fikes and Nils J. Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence Journal*, 2:189–208, 1971.

[FN93] Richard E. Fikes and Nils J. Nilsson. STRIPS, a retrospective. *Artificial Intelligence Journal*, 59:227–232, 1993.

[Gir87] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987.

[GLL+04] Enrico Giunchiglia, Joohyung Lee, Vladimir Lifschitz, Norman McCain and Hudson Turner. Nonmonotonic Causal Theories. *Artificial Intelligence Journal*, 153(1-2):49–104, 2004.

[Hob90] Jerry R. Hobbs. *Literature and Cognition.* Center for the Study of Language and Information (CSLI), Stanford CA, 1990.

[Kha02] Hesham Khalil. *Logical Foundations of Reactive Default Reasoning.* Akademische Verlagsgesellschaft Aka, Berlin, 2002.

[Len95] Douglas B. Lenat. Cyc: A Large-Scale Investment in Knowledge Infrastructure. *Communications of the ACM*, 38(11):33–38, November 1995.

[Lif86] Vladimir Lifschitz. On the Semantics of STRIPS. In M. P. Georgeff and A. L. Lansky (Eds.), *Reasoning about Actions and Plans*, pp. 1–8, Los Altos, 1986. Morgan Kaufmann.

[Lin03] Fangzhen Lin. Compiling causal theories to successor state axioms and STRIPS-like systems. *Journal of AI Research*, 19:279–314, 2003.

[Llo93] J. W. Lloyd. *Foundations of Logic Programming.* Springer, Berlin, 1993.

[Mar99] Arthur B. Markman. *Knowledge Representation.* Lawrence Erlbaum Associates, Mahwah NJ, 1999.

[McC98] John McCarthy. Elaboration tolerance. In *Working papers of the Fourth International Symposium on Logical Formalizations of Commonsense Reasoning, Common Sense 98*, 1998.

[MT97] Norman McCain and Hudson Turner. Causal theories of action and change. In *Proceedings of the National Conference on Artificial Intelligence*, pp. 460–465, Palo Alto, CA, 1997.

[Nil98] Nils J. Nilsson. *Artificial Intelligence: A New Synthesis.* Morgan Kaufmann, San Francisco, CA, 1998.

[OB03] Jens Otten and Wolfgang Bibel. leanCoP: Lean Connection-Based Theorem Proving. *Journal of Symbolic Computation*, 36:139–161, 2003.

[PPU03] Ramón Pino-Péres and Carlos Uzcátegui. Preferences and explanations. *Artificial Intelligence Journal*, 149(1):1–30, 2003.

[Rei91] Raymond Reiter. The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression. In V. Lifschitz (Ed.), *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pp. 359–380. Academic Press, New York, 1991.

[Rei01] Raymond Reiter. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems.* MIT Press, Cambridge MA, 2001.

[RN03] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall Series in Artificial Intelligence. Pearson Education, Upper Saddle River, N.J., 2nd edition, 2003.

[SBD+00] V.S. Subrahmanian, Piero Bonatti, Jürgen Dix, Thomas Eiter, Sarit Kraus, Fatma Ozcan and Robert Ross. *Heterogeneous Agent Systems.* The MIT Press, Cambridge MA, 2000.

[Sha97] Murray Shanahan. *Solving the Frame Problem – A Mathematical Investigation of the Common Sense Law of Inertia.* Artificial Intelligence. The MIT Press, Cambridge, Massachusetts, 1997.

[Sin02] Push Singh. The Open Mind Common Sense Project. KurzweilAI.net, http://www.kurzweilai.net/articles/art0371.html?printable=1, 2002.

[Thi99] Michael Thielscher. From Situation Calculus to Fluent Calculus: State Update Axioms as a Solution to the Inferential Frame Problem. *AI Journal*, 111(1–2):277–299, 1999.

[Thi04] Michael Thielscher. Logic-Based Agents and the Frame Problem: A Case for Progression. In V. Hendricks, F. Neuhaus, S.A. Pedersen, U. Scheffler and H. Wansing (Eds.), *First-Order Logic Revisited*, Logische Philosophie, pp. 323–336, Berlin, Germany, 2004. Logos.

[Thi05] Michael Thielscher. FLUX: A Logic Programming Method for Reasoning Agents. *Theory and Practice of Logic Programming*, (to appear) 2005.

[THN03] Sylvie Thiébaux, Jörg Hoffmann and Bernhard Nebel. In Defense of PDDL Axioms. In G. Gottlob and T. Walsh (Eds.), *IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, pp. 961–966. Morgan Kaufmann, 2003.

[Win88] Marianne Winslett. Reasoning about action using a possible models approach. In *Proceedings AAAI-88 – Seventh National Conference on Artificial Intelligence*, volume 1, pp. 89–93. AAAI, 1988.