ELSEVIER

Artificial Intelligence 80 (1996) 243–308

Artificial Intelligence

A logic of time, chance, and action for representing plans

Peter Haddawy *

Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, Milwaukee, WI 53201, USA

Received November 1992; revised September 1994

Abstract

This paper integrates logical and probabilistic approaches to the representation of planning problems by developing a first-order logic of time, chance, and action. We start by making explicit and precise commonsense notions about time, chance, and action central to the planning problem. We then develop a logic, the semantics of which incorporates these intuitive properties. The logical language integrates both modal and probabilistic constructs and allows quantification over time points, probability values, and domain individuals. Probability is treated as a sentential operator in the language, so it can be arbitrarily nested and combined with other logical operators. The language can represent the chance that facts hold and events occur at various times. It can represent the chance that actions and other events affect the future. The model of action distinguishes between action feasibility, executability, and effects. We present a proof theory for the logic and show how the logic can be used to describe actions in such a way that the action descriptions can be composed to infer properties of plans via the proof theory.

1. Introduction

Most AI planning systems to date [7, 18, 47, 55] have used variations of situation calculus to represent planning knowledge. This has limited their ability to represent temporal knowledge as well as to deal with uncertainty in the planning domain. Researchers in AI have used temporal logics to formalize the temporal aspects of planning problems [1, 20, 40, 45, 50] and they have developed probability logics for representing uncertainty [3, 24, 26]. But little work has been done toward developing a formal language that integrates these two representational capabilities in the context of planning.

* E-mail: haddawy@cs.uwm.edu.

SSDI 0004-3702(94)00070-0

This paper presents a logic for representing time, chance, and action. We show that by integrating temporal and probabilistic representations in a common formalism we can represent aspects of planning problems that cannot be represented if the two are treated separately. We exploit one of the advantages of using a logical language by building commonsense notions of time, chance, and action directly into the model theory. Thus commonsense inferences follow directly from the semantics of the language. Since such a language can be used to analyze planning problems and prove the correctness of planning algorithms, it represents a significant step toward building planners for more realistic domains. [1]

In the present work, planning is viewed as the process of formulating and choosing a set of actions which when executed will likely achieve a desirable outcome. Actions in a plan may be performed to affect the state of knowledge of the performing agent, to affect the state of the world, or simply for their own sake. The present work focuses on the last two types of action. To choose appropriate courses of such action, an agent must reason about the state of the world, conditions in the world that his actions can influence and the extent to which he can influence them, and, conversely, conditions in the world that influence his actions and the extent of that influence.

Many aspects of the world are inherently stochastic, so a representation for reasoning about plans must be able to express chances of conditions in the world as well as indeterminacy in the effects of actions and events. For example, smoking does not deterministically cause lung cancer; it only greatly increases one's chance of contracting lung cancer. Uncertain environmental factors can influence a smoker's chance of contracting cancer as can uncertainty in the effects of smoking.

Reasoning about plans requires the ability to reason about time. Facts tend to be true for periods of time, and actions and events occur at particular times. Actions comprising a plan may occur sequentially or concurrently. Actions and events affect the future, but not the past. Chance evolves with time: the chance of rain tomorrow may not be the same now as it will be tonight. Ambiguities in the world are resolved with time: before a fair coin is flipped the chance of heads is 50% but after it is flipped it either certainly landed heads or it certainly did not.

This paper presents a first-order logic of time, chance, and action for representing and reasoning about plans. The developed logic represents time in terms of possible world-histories. Possibility is represented in terms Kripke structures [34] by defining an accessibility relation over the world-histories. Chance is represented by defining generalized Kripke structures in terms of probability distributions over the world-histories. By integrating both modal and probabilistic constructs, the logical language can represent and distinguish between possibility, probability, and truth. The language allows quantification over time points, probability values, and domain individuals. Probability is treated as a sentential operator in the language, so it can be arbitrarily nested and combined with other logical operators. The probability operator is temporally indexed so it can capture the dynamic nature of chance.

[1] Haddawy [21] uses the logic presented here to analyze and prove correct components of a construction planning system.

Our model of possibility is similar to that presented by Pelavin [43]. Our model of chance in the context of branching time is based on that of van Fraassen [57]. The probabilistic component of our logic is similar to Halpern's [26] probability logic $\mathcal{L}_2$, to Bacchus' [3] logic of propositional probabilities, and to Haddawy and Frisch's [24] logic of staged probabilities.

1.1. Temporal logic

Temporal logics represent change by specifying what is true in the world at various times. They can be classified as being either interval- or point-based. A point-based logic associates a time point with each temporal object. Most work on plan representation in AI has used interval-based logics which associate a time interval with each temporal object. Through the use of intervals, actions, events, and facts can have temporal extent. This means that these temporal languages can represent plans with concurrent actions as well as conditions during the execution of an action that influence the action. For this reason, the logic of time, chance, and action developed in this paper uses time intervals as well.

Temporal logics can be further classified as being either linear or branching. Linear time logics [1,2] model only the actual world and thus can only represent that an event actually occurs at a given time. In contrast, branching time logics [20,40,45] model all possible worlds and thus can represent whether or not an event can possibly occur, as well as its various possible times of occurrence. Since we are interested in representing chance, which can be roughly thought of as a degree of possibility, we will use a branching time logic. Chance is modeled by defining probability distributions over the branching time structure.

1.2. Probability

Semantic theories of probability may be classified into three main schools: logical [5,32], frequentist [58,59], and subjectivist [37,46,48]. The logical theory takes probabilities to represent a logical relationship between a given hypothesis and given evidence. The frequentist view identifies probability with some suitably defined relative frequency. The subjectivist school takes probability to represent the degrees of belief of a rational agent. Of the three views, the subjective interpretation has the most solid semantic foundation. But while subjective probability theory dictates how degrees of belief should be represented, it does not provide us with guidance for structuring knowledge about the world. Since we are interested in reasoning about the effects of actions on the state of the world, we would like our representation of uncertainty to account for some aspects of probabilistic causality. Subjectivists have proposed subjective theories of objective chance that account for causality by imposing some additional constraints on beliefs [39,53,57]. The current work builds upon van Fraassen's model of objective chance.

But probability theory alone is not representationally adequate for reasoning about plans. Probability theory does not include a notion of quantification. First-order quantification provides great representational economy by allowing us to describe properties

shared by general classes of actions, events, and facts. For example, rather than having to define a different lifting action for each possible object, we can describe the class of lifting actions, where the object being lifted is left as a quantified variable. Furthermore, probability theory provides no vocabulary for describing planning problems [60, p. 570]. We would like a language that facilitates representing the salient features of planning problems. This paper addresses both these limitations.

1.3. Use of the logic

We show in Section 4 that the logic of time, chance, and action is not completely axiomatizable. So we do not foresee building a theorem prover for the logic as a means of solving planning problems. Even if this were possible, it would likely be too inefficient to be useful. Rather, the logic is intended to be used as a tool in the representation and analysis of planning problems involving time and chance. It enables us to design planning algorithms for these problems in a principled manner. Representing the data structures and assumptions of a planner in our logic makes explicit and precise the meaning of the knowledge embodied in the planning algorithm in such a way that we can analyze what is entailed by this knowledge without reference to the planning algorithm itself. We can then refer to this formal semantics to prove the algorithm correct.

1.4. Organization of the paper

Section 2 presents the ontology of the logic. It discusses the desired properties of time, chance, and action at an intuitive level, without recourse to logical formalism. Section 3 formalizes the concepts discussed in the ontology by presenting the syntax and semantics of the logic. Constraints are imposed on the semantic model in order to obtain the desired intuitive properties. Section 4 presents a proof theory for the logic and derives several useful theorems illustrating the use of the proof theory. Numerous axioms and theorems are statements of the intuitive properties discussed in the ontology. Section 5 discusses how the logic can be used to describe properties of actions and plans. We show that the logic captures the natural temporal relation between cause and effect. Section 6 shows how the logic can be used to describe and reason about planning problems. We describe individual components of a planning problem and compose them to reason about the overall problem. Section 7 discusses related work, and Section 8 draws conclusions.

2. Ontology

In this section we seek to sharpen the reader's intuitions concerning those aspects of the world described by the logic: time, facts/events, actions, possibility, and chance.

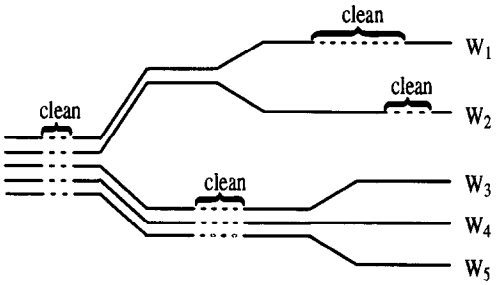

Fig. 1. A temporal tree.

2.1. Time

Time is modeled as a collection of world-histories, each of which is one possible chronology or history of events throughout time. At any given point in time, some of the world-histories will naturally share a common past up to that time. Thus the world-histories form a tree structure that branches into the future. No special status is given to the time "now", so the temporal tree branches into the future relative to each point in time. Fig. 1 shows the structure of a typical temporal tree. The present work is only concerned with future-branching time because actions and events can only affect the state of the world at times after their occurrence. That is to say, at each point in time, the past is fixed—no occurrences in the world will cause it to change. But at each point in time the future might unfold in any number of ways, which are influenced by the agent's actions as well as by other events. So relative to any point in time, only one possible past exists, but numerous possible futures exist.[2]

2.2. Facts and events

The world is described in terms of facts and events. Facts tend to hold and events tend to occur over intervals of time. So facts and events are associated with the time intervals over which they hold or occur in the various world-histories. Facts are distinguished from events on the basis of their temporal properties. A fact may hold over several intervals in any given world-history and if a fact holds over an interval then it holds over all subintervals of that interval. So, for example, my car may be alternately clean and dirty over several different time periods in a given world-history and if my car is clean over a period of time, then it is clean throughout that time. See Fig. 1.

Events are somewhat more complex than facts. First, one must distinguish between *event types* and *event tokens*. An event type is a general class of events and an event token is a specific instance of an event type. For example, a vase breaking is an event type, and the brown vase in my office breaking at 9:00 am is an event token of that type. So event types are sets of event tokens. Event tokens are unique individuals—the interval over which an event token occurs is the unique interval containing the event

[2] For a more thorough discussion of why a future-branching time model is appropriate for representing effects see the article by Mellor [41].

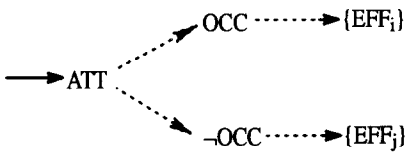

Fig. 2. Action representation.

token and an event token can occur at most once in any world-history. If a vase breaks during a time period, there is no smaller period of time during which the event token of the vase breaking can be said to have occurred. On the other hand, numerous tokens of a given event type may occur during a particular interval. For example, two different vases can break concurrently. So if a token of an event type occurs over an interval, it is possible for another token of that type to occur over a subinterval, but it is not necessary as it is in the case of facts. The present work deals with event types, which for brevity are simply referred to as events.

The fact/event dichotomy just described is a simplification of the true situation. As Shoham [50] has shown, there are many different types of facts and events, characterized by their temporal properties. Although Shoham's refined categories of fact types constitute a more useful and accurate picture of the world than the simple fact/event dichotomy, the fact/event categorization will be used for simplicity of exposition. Extending the work to encompass Shoham's categories is completely straightforward.

2.3. Actions

Actions are similar to events but are distinguished from events by the fact that an action is brought about by an agent. We view the planning problem from the perspective of the planning agent. From this perspective, only the planning agent's own actions are acts; all other actions appear as events. An action is initiated by an *attempt*: the agent attempts an action and, if conditions are right, the action occurs. The occurrence of the action will have certain effects. Likewise, if conditions are such that the attempt of the action does not result in its occurrence, the attempt will have other effects.[3] The situation is depicted in Fig. 2. Once an agent attempts an action, whether or not the action occurs is a function of chance; furthermore, what effects the action's success or failure will have is also a function of chance.

Distinguishing the attempt of an action from its occurrence facilitates several useful inferences. First, it facilitates reasoning about actions as goals. Examples of such goals are going skiing or eating at one's favorite restaurant. In planning for such goals, we are interested in finding circumstances under which the attempt of the action will result in its occurrence.

Second, separating the attempt of an action from its occurrence allows us to distinguish between effects due to the occurrence from effects due to the attempt. For example, lifting an object will result in the agent's holding it but attempting to lift an object that is too heavy may result in the agent straining his back. One advantage of being able to

[3] In the AI literature the effects of the action's occurrence are typically called *intended effects*.

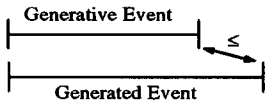

Fig. 3. Relative times of generative and generated events.

make this distinction is that we can distinguish between a plan achieving a goal through its occurrence and a plan achieving a goal as a side-effect of its attempt.

Third, the separation facilitates temporal reasoning about actions. Action occurrences can span time, allowing us to reason about conditions during an action. Furthermore, the amount of time required for an action to occur once it is attempted can be allowed to vary, depending on conditions.

As with events, we distinguish action tokens and action types. Action tokens and action types are analogous to event tokens and event types. Our representation of action tokens is based on Goldman's [19] theory of action individuation in which he defines a generation relation that holds between certain actions. An act token A *generates* an act token A' if it can be said that A' was done by doing A. For example, "flipping the light switch" and "turning on the light" are related by generation. So rather than saying that the two are different descriptions of the same action or that they are two completely different actions, Goldman says that they are two different actions but are related by generation. Generally A and A' will be causally related but other relationships are also possible. Goldman [19, Chapter 2] details four different relationships that he classifies as generation.

We define an *action token* to be composed of a *generative* event token, representing the action's attempt and a *generated* event token, representing the event associated with the successful occurrence of the action. The two event tokens are related by generation: the agent brings about the generated event token by bringing about by the generative event token. Take the example action token of starting my car. This might consist of the generative event token turn-key and the generated event token car-starts. An *action type* is simply a set of action tokens. We will be primarily concerned with action types and will refer to them simply as actions.

We impose constraints on the temporal relation between the generative events and the generated event of an action token. The beginning time of the generative event must coincide with the beginning time of the generated event and the end time of the generative event must not occur after the end time of the generated event. See Fig. 3.

2.4. Possibility

Because the present work is concerned with representing actions that affect the state of the world, the kind of possibility we are interested in describing is objective possibility. Something is objectively possible if it either was true in the past or could be true in the future. So possibility is taken relative to a point in time in the temporal tree. Since actions and events can only affect the future, conditions in the present and past relative to a given time are either inevitably true or inevitably false.

2.5. Chance

Chance is introduced by defining probabilities over the tree of possible futures. Like possibility, chance is taken relative to a given point in time. As a consequence of the property that the past is inevitably true or inevitably false and the fact that inevitability implies probability one, it follows that the chance of the past is either zero or one. In this way, actions and other events can only affect the probabilities of future facts and events. This type of probability is objective, as opposed to subjective. Subjectively the past can be uncertain but objectively it is completely determined. For example, subjectively I may be uncertain as to whether my company's stock rose or fell yesterday, but objectively it either certainly rose or it certainly did not and, furthermore, there is nothing I can do now to change that. The other property imposed on objective chance is that the chance be completely determined by the history up to the current time. So objective chance is purely a function of the state of the world. In contrast, subjective probability is a function not only of the state of the world but also of the epistemic state of an agent.

There is a subtle reason for distinguishing between possibility and chance. It is tempting to think that possibility could just be represented by saying that the probability is non-zero but this is not the case. The problem is that in an uncountably infinite sample space possible events can have probability zero. For example, suppose you pick a real number randomly in the interval [0, 1]. For each number the probability that it will be picked is zero, yet it is possible that it will be picked. The ability to make this distinction will become essential when we discuss action feasibility in Section 5. Possibility and chance are related by the fact that impossibility implies probability zero and inevitability implies probability one.

2.6. Planning

Within the ontological framework just outlined, planning becomes the task of navigating from the present to the future along a temporal tree in an effort to attain a world-history in which the goal condition is satisfied. But an agent has only partial control over its path through the tree: the state of the world is uncertain, action effects are not necessarily deterministic, and external events may influence the path of the agent.

For example, consider the temporal tree shown in Fig. 4. The temperature is below freezing in worlds w_1 and w_2 and above freezing in the rest of the worlds. So there is a 70% chance that the temperature is below freezing. The car is certain to start if the key is turned when the temperature is above freezing but there is only 57% chance that the car will start given that the key is turned when the temperature is below freezing. Thus, by turning the key, the agent can only partially control its path through the tree.

3. The logic of time, chance, and action

To formalize the concepts discussed in the ontology, we now define the language of the logic of time, chance, and action, $\mathcal{L}_{\text{tca}}$. First in order to provide a vocabulary for

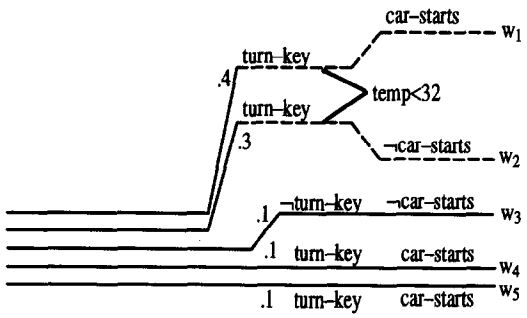

Fig. 4. Navigating the temporal tree.

referring to the elements of our ontology, we specify the syntax of the language. The language of $\mathcal{L}_{\text{tca}}$ is a first-order language with modal operators to express possibility and chance. Since chance is treated as a sentential operator, it can be combined freely with other logical operators. The language allows quantification over probabilities, time points, and domain individuals. The syntax of the probabilistic component of the language is similar to that of Bacchus' [3] logic of propositional probabilities and the syntax of the temporal component of the logic is based on both Shoham's [50] and Pelavin's [43] temporal logics.

3.1. Syntax

The language of $\mathcal{L}_{\text{tca}}$ contains four predicates. To refer to facts and event types occurring in time, the language contains two predicates: $HOLDS(FA, t_1, t_2)$ is true if fact FA holds over the time interval t_1 to t_2, and $OCCURS(EV, t_1, t_2)$ is true if event EV occurs during the interval t_1 to t_2. Henceforth we will use the symbol t, possibly subscripted, to denote time points; ϕ, ψ, and γ to denote formulas; and α and β to denote probability values.

The language contains two predicates to describe actions: $ATT(A, t_A)$ is true if action A is attempted beginning at time t_A, and $OCC(A, t_A, t'_A)$ is true if action A occurs during the interval t_A to t'_A.

In addition to the usual first-order logical operators, the language contains three modal operators to express inevitability, possibility, and chance. The operators are subscripted with a time since, according to the ontology, possibility, inevitability, and chance are taken relative to a point in time. We write $\Box_t(\phi)$ to indicate that ϕ is inevitably true at time t and $\Diamond_t(\phi)$ to say that ϕ is possibly true at time t. We write $P_t(\phi)$ to denote the probability of ϕ at time t. Probability is treated as a sentential operator in the object language. So the probability operator can be arbitrarily nested and combined with the inevitability and possibility operators, allowing us to write complex sentences like:

$$P_{t_1}(\Box_{t_2}\phi \wedge P_{t_3}(\psi) = \alpha) = \beta.$$

Such a treatment is useful for writing sentences about conditional probability. The probability of ϕ given ψ is traditionally defined as

$$prob(\phi \mid \psi) = prob(\phi \wedge \psi)/prob(\psi).$$

If the probability of the conditioning sentence ψ is zero, then the conditional probability is undefined. In this case, a conditional probability sentence like $prob(\phi \mid \psi) = \alpha$ can be assigned neither the value true nor the value false. Rather than introducing a new conditional probability operator and dealing with this truth assignment problem, sentences about conditional probability can simply be written in the form

$$P_t(\phi \wedge \psi) = \alpha \cdot P_t(\psi).$$

Note that this sentence is true for all values of α if $P_t(\psi) = 0$. The standard conditional probability notation will be used to syntactically denote a sentence of the above form:

Definition 1 (*c-prob*).

$$P_t(\phi \mid \psi) = \alpha \equiv P_t(\phi \wedge \psi) = \alpha \cdot P_t(\psi).$$

The language of $\mathcal{L}_{tca}$ contains three types of terms: ordinary domain object terms, temporal terms, and probability terms and four types of relations: numeric relations, fact relations, event relations, and action relations.

The *lexicon* of the language consists of the following disjoint sets of non-logical symbols:

- C a set of object constant symbols;
- TC, a set of time constant symbols;
- NC, a set of numeric constant symbols;
- V, a set of object variables;
- TV, a set of temporal variables;
- PV, a set of probability variables;
- FCT, a set of object function symbols;
- $NFCT$, a set of numeric function symbols, including $+$, $-$, $\cdot$, $/$;
- $PFCT$, a set of probability function symbols for representing distribution functions;
- FR, a set of fact relation symbols;
- ER, a set of event relation symbols;
- AR, a set of action relation symbols; and
- NR, the set $\{<, \leqslant, =, \geqslant, >\}$ of numeric relation symbols.

Note that we will use the symbols $<$, $\leqslant$, $=$, $\geqslant$, $>$ to denote numeric relations in both the object and the meta-language. It will be clear from context which interpretation is meant.

The set of well-formed formulas combining the logical and non-logical symbols is recursively defined as follows.

(1) The set of *object terms* (o-terms) contains members of C, all members of V, as well as all terms of the form $f(otrm_1, \ldots, otrm_n)$, where $f \in FCT$ and $otrm_1, \ldots, otrm_n$ are o-terms.

(2) The set of *temporal terms* (t-terms) contains all members of TC, all members of TV, as well as all terms of the form $f(ttrm_1, \ldots, ttrm_n)$, where $f \in NFCT$ and $ttrm_1, \ldots, ttrm_n$ are t-terms.

(3) The set of *probability terms* (p-terms) contains all members of *NC*, all members of *PV*, all terms of the form $f(trm_1, \ldots, trm_n)$, where $f \in PFCT \cup NFCT$ and $trm_1, \ldots, trm_n$ are t-terms or p-terms, as well as all terms of the form $P_{ttrm}(\phi)$, where ϕ is a wff and *ttrm* is a t-term.

(4) If trm_1 and trm_2 are both t-terms or p-terms then $ttrm_1 < ttrm_2$ is a wff, and similarly for $\leqslant$, $=$, $\geqslant$, $>$.

(5) If $ttrm_1$ and $ttrm_2$ are t-terms, $trm_1, \ldots, trm_n$ are o-terms or t-terms, and *fr* is an *n*-ary fact relation symbol then $HOLDS(fr(trm_1, \ldots, trm_n), ttrm_1, ttrm_2)$ is a wff.

(6) If $ttrm_1$ and $ttrm_2$ are t-terms, $trm_1, \ldots, trm_n$ are o-terms or t-terms, and *er* is an *n*-ary event relation symbol then $OCCURS(er(trm_1, \ldots, trm_n), ttrm_1, ttrm_2)$ is a wff.

(7) If $ttrm_1$ is a t-term, $trm_1, \ldots, trm_n$ are o-terms or t-terms, and *ar* is an *n*-ary action relation symbol then $ATT(ar(trm_1, \ldots, trm_n), ttrm_1)$ is a wff.

(8) If $ttrm_1$ and $ttrm_2$ are t-terms, $trm_1, \ldots, trm_n$ are o-terms or t-terms, and *ar* is an *n*-ary action relation symbol then $OCC(ar(trm_1, \ldots, trm_n), ttrm_1, ttrm_2)$ is a wff.

(9) If $otrm_1$ and $otrm_2$ are o-terms then $otrm_1 = otrm_2$ is a wff.

(10) If ϕ_1 and ϕ_2 are wffs then so are $\neg\phi_1$, $\phi_1 \wedge \phi_2$, $\phi_1 \vee \phi_2$, and $\phi_1 \rightarrow \phi_2$.

(11) If ϕ is a wff and $z \in V \cup TV \cup PV$ then $\forall z\phi$ and $\exists z\phi$ are wffs.

(12) If ϕ is a wff and *ttrm* is a t-term then $\Box_{ttrm}(\phi)$ and $\Diamond_{ttrm}(\phi)$ are wffs.

We use the symbol = to denote equality between object terms, probability terms, and temporal terms. The meaning will be clear from the context.

Notice that the syntax of the language is restricted to disallow sentences that would be meaningless in the intended interpretation of the language. For example the following sentence is not well-formed because it does not make sense for a probability term to be used as a time point:

$$HOLDS(ON(A, B), t_1, P_{t_1}(\phi)).$$

3.1.1. Examples of representing planning knowledge

Now that we have defined the syntax of the language, we can see how the language can represent various types of planning knowledge. The language allows us to write sentences that describe:

- Uncertainty in the state of the world: there is a 60% chance of a power glitch between noon and 5:00.

$$P_{now}(\exists t_1, t_2\ (noon \leqslant t_1 \leqslant t_2 \leqslant 5{:}00) \wedge OCCURS(\text{power-glitch}, t_1, t_2)) = 0.6.$$

- Uncertainty of action effects: there is a 50% chance that the coin will land heads when flipped.

$$\forall t, t_1, t_2\ (t \leqslant t_1) \rightarrow P_t(\exists t_3\ OCCURS(\text{land(coin,heads)}, t_2, t_3) \mid OCC(\text{flip(coin)}, t_1, t_2)) = 0.5.$$

- Conditions during an action that influence its effects: holding the oven temperature constant increases the likelihood that the souffle will turn out right.

$$\forall t, t_1, t_2\ (t \leqslant t_1) \rightarrow P_t(\exists t_3\ \mathit{HOLDS}(\text{done-right}(\text{souffle}), t_2, t_3) \mid \mathit{OCCURS}(\text{bake}(\text{souffle}), t_1, t_2) \wedge \exists x\ \mathit{HOLDS}(\text{temp}(\text{oven},x), t_1, t_2)) > P_t(\exists t_3\ \mathit{HOLDS}(\text{done-right}(\text{souffle}), t_2, t_3) \mid \mathit{OCCURS}(\text{bake}(\text{souffle}), t_1, t_2)).$$

- Conditions not influenced by an action: the chance of rain is independent of my clapping my hands.

$$\forall t, t_1, t_2, t_3, t_4\ (t_1 \leqslant t_3) \rightarrow P_t(\mathit{HOLDS}(\text{raining}, t_3, t_4) \mid \mathit{OCC}(\text{clap}, t_1, t_2)) = P_t(\mathit{HOLDS}(\text{raining}, t_3, t_4)).$$

- Concurrent actions: it is not possible for me to raise and lower my arm at the same time.

$$\forall t, t_1, t_2, t_3, t_4\ \Box_t\ [\mathit{OCC}(\text{raise}(\text{arm}), t_1, t_2) \wedge \mathit{OCC}(\text{lower}(\text{arm}), t_3, t_4) \rightarrow (t_2 < t_3) \vee (t_4 < t_1)].$$

- External events: there is a 90% chance that the computer will crash if a power glitch occurs.

$$\forall t_1, t_2\ P_{now}(\exists t_3, t_4\ (t_1 < t_3 \leqslant t_2) \wedge \mathit{OCCURS}(\text{crash}(\text{computer}), t_3, t_4) \mid \mathit{OCCURS}(\text{power-glitch}, t_1, t_2)) = 0.9.$$

- Temporally qualified goals: be at the bank before 5:00pm.

$$\exists t_1, t_2\ (t_1 < 5{:}00) \wedge \mathit{HOLDS}(\text{loc}(\text{me},\text{bank}), t_1, t_2).$$

Furthermore, the language allows us to write sentences that

- Combine possibility and chance: there is a 50% chance that by noon a train crash will inevitably occur between 3:00 and 5:00.

$$P_{now}(\exists t_1\ (t_1 \leqslant \mathit{noon}) \wedge \Box_{t_1}(\exists t_2, t_3\ (3{:}00 \leqslant t_2 \leqslant t_3 \leqslant 5{:}00) \wedge \mathit{OCCURS}(\text{crash}(\text{train}), t_2, t_3))) = 0.5.$$

- Distinguish between truth and chance: I won the lottery even though it was unlikely.

$$(t_0 < t_1 < t_2 < \mathit{now}) \wedge P_{t_0}(\mathit{OCCURS}(\text{win-lottery}, t_1, t_2)) = 0.0001 \wedge \mathit{OCCURS}(\text{win-lottery}, t_1, t_2).$$

- Express information about probability distributions: the arrival time of the train is normally distributed about noon.

$$\forall t\ P_{now}(\exists t'\ \mathit{OCCURS}(\text{arrive}, t, t')) = N(t, \text{noon}, 10\text{min}),$$

where $N(t, \text{noon}, 10\text{min})$ is a normal distribution over the variable t with mean noon and variance 10min.

3.2. Semantics

So far we have only specified a formal way of writing down sentences in the language of $\mathcal{L}_{\text{tca}}$. We need some way of assigning meaning to these sentences. This is done through the use of model-theoretic semantics. The elements of the model correspond closely to the elements identified in the ontology. The models contain a set of possible world-histories as well as sets of facts, events, and actions. The temporal tree structure is created by defining an accessibility relation over the world-histories. Possibility is then defined with respect to this accessibility relation. Chance is defined in terms of probability distributions over the temporal tree. The desired properties identified in the ontology are obtained by placing a number of semantic constraints on the models. A model is a tuple

$$\langle W, D, FN, NFN, PFN, FRL, ERL, ARL, NRL, FA, EVENTS, EV, ACTS, ACTIONS, R, \mathcal{X}, PR, F\rangle,$$

where:

- W is the set of possible world-histories, called worlds.
- D is the non-empty domain of individuals.
- FN is the set of object functions: $D^k \rightarrow D$.
- NFN is the set of numeric functions: $\mathbb{R}^k \rightarrow \mathbb{R}$.
- PFN is the set of probability functions: $\mathbb{R}^k \rightarrow \mathbb{R}$.
- FRL is the set of fact relations: $D^k \rightarrow 2^{(\mathbb{R}\times\mathbb{R})\times W}$.
- ERL is the set of event relations: $D^k \rightarrow 2^{(\mathbb{R}\times\mathbb{R})\times W}$.
- ARL is the set of action relations: $D^k \rightarrow 2^{ACTS}$.
- NRL is the set of numeric relations, a subset of $2^{\mathbb{R}^k}$.
- FA is the set of facts, a subset of $2^{(\mathbb{R}\times\mathbb{R})\times W}$. A fact is a set of ⟨temporal interval, world⟩ pairs: $\{\langle\langle t_1, t_1'\rangle, w_1\rangle, \ldots, \langle\langle t_n, t_n'\rangle, w_n\rangle\}$. If fa is a fact and $\langle\langle t_1, t_2\rangle, w\rangle \in fa$ then fa holds throughout interval $\langle t_1, t_2\rangle$ in world-history w.
- $EVENTS$ is the set of event tokens, a subset of $(\mathbb{R} \times \mathbb{R}) \times W$. An event token is a single ⟨temporal interval, world⟩ pair.
- EV is the set of event types, a subset of 2^{EVENTS}. An event type is a set of event tokens: $\{\langle\langle t_1, t_1'\rangle, w_1\rangle, \ldots, \langle\langle t_n, t_n'\rangle, w_n\rangle\}$. If ev is an event and $\langle\langle t_1, t_2\rangle, w\rangle \in ev$ then ev occurs during interval $\langle t_1, t_2\rangle$ in world-history w.
- $ACTS$ is the set of action tokens, a subset of $2^{EV\times EV}$. An action token is an ordered pair consisting of a generative event token and a generated event token: $\langle gev, Gev\rangle$. We will find it useful to define two functions in the meta-language in order to pick out the generative and generated event tokens: $gen(act)$ is the generative event token of act and $Gen(act)$ is the generated event token of act.
- $ACTIONS$ is the set of action types, a subset of 2^{ACTS}. An action type is a set of action tokens: $\{\langle gev_1, Gev_1\rangle, \langle gev_2, Gev_2\rangle, \ldots\}$. For example, the start-car action might be represented as $\{\langle\langle\text{turn-key}, t_1, t_2\rangle, \langle\text{car-starts}, t_1, t_3\rangle\rangle, \ldots\}$. Note that

the action type, event type, and fact corresponding to the empty set denote the impossible action, event, and fact, respectively.

- R is an accessibility relation defined on $\mathbb{R} \times W \times W$. $R(t, w_1, w_2)$ means that world-histories w_1 and w_2 are indistinguishable up to and including time t. Making worlds indistinguishable through time t disallows instantaneous effects, i.e. there must be some time between the occurrence of an event and its effects. If $R(t, w_1, w_2)$ we say a world-history w_2 is R-accessible from w_1 at time t. The set of all world-histories R-accessible from w at time t will be designated R_t^w. For each time t, the R_t^w partition the world-histories into sets of equivalence classes indistinguishable up to t.
- $\mathcal{X}$ is a σ-algebra over W,[4] containing all the sets corresponding to wffs in the language, as well as all R-equivalence classes of world-histories.
- PR is a probability assignment function that assigns to each time $t \in \mathbb{R}$ and world-history $w \in W$ a countably additive probability distribution μ_t^w defined over $\mathcal{X}$.
- F is the denotation function, defined as follows:

$$
\begin{aligned}
&C \rightarrow D,\\
&NC \rightarrow \mathbb{R},\\
&TC \rightarrow \mathbb{R},\\
&FCT \rightarrow FN,\\
&NFCT \rightarrow NFN,\\
&PFCT \rightarrow PFN,\\
&FR \rightarrow FRL,\\
&ER \rightarrow ERL,\\
&NR \rightarrow NRL.
\end{aligned}
$$

Henceforth, M will be used to refer a model with the eighteen components named above.

The reader will note that constants, functions, and relations are rigid with respect to both time and world-history. Rigidity with respect to world-history simplifies the logic somewhat and could be relaxed without much effort. Non-rigidity with respect to time seems to be basically incompatible with the current framework since it would result in undesirable interpretations like the following. At time t_1 a term like Blue(car) could denote the fact that my car is blue from t_0 to t_3 and at time t_2 it could denote the fact that my car is blue from t_4 to t_5.

3.2.1. *Semantic constraints*

In Section 2 the ontology of the logic was discussed from an intuitive standpoint. In order to obtain the desired intuitive properties, a number of constraints must be imposed on the models. These constraints, labeled (C1)–(C8) are presented in the following discussion.

[4] A σ-algebra over W is a class of subsets that contains W and is closed under complement and countable union.

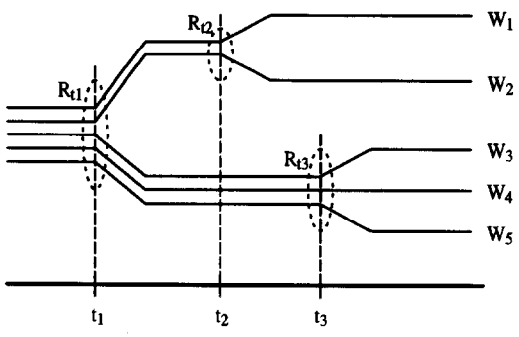

Fig. 5. Structure imposed by the R-accessibility relation.

The future-branching temporal tree is defined in terms of the R-relation over world-histories. To capture the property that time does not branch into the past, we say that if two world-histories are indistinguishable up to time t_2 then they are indistinguishable up to any earlier time:

(C1) If $t_1 \leqslant t_2$ and $R(t_2, w_1, w_2)$ then $R(t_1, w_1, w_2)$.

Since R just represents the indistinguishability of histories up to a time t, for a fixed time R is an equivalence relation, i.e., reflexive, symmetric, and transitive:

(C2) $R(t, w, w)$.
If $R(t, w_1, w_2)$ then $R(t, w_2, w_1)$.
If $R(t, w_1, w_2)$ and $R(t, w_2, w_3)$ then $R(t, w_1, w_3)$.

Fig. 5 illustrates that the R-relation ties together the different world-histories to form the temporal tree structure discussed in Section 2.

As mentioned earlier, facts and events differ in their temporal properties. This distinction is captured by the following two semantic constraints. If a fact holds over an interval, it holds over all subintervals, except possibly at the endpoints:

(C3) If $t_1 \leqslant t_2 \leqslant t_3 \leqslant t_4$, $t_1 \neq t_3$, $t_2 \neq t_4$, $fa \in FA$ and $\langle\langle t_1, t_4\rangle, w\rangle \in fa$ then $\langle\langle t_2, t_3\rangle, w\rangle \in fa$.

An event token occurs only once in each world-history:

(C4) If $evt \in EVENTS$, $\langle\langle t_1, t_2\rangle, w\rangle \in evt$, and $\langle\langle t_3, t_4\rangle, w\rangle \in evt$ then $t_1 = t_3$ and $t_2 = t_4$.

If two worlds are indistinguishable up to a time then they must share a common past up to that time. And if they share a common past up to a given time, they must agree on all facts and events up to that time. To enforce this relationship, we impose the constraint that if two world-histories are R-accessible at time t, they must agree on all facts (events) that hold (occur) over intervals ending before or at the same time as t:

(C5) If $t_0 \leqslant t_1 \leqslant t_2$ and $R(t_2, w_1, w_2)$ then $\langle\langle t_0, t_1\rangle, w_1\rangle \in A$ iff $\langle\langle t_0, t_1\rangle, w_2\rangle \in A$, where A is a fact or event.

Section 2 mentions two desired characteristics of the probability operator. The first is that the probability at a time t be completely determined by the history up to that time. The second desired characteristic is that the probability of the present and past should be either zero or one. These two properties follow as meta-theorems from the following two constraints:

(C6) For all $X \in \mathcal{X}$, $t \leqslant t'$, and w, w' such that $R(t, w, w')$, $\mu_t^w(R_t^{w'}) > 0 \rightarrow \mu_{t'}^{w'}(X) = \mu_t^w(X \mid R_{t'}^{w'})$.

(C7) $\mu_t^w(R_t^w) > 0$.

Meta-Theorem 2. *The probability of the present and past is either zero or one.*

$\mu_t^w(R_t^w) = 1$.

Proof.

1.	$\mu_t^w(R_t^w) > 0$	(C7)
2.	$\mu_t^w(R_t^w) = \mu_t^w(R_t^w \mid R_t^w)$	Modus Ponens: (C6), 1
3.	$\mu_t^w(R_t^w) = 1$	def of conditional probability.

Defining the probabilities in this way makes good intuitive sense if we look at the meaning of R. R_t^w designates the set of world-histories that are objectively possible with respect to w at time t. It is natural that the set of world-histories that are objectively likely with respect to w at time t should be a subset of the ones that are possible.

Theorem 3. *If two worlds are indistinguishable up to time t then they have identical probability distributions at that time.*

If $R(t, w, w')$ *then* $\mu_t^{w'}(X) = \mu_t^w(X)$

Proof.

1.	$\mu_t^w(R_t^{w'}) > 0$	(C2), (C7)
2.	$\mu_t^w(R_t^{w'}) = \mu_t^w(X \mid R_t^{w'})$	Modus Ponens: (C6), 1
3.	$\mu_t^w(X \mid R_t^{w'}) = \mu_t^w(X \mid R_t^w)$	(C2)
4.	$\mu_t^w(R_t^w) = 1$	Meta-Theorem 2
5.	$\mu_t^{w'}(X) = \mu_t^w(X)$	definition of conditional probability.

The following constraint on actions performs two functions. First, it enforces the desired temporal relation between generative and generated event tokens. Second, it guarantees that actions can actually occur, by requiring that the generative and generated event tokens of an action token occur in the same world-history.

(C8) If $act \in ACTS$ and $\langle\langle t_1, t_2\rangle, w_1\rangle = gen(act)$ and $\langle\langle t_3, t_4\rangle, w_2\rangle = Gen(act)$ then $t_1 = t_3$, $t_2 \leqslant t_4$, and $w_1 = w_2$.

3.2.2. Semantic definitions

Given the models described above, the semantic definitions for the well-formed formulas can now be defined. Denotations are assigned to expressions relative to a model, a world-history within the model, and an assignment of individuals in the domain to variables. The denotation of an expression ϕ relative to a model M and a world-history w, and a variable assignment g is designated by $[\![\phi]\!]^{M,w,g}$. The variable assignment function g maps each temporal and probability variable to a real number and each object variable to a domain individual. In the definitions below, the expression $g[d/z]$ denotes the assignment of values to variables that is identical to assignment g with the possible exception that element d is assigned to variable z. The semantic definitions for the well-formed formulas are given below. The definitions specify the conditions under which sentences are assigned the value true. Since we have a two-valued logic, if a sentence is not true it is false.

(1) If u is a variable then $[\![u]\!]^{M,w,g} = g(u)$.

(2) If α is a non-logical constant then $[\![\alpha]\!]^{M,w,g} = F(\alpha)$.

(3) If $\tau = f(trm_1, \ldots, trm_n)$ is a o-term, t-term, or p-term then

$$[\![\tau]\!]^{M,w,g} = [\![f]\!]^{M,w,g}([\![trm_1]\!]^{M,w,g}, \ldots, [\![trm_n]\!]^{M,w,g}).$$

(4) $[\![trm_1 < trm_2]\!]^{M,w,g}$ = true iff $[\![trm_1]\!]^{M,w,g} < [\![trm_2]\!]^{M,w,g}$.

(5) $[\![trm_1 = trm_2]\!]^{M,w,g}$ = true iff $[\![trm_1]\!]^{M,w,g} = [\![trm_2]\!]^{M,w,g}$.

(6) $[\![HOLDS(rf(trm_1, \ldots, trm_n), ttrm_1, ttrm_2)]\!]^{M,w,g}$ = true iff

$$\langle\langle [\![ttrm_1]\!]^{M,w,g}, [\![ttrm_2]\!]^{M,w,g}\rangle, w\rangle \in F(rf)([\![trm_1]\!]^{M,w,g}, \ldots, [\![trm_n]\!]^{M,w,g}).$$

(7) $[\![OCCURS(re(trm_1, \ldots, trm_n), ttrm_1, ttrm_2)]\!]^{M,w,g}$ = true iff

$$\langle\langle [\![ttrm_1]\!]^{M,w,g}, [\![ttrm_2]\!]^{M,w,g}\rangle, w\rangle \in e$$
$$\text{for some } e \in F(re)([\![trm_1]\!]^{M,w,g}, \ldots, [\![trm_n]\!]^{M,w,g}).$$

(8) $[\![ATT(ra(trm_1, \ldots, trm_n), ttrm_1)]\!]^{M,w,g}$ = true iff

$$\exists act \in F(ra)([\![trm_1]\!]^{M,w,g}, \ldots, [\![trm_n]\!]^{M,w,g})$$
$$\text{such that } \exists t\, gen(act) = \langle\langle [\![ttrm1]\!]^{M,w,g}, t\rangle, w\rangle.$$

(9) $[\![OCC(ra(trm_1, \ldots, trm_n), ttrm_1, ttrm_2)]\!]^{M,w,g}$ = true iff

$$\exists act \in F(ra)([\![trm_1]\!]^{M,w,g}, \ldots, [\![trm_n]\!]^{M,w,g})$$
$$\text{such that } \exists t\, gen(act) = \langle\langle [\![ttrm1]\!]^{M,w,g}, t\rangle, w\rangle$$
$$\text{and } Gen(act) = \langle\langle [\![ttrm_1]\!]^{M,w,g}, [\![ttrm_2]\!]^{M,w,g}\rangle, w\rangle.$$

(10) $[\![\neg\phi]\!]^{M,w,g}$ = true iff $[\![\phi]\!]^{M,w,g} \neq$ true.

(11) $[\![\phi_1 \wedge \phi_2]\!]^{M,w,g}$ = true iff $[\![\phi_1]\!]^{M,w,g}$ = true and $[\![\phi_2]\!]^{M,w,g}$ = true.

(12) $[\![\forall z\, \phi]\!]^{M,w,g}$ = true iff $z \in V$ and $[\![\phi]\!]^{M,w,g[d/z]}$ = true for all $d \in D$ or $z \in TV \cup PV$ and $[\![\phi]\!]^{M,w,g[d/z]}$ = true for all $d \in \mathbb{R}$.

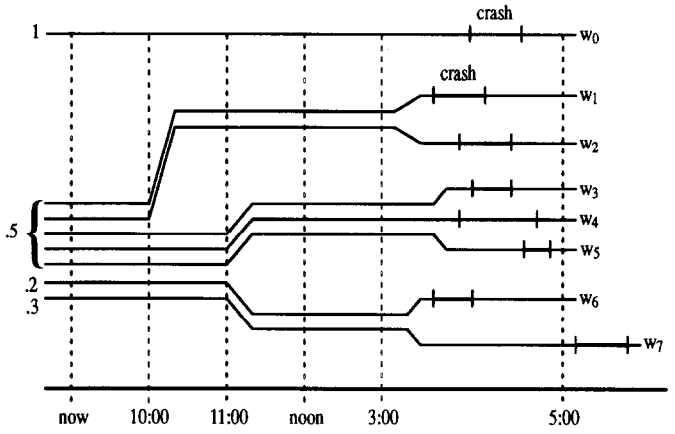

Fig. 6. Model for train crash example.

(13) $[\![\Box_{ttrm}(\phi)]\!]^{M,w,g}$ = true iff $[\![\phi]\!]^{M,w',g}$ = true for every w' such that

$$R([\![ttrm]\!]^{M,w,g}, w, w').$$

(14) $[\![P_{ttrm}(\phi)]\!]^{M,w,g} = \mu^{w}_{[\![ttrm]\!]^{M,w,g}}(\{w' \in R^{w}_{[\![ttrm]\!]^{M,w,g}}: [\![\phi]\!]^{M,w',g} = \text{true}\})$.

The logical operators $\vee$, $\rightarrow$, and $\exists$ are defined in terms of $\neg$, $\wedge$, and $\forall$ in the usual way. The numeric relations $\leqslant$, $=$, $\geqslant$, $>$ are defined in terms of $<$, e.g.,

$$(A \leqslant B) \equiv \neg(B < A).$$

Possibility $\Diamond$ is defined in terms of inevitability as $\Diamond_t(\phi) \equiv \neg\Box_t(\neg\phi)$. The interesting definitions are the last two. Definition (13) says that a sentence is inevitable in a world w at a time t iff it is true in all worlds indistinguishable from w up to and including time t. Definition (14) says that the probability of a sentence ϕ in a world w at a time t is the probability of those accessible worlds in which ϕ is true.

A sentence ϕ is *satisfied* by a model M at a world w if it is assigned the value true by that model and world. A sentence is *valid* if it is satisfied by every model at every world. The value assignment g is irrelevant to the truth value of a sentence since by definition sentences contain no free variables.

3.3. Example

To illustrate the correspondence between the syntax and semantics of $\mathcal{L}_{\text{tca}}$ we show a possible model for one of the example sentences presented earlier:

- There is a 50% chance that by noon a train crash will inevitably occur between 3:00 and 5:00.

$$P_{now}(\exists t_1\ (t_1 \leqslant noon) \wedge \Box_{t_1}(\exists t_2, t_3\ (3{:}00 \leqslant t_2 \leqslant t_3 \leqslant 5{:}00) \wedge OCCURS(\text{crash}(\text{train}), t_2, t_3))) = 0.5.$$

One possible model for this sentence is shown in Fig. 6. The sentence is not satisfied in world w_0 since there the chance that a crash between 3:00 and 5:00 is inevitable by noon is one. The sentence is satisfied in worlds w_1–w_7. In worlds w_1 and w_2 a crash between 3:00 and 5:00 is inevitable at 10:00am. In each of the worlds w_3–w_5 a crash between 3:00 and 5:00 is inevitable at 11:00am. The chance of worlds w_1–w_5 is 50%. In worlds w_6 and w_7 a crash is inevitable but it does not necessarily occur between 3:00 and 5:00. In worlds w_1–w_7 there is a 70% chance of a crash occurring between 3:00 and 5:00. This is higher than the chance of a crash *inevitably* occurring between 3:00 and 5:00. Notice that it follows from the above logical sentence that there is at least a 50% chance an agent controlling the train can do nothing to avoid a crash after noon.

4. Proof theory

In this section we present a partial axiomatization for $\mathcal{L}_{\text{tca}}$. We then use the axiomatization to prove numerous theorems that illustrate properties of the logic and that will be useful later in reasoning about actions and plans. In addition to facilitating reasoning with the logic, the axioms and theorems presented in this section show that the constraints imposed on the models in Section 3 are sufficient to capture the desired intuitive properties described in the ontology.

Before presenting the axioms we first note that $\mathcal{L}_{\text{tca}}$ is not completely axiomatizable. This fact follows from a result due to Halpern [26]. Halpern presents a first-order probability logic $\mathcal{L}_2$ that allows quantification over real probability values and domain individuals. The language of $\mathcal{L}_2$ is that of ordinary first-order logic with the addition of the probability operator P. Probability in models for $\mathcal{L}_2$ is defined in terms of a single distribution over possible worlds. Halpern shows that the valid formulas of $\mathcal{L}_2$ are not recursively enumerable and hence the logic is not completely axiomatizable. The argument that this entails the non-axiomatizability of $\mathcal{L}_{\text{tca}}$ is given below.

Sentences of $\mathcal{L}_2$ can be translated to sentences of $\mathcal{L}_{\text{tca}}$ as follows. Halpern's P operator is translated as a probability at a given fixed time, say t_0: P_{t_0}. All atomic formulas are translated as instantaneous facts at that same fixed time, e.g., $\text{on}(A,B)$ becomes $HOLDS(\text{on}(A,B), t_0, t_0)$. So, for example, the $\mathcal{L}_2$ sentence

$$\forall x\, P(\exists y\, Q(x,y)\ \wedge\ P(\forall z\, R(z)) = 0.7) = 0.2$$

would be translated as

$$\forall x\, P_{t_0}(\exists y\, HOLDS(Q(x,y), t_0, t_0)\ \wedge\ P_{t_0}(\forall z\, HOLDS(R(z), t_0, t_0)) = 0.7) = 0.2.$$

Now an $\mathcal{L}_{\text{tca}}$ sentence which is a translation of an $\mathcal{L}_2$ sentence is valid in our logic if and only if the corresponding $\mathcal{L}_2$ sentence is valid in Halpern's logic. So if the valid formulas of $\mathcal{L}_{\text{tca}}$ were recursively enumerable so would be the valid formulas of $\mathcal{L}_2$. Since the valid formulas of $\mathcal{L}_2$ are not recursively enumerable, the valid formulas of $\mathcal{L}_{\text{tca}}$ cannot be either. Hence $\mathcal{L}_{\text{tca}}$ is not completely axiomatizable.

The lack of existence of a complete axiomatization for $\mathcal{L}_{tca}$ is not seen as a serious drawback for two reasons. First, the logic is intended to be used in the design and

analysis of planning algorithms that are sound with respect to the logic and possibly complete with respect to a subset of the logic. Second, a sound set of axioms and rules of inference are provided that are rich enough to allow us to make all inferences considered in this paper.

4.1. Axioms

The axioms presented below are divided into classes. First there are six basic classes describing six types of reasoning:

(1) first-order axioms with equality and the rules of Modus Ponens and Universal Generalization to describe first-order logical reasoning,
(2) axioms of real closed fields to describe numeric reasoning,
(3) S5 axioms and the rule of necessitation for reasoning about inevitability,
(4) probability axioms and the rule of probability of logical equivalents to describe probabilistic reasoning,
(5) temporal logic axioms to describe temporal reasoning, and
(6) an action axiom for reasoning about action.

Following these are three sets of axioms that relate some of the basic classes:

- the inevitability and equality axiom,
- axioms describing the temporal properties of inevitability, and
- axioms relating inevitability and probability.

Subsets of the axioms have appeared elsewhere in the literature as parts of axiomatizations of other logics. The first-order axioms and rules of inference are taken from [30, Chapter 1] and appear originally in [63]. The S5 axioms and rule of necessitation are taken from [30, Chapter 3]. The S5 axioms, the rule of necessitation, axiom IE, and axioms IT1–IT4 are part of the axiomatization of Pelavin's [43] planning logic. The field axioms and the probability axioms P1, P2, and PE have appeared as parts of the axiomatization of the probability logics in [17] and [3]. The field axioms appear originally in [49]. Soundness proofs for the less commonly known axioms are given in the appendix.

First-order axioms with equality

FOL1. $(\phi \vee \phi) \rightarrow \phi$.

FOL2. $\psi \rightarrow (\phi \vee \psi)$.

FOL3. $(\phi \vee \psi) \rightarrow (\psi \vee \phi)$.

FOL4. $(\psi \rightarrow \gamma) \rightarrow ((\phi \vee \psi) \rightarrow (\phi \vee \gamma))$.

FOL5. $\forall x\phi \rightarrow \phi(x/trm)$, where trm is substitutable[5] for x in ϕ.

[5] Roughly, trm is substitutable for x in ϕ if ϕ does not contain a quantifier that could capture trm.

EQ1. $trm = trm$, where trm is any term.

EQ2. $(trm_1 = trm_2) \rightarrow (\phi \rightarrow \phi')$, where ϕ is atomic and ϕ' is obtained by replacing trm_1 in zero or more places by trm_2.

First-order rules of inference

MoPo. Modus Ponens: From ϕ and $\phi \rightarrow \psi$ infer ψ.

UG. Universal Generalization: From $\phi \rightarrow \psi$ infer $\phi \rightarrow \forall x\psi$, where x does not occur free in ϕ.

Axioms of real closed fields

These axioms capture numeric reasoning over the reals. The variables range over time points and probability values.

F1. $\forall xyz\ ((x + y) + z = x + (y + z))$.

F2. $\forall x(x + 0 = x)$.

F3. $\forall x\ (x + (-1 \cdot x) = 0)$.

F4. $\forall xy\ (x + y = y + x)$.

F5. $\forall xyz\ ((x \cdot y) \cdot z = x \cdot (y \cdot z))$.

F6. $\forall x\ (x \cdot 1 = x)$.

F7. $\forall x\ (x \neq 0 \rightarrow \exists y(x \cdot y = 1))$.

F8. $\forall xy\ (x \cdot y = y \cdot x)$.

F9. $\forall xyz\ (x \cdot (y + z) = (x \cdot y) + (x \cdot z))$.

F10. $0 \neq 1$.

F11. $\forall x\ (\neg(x < x))$.

F12. $\forall xyz\ ((x < y) \wedge (y < z) \rightarrow (x < z))$.

F13. $\forall xy\ ((x < y) \vee (x = y) \vee (y < x))$.

F14. $\forall xyz\ ((x < y) \rightarrow ((x + z) < (y + z)))$.

F15. $\forall xy\ (((0 < x) \wedge (0 < y)) \rightarrow (0 < x \cdot y))$.

F16. $\forall x\ ((0 < x) \rightarrow \exists y(y \cdot y = x))$.

F17. Every polynomial of odd degree has a root, e.g., the axiom for a polynomial of degree 3 is

$$\forall y_0 y_1 y_2 y_3\ ((y_0 \neq 0) \rightarrow \exists x\ (y_0 \cdot x \cdot x \cdot x + y_1 \cdot x \cdot x + y_2 \cdot x + y_3 = 0)).$$

S5 axioms

These axioms capture the fact that for a fixed time, inevitability is an S5 type modal operator.

I1. $\Box_t \phi \rightarrow \phi$.

I2. $\Box_t(\phi \rightarrow \psi) \rightarrow (\Box_t \phi \rightarrow \Box_t \psi)$.

I3. $\Box_t \phi \rightarrow \Box_t \Box_t \phi$.

I4. $\Diamond_t \phi \rightarrow \Box_t \Diamond_t \phi$.

Rule of necessitation

NEC. From ϕ conclude $\Box_t \phi$.

Probability axioms

These axioms capture the probabilistic component of the logic. The first two axioms describe probability at a fixed time, while the third axiom describes the behavior of probability over time.

P1. Non-negativity: $P_t(\phi) \geqslant 0$.

P2. Additivity: $P_t(\phi) = P_t(\phi \wedge \psi) + P_t(\phi \wedge \neg\psi)$.

P3. Miller's principle: $(t_1 \leqslant t_2) \rightarrow P_{t_1}(\phi \mid P_{t_2}(\phi) \geqslant \alpha) \geqslant \alpha$.

Axioms P1 and P2 are variants of two of the three well-known axioms of probability [33]. As mentioned earlier, variants of these axioms have appeared in [17] and [3]. Those axiomatizations also contain the third axiom of probability, represented by Fagin et al. as $P_t(\text{true}) = 1$ and by Bacchus as $P_t(\phi) + P_t(\neg\phi) = 1$. These axioms follow in our logic as a consequence of other axioms describing inevitability and probability. The validity of the third axiom of probability is proven as Theorem 7 in the next section.

Axiom P3 is called *Miller's principle* and several nontemporal variants of it were first suggested by Brian Skyrms [52] as possible constraints on higher-order probabilities. Miller's principle is useful for two reasons. First, it formalizes our intuitions about the relation between chance at various times. For example, suppose that I have two coins, one fair coin and one with a 70% chance of heads and that I am going to choose one coin and flip it. What is the chance now of heads given that I will choose the biased

coin? By intuition and Miller's principle it is 70%. Notice that, as Skyrms [52, Appendix 2] has pointed out, if we have a sentence in the language corresponding to the-biased-coin-is-chosen and appropriate conditional probabilities we can always represent such higher-order probability statements with simple probabilities. But we may not always have such sentences readily available.

The second useful consequence of Miller's principle is that it allows the current chance of facts and events to be inferred from the chances of their future chances: the probability at a given time is the expected value of the probability at any future time. Suppose I am going to choose at random between the two coins above. There is a 50% chance that the chosen coin will have a 70% chance of landing heads and a 50% chance that the coin will have a 50% chance of landing heads. By Miller's principle, it follows that there is now a 60% chance that the coin flip will result in heads.

Probability of logical equivalents rule

This inference rule is valid since if two sentences are logically equivalent then they hold in exactly the same possible worlds.

PE. From $\phi \leftrightarrow \psi$ infer $P_t(\phi) = P_t(\psi)$.

Temporal logic axioms

These axioms capture the temporal component of the logic. No specific temporal axioms are needed for reasoning about relations among time intervals since this is captured by the field axioms above.

TL1. Facts hold over their subintervals, except possibly at the endpoints.

$$(t_1 \leqslant t_2 \leqslant t_3 \leqslant t_4) \wedge (t_1 \neq t_3) \wedge (t_2 \neq t_4) \rightarrow [HOLDS(FA, t_1, t_4) \rightarrow HOLDS(FA, t_2, t_3)].$$

TL2. $HOLDS(FA, t_1, t_2) \rightarrow (t_1 \leqslant t_2)$.

TL3. $OCCURS(EV, t_1, t_2) \rightarrow (t_1 \leqslant t_2)$.

TL4. $OCC(A, t_1, t_2) \rightarrow (t_1 \leqslant t_2)$.

Action axiom

This axiom is valid since *OCC* is true iff both the generative and generated event tokens of the action occur, while *ATT* is true iff the generative event token occurs.

ACT1. $OCC(A, t_A, t'_A) \rightarrow ATT(A, t_A)$.

Inevitability and equality axiom

This axiom is valid because the interpretation of the equality predicate is rigid across worlds.

IE. Equality is inevitable: $\Diamond_t(trm_1 = trm_2) \rightarrow \Box_t(trm_1 = trm_2)$.

Inevitability temporal axioms

This set of axioms relates the inevitability operator and the temporal component of the logic. The first axiom is valid since the set of accessible worlds becomes smaller as we move into the future. The second axiom is valid since the interpretation of relations among time points is rigid. The remaining axioms capture the intuitive property that the present and past are determined. They are valid since the inevitability operator is defined over worlds that are indistinguishable up to and including a given time.

IT1. Inevitability persists.

$$(t_1 \leqslant t_2) \rightarrow (\Box_{t_1}\phi \rightarrow \Box_{t_2}\phi).$$

IT2. Temporal relations are inevitable.

$$\Diamond_t(t_1 < t_2) \rightarrow \Box_t(t_1 < t_2).$$

IT3. Present and past facts are inevitable.

$$(t_1 \leqslant t_2) \rightarrow [\Box_{t_2} HOLDS(FA, t_0, t_1) \vee \Box_{t_2}\neg HOLDS(FA, t_0, t_1)].$$

IT4. Present and past events are inevitable.

$$(t_1 \leqslant t_2) \rightarrow [\Box_{t_2} OCCURS(EV, t_0, t_1) \vee \Box_{t_2}\neg OCCURS(EV, t_0, t_1)].$$

IT5. Present and past action occurrences are inevitable.

$$(t_1 \leqslant t_2) \rightarrow [\Box_{t_2} OCC(A, t_0, t_1) \vee \Box_{t_2}\neg OCC(A, t_0, t_1)].$$

IT6. Present and past action attempts are inevitable.

$$(t_0 \leqslant t_1) \rightarrow [\Box_{t_1} ATT(A, t_0) \vee \Box_{t_1}\neg ATT(A, t_0)].$$

Axioms relating inevitability and probability

IP1. Inevitability implies certainty: $\Box_t\phi \rightarrow P_t(\phi) = 1$.

IP2. Current chance is inevitable: $\Diamond_t[P_t(\phi) \geqslant \alpha] \rightarrow \Box_t[P_t(\phi) \geqslant \alpha]$.

4.2. Theorems

In this section we present theorems that will either be useful later or help to illustrate properties of the logic. For illustrative purposes, we provide proofs of a few of the less obvious theorems. Proofs are presented in two-column format with the justifying axiom or inference rule in the right column. A reference to an inference rule is followed by a list of the axioms and/or steps to which the rule was applied. We first present two theorems that will be used in later derivations.

Theorem 4. *From* $\phi \rightarrow \psi$ *and* $\psi \rightarrow \gamma$ *infer* $\phi \rightarrow \gamma$.

Theorem 5. $\phi \rightarrow \Diamond_t\phi$.

4.2.1. Field theorems

We show the derivation of the following simple theorem to illustrate the use of the field and equality axioms. Henceforth we will simply cite the field and equality axioms to justify multiple derivation steps using them.

Theorem 6. $(trm_1 = trm_2 + s - s) \rightarrow (trm_1 = trm_2)$.

Proof.

1.	$s + (-1 \cdot s) = 0$	F3
2.	$[s + (-1 \cdot s) = 0] \rightarrow$ $[(trm_1 = trm_2 + s - s) \rightarrow (trm_1 = trm_2 + 0)]$	EQ2
3.	$(trm_1 = trm_2 + s - s) \rightarrow (trm_1 = trm_2 + 0)$	MoPo: 1, 2
4.	$trm_2 + 0 = trm_2$	F2
5.	$[trm_2 + 0 = trm_2] \rightarrow [(trm_1 = trm_2 + 0) \rightarrow (trm_1 = trm_2)]$	EQ2
6.	$(trm_1 = trm_2 + 0) \rightarrow (trm_1 = trm_2)$	MoPo: 4, 5
7.	$(trm_1 = trm_2 + s - s) \rightarrow (trm_1 = trm_2)$	MoPo: 3, 6

4.2.2. Probability theorems

If we think of probability one as analogous to inevitability and probability greater than zero as analogous to possibility, then several of the theorems in this section can be seen as probabilistic analogues of theorems describing inevitability and possibility.

Theorem 7. *From* ϕ *infer* $P_t(\phi) = 1$.

This is the third Kolmogorov axiom of probability mentioned above. It follows from properties of inevitability and the relation between probability and inevitability:

Proof.

1.	ϕ	Hypothesis
2.	$\Box_t \phi$	NEC
3.	$\Box_t \phi \rightarrow P_t(\phi) = 1$	IP1
4.	$P_t(\phi) = 1$	MoPo.

Theorem 8. $P_t(\phi) + P_t(\neg\phi) = 1$.

Theorem 9. $P_t(\phi \vee \psi) = P_t(\phi) + P_t(\psi) - P_t(\phi \wedge \psi)$.

Theorem 10. *Stronger sentences have lower probability. From* $\phi \rightarrow \psi$ *infer* $P_t(\phi) \leqslant P_t(\psi)$.

Theorem 11. $P_t(\phi) = 1 \rightarrow P_t(\phi \wedge \psi) = P_t(\psi)$.

Theorem 12. *Chance is the expected value of future chance.*

$$(t_1 \leqslant t_2) \rightarrow [P_{t_1}(P_{t_2}(\phi) \geqslant \alpha) \geqslant \beta \rightarrow P_{t_1}(\phi) \geqslant \alpha \cdot \beta]$$

Proof.

1. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(\phi \mid P_{t_2}(\phi) \geqslant \alpha) \geqslant \alpha$ — P3
2. $P_{t_1}(\phi) = P_{t_1}(\phi \wedge P_{t_2}(\phi) \geqslant \alpha) + P_{t_1}(\phi \wedge \neg P_{t_2}(\phi) \geqslant \alpha)$ — P2
3. $P_{t_1}(\phi) = P_{t_1}(\phi \mid P_{t_2}(\phi) \geqslant \alpha) \cdot P_{t_1}(P_{t_2}(\phi) \geqslant \alpha) + P_{t_1}(\phi \wedge \neg P_{t_2}(\phi) \geqslant \alpha)$ — definition of c-prob: 2
4. $(t_1 \leqslant t_2) \rightarrow [P_{t_1}(P_{t_2}(\phi) \geqslant \alpha) \geqslant \beta \rightarrow P_{t_1}(\phi) \geqslant \alpha \cdot \beta + P_{t_1}(\phi \wedge \neg P_{t_2}(\phi) \geqslant \alpha)]$ — Field axioms, EQ2: 1, 3
5. $(t_1 \leqslant t_2) \rightarrow [P_{t_1}(P_{t_2}(\phi) \geqslant \alpha) \geqslant \beta \rightarrow P_{t_1}(\phi) \geqslant \alpha \cdot \beta]$ — Field axioms

Theorem 13. *Certainly later certainty implies truth.*

$$(t_1 \leqslant t_2) \rightarrow P_{t_1}(P_{t_2}(\phi) = 1 \rightarrow \phi) = 1.$$

This is the probabilistic analogue of axiom I1.

Proof.

1. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(\phi \mid P_{t_2}(\phi) = 1) = 1$ — P3
2. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(\phi \wedge P_{t_2}(\phi) = 1) = P_{t_1}(P_{t_2}(\phi) = 1)$ — definition of c-prob

Let $A \equiv \phi$ and $B \equiv P_{t_2}(\phi) = 1$. Then

3. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(A \wedge B) = P_{t_1}(B)$
4. $P_{t_1}(B \rightarrow A) = P_{t_1}(\neg B) + P_{t_1}(A) - P_{t_1}(\neg B \wedge A)$ — Theorem 9
5. $P_{t_1}(B \rightarrow A) = 1 - P_{t_1}(B) + P_{t_1}(A) - P_{t_1}(\neg B \wedge A)$ — Theorem 8
6. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(B \rightarrow A) = 1 - P_{t_1}(A \wedge B) + P_{t_1}(A) - P_{t_1}(\neg B \wedge A)$ — EQ2: 3, 5
7. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(B \rightarrow A) = 1 - P_{t_1}(A) + P_{t_1}(A)$ — P2, EQ2
8. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(B \rightarrow A) = 1$ — Theorem 6
9. $(t_1 \leqslant t_2) \rightarrow P_{t_1}(P_{t_2}(\phi) = 1 \rightarrow \phi) = 1$ — substituting back

Theorem 14. *Certainty certainly persists.*

$$(t_0 \leqslant t_1 \leqslant t_2) \rightarrow P_{t_0}(P_{t_1}(\phi) = 1 \rightarrow P_{t_2}(\phi) = 1) = 1.$$

This is the probabilistic analogue of axiom IT1. Note that it is not valid that certainty simply persists since semantic constraint (C6) only applies to equivalence classes of worlds of positive probability.

Theorem 15. *Facts have higher chance of holding over their subintervals.*

$$(t_1 \leqslant t_2 \leqslant t_3 \leqslant t_4) \wedge (t_1 \neq t_3) \wedge (t_2 \neq t_4) \rightarrow P_t(HOLDS(FA, t_2, t_3)) \geqslant P_t(HOLDS(FA, t_1, t_4)).$$

4.2.3. Inevitability of the past

Theorem 16. *Possibility persists into the past.*

$$(t_1 \leqslant t_2) \rightarrow (\Diamond_{t_2}\phi \rightarrow \Diamond_{t_1}\phi).$$

Theorem 17. *Present and past inevitability are inevitable.*

$$(t_1 \leqslant t_2) \rightarrow (\Diamond_{t_2} \Box_{t_1} \phi \rightarrow \Box_{t_2} \Box_{t_1} \phi).$$

Theorem 18. *Present and past possibility are inevitable.*

$$(t_1 \leqslant t_2) \rightarrow (\Diamond_{t_2} \Diamond_{t_1} \phi \rightarrow \Box_{t_2} \Diamond_{t_1} \phi).$$

Theorem 19. *Present and past chance are inevitable.*

$$(t_1 \leqslant t_2) \rightarrow (\Diamond_{t_2} P_{t_1}(\phi) \geqslant \alpha \rightarrow \Box_{t_2} P_{t_1}(\phi) \geqslant \alpha).$$

Proof.

1.	$(t_1 \leqslant t_2) \rightarrow \Diamond_{t_2} P_{t_1}(\phi) \geqslant \alpha \rightarrow \Diamond_{t_1} P_{t_1}(\phi) \geqslant \alpha$	Theorem 16
2.	$\Diamond_{t_1} P_{t_1}(\phi) \geqslant \alpha \rightarrow \Box_{t_1} P_{t_1}(\phi) \geqslant \alpha$	IP2
3.	$(t_1 \leqslant t_2) \rightarrow \Box_{t_1} P_{t_1}(\phi) \geqslant \alpha \rightarrow \Box_{t_2} P_{t_1}(\phi) \geqslant \alpha$	IT1
4.	$(t_1 \leqslant t_2) \rightarrow \Diamond_{t_2} P_{t_1}(\phi) \geqslant \alpha \rightarrow \Box_{t_2} P_{t_1}(\phi) \geqslant \alpha$	Theorem 4: 1–3

4.2.4. Certainty of the present and past

By axiom IP1 that inevitability implies certainty, the following theorems that the past is certain follow directly from the corresponding axioms and theorems for inevitability.

Theorem 20. *Present and past facts are certain.*

$$(t_1 \leqslant t_2) \rightarrow [P_{t_2}(HOLDS(FA, t_0, t_1)) = 1 \vee P_{t_2}(HOLDS(FA, t_0, t_1)) = 0].$$

Similar theorems hold for OCCURS, OCC, and ATT.

Theorem 21. *Present and past inevitability is certain.*

$$(t_1 \leqslant t_2) \rightarrow (P_{t_2}(\Box_{t_1}\phi) > 0 \rightarrow P_{t_2}(\Box_{t_1}\phi) = 1).$$

Theorem 22. *Present and past possibility is certain.*

$$(t_1 \leqslant t_2) \rightarrow (P_{t_2}(\Diamond_{t_1}\phi) > 0 \rightarrow P_{t_2}(\Diamond_{t_1}\phi) = 1).$$

Theorem 23. *Present and past chance is certain.*

$$(t_1 \leqslant t_2) \rightarrow (P_{t_2}(P_{t_1}(\phi) \geqslant \alpha) > 0 \rightarrow P_{t_2}(P_{t_1}(\phi) \geqslant \alpha) = 1).$$

4.2.5. Modal operators and quantifiers

The following theorems capture the relationships between the quantifiers and the modal operators. These theorems hold because the domain of individuals does not vary from world to world.

Theorem 24 (Barcan formula).

$$\forall x \Box_t \phi \rightarrow \Box_t \forall x \phi, \text{ where } x \text{ does not occur free in } t.$$

Theorem 25 (Converse Barcan formula).

$$\Box_t \forall x \phi \rightarrow \forall x \Box_t \phi, \text{ where } x \text{ does not occur free in } t.$$

Theorem 26 (Probabilistic converse Barcan formula).

$$P_t(\forall x \phi) = \alpha \rightarrow \forall x P_t(\phi) \geqslant \alpha.$$

Proof.

1.	$\forall x \phi \rightarrow \phi$	FOL4
2.	$P_t(\forall x \phi \rightarrow \phi) = 1$	Theorem 7
3.	$P_t(\forall x \phi \rightarrow \phi) = 1 \rightarrow P_t(\phi \wedge \forall x \phi) = \alpha$	Theorem 11
4.	$P_t(\phi \wedge \forall x \phi) = \alpha$	MoPo: 2, 3
5.	$P_t(\phi) \geqslant \alpha$	Theorem 10

Note that the probabilistic analogue of the Barcan formula

$$\forall x \; P_t(\phi) = 1 \rightarrow P_t(\forall x \; \phi) = 1$$

is not valid because the antecedent can be satisfied by a model in which each x is true in *almost* all worlds.

4.2.6. Actions

Theorem 27. *Attempted acts are feasible.*[6]

$$ATT(A, t_A) \rightarrow FEAS(A, t_A).$$

[6] The definition of feasibility is given in Section 5.1

Proof.

1.	$ATT(A, t_A) \rightarrow \Diamond_{t_A} ATT(A, t_$	Theorem 5
2.	$(t \leqslant t_A) \rightarrow (\Diamond_{t_A} ATT(A, t_A))$	Theorem 16
3.	$\forall t\ (t \leqslant t_A) \rightarrow (\Diamond_{t_A} ATT(A, t_A))$	UG
4.	$\forall t\ \Diamond_{t_A} ATT(A, t_A) \rightarrow [(t \leqslant t_A)]$	FOL axioms
5.	$ATT(A, t_A) \rightarrow [\forall t\ (t \leqslant t_A)$	FOL axioms, MoPo
6.	$ATT(A, t_A) \rightarrow FEAS(A, t_A)$	definition of FEAS

Theorem 28. *Past and present feasit*

$$(t_A \leqslant t) \rightarrow [\Diamond_t FEAS(A, t_A) \rightarrow [$$

4.3. Examples

In this section two simple examples llustrate the use of the logic in reasoning about actions. The first exam logic can be used to distinguish between factors that an agent cannot i they are inevitable and factors the agent cannot influence because the endent of his available actions. Consider a game in which balls are (Suppose that at time t_1 you will randomly be given one of three l ning two urns. In box 1 both urns contain all red balls. In box 2 th h red and white balls and the proportions in the two urns are identica contains all red balls and the other contains all white balls. At time ose one of the urns from the box you are given and give the urn to me t_3 she is then to choose a ball at random from that urn. What is t can influence the chance that she will draw a red ball? The situation y the following sentences:

$$(now < t_1 < t_2 < t_3 < t_4), \tag{1}$$

$$P_{now}(\Box_{t_1}\ OCCURS(\text{red-picked}, t_3, \tag{2}$$

$$P_{now}(P_{t_2}(OCCURS(\text{red-picked}, t_3, \forall x\ P_{t_2}(OCCURS(\text{red-picked}, k(x), t_2, t_3)) = P_{t_2}(OCCURS(\text{red-picked} \tag{3}$$

One model in which these sentences are rlds is shown in Fig. 7.

Since inevitability persists (axiom IT1 tences have lower probability (Theorem 10) it follows from sentence

$$P_{now}(\Box_{t_2} OCCURS(\text{red-picked}, t_3, t_4 \tag{4}$$

Since inevitability implies certainty (axi from (4) that

$$P_{now}(P_{t_2}(OCCURS(\text{red-picked}, t_3, t_4 \tag{5}$$

By Theorem 11,

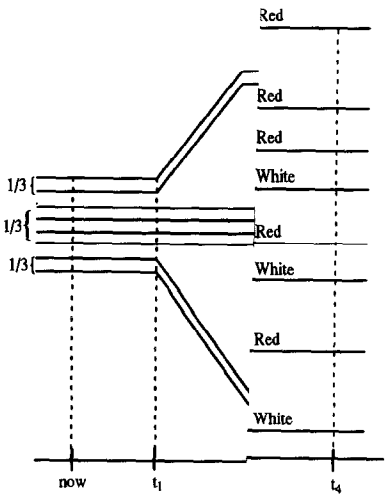

Fig. 7. Possible.

$$P_{t_2}(OCCURS(\text{red-picked}, t_3, t_4)) \\ P_{t_2}(OCCURS(\text{red-picked}, t_3, t_4, t_2, t_3)) = \\ P_{t_2}(OCC(\text{pick}(x), t_2, t_3)). \tag{6}$$

By inference rule UG and Theorem 7) that

$$P_{now}(P_{t_2}(OCCURS(\text{red-picked}, t_2 \\ \forall x\ P_{t_2}(OCCURS(\text{red-pick}\text{pick}(x), t_2, t_3)) = \\ P_{t_2}(OCC(\text{pick}(x), t_2, t \tag{7}$$

Theorem 11 applied to (5) and (7) y

$$P_{now}(P_{t_2}(OCCURS(\text{red-picked}, t \\ \forall x\ P_{t_2}(OCCURS(\text{red-pick}\text{pick}(x), t_2, t_3)) = \\ P_{t_2}(OCC(\text{pick}(x), t_2, t \tag{8}$$

And by the definition of conditional p

$$P_{now}(P_{t_2}(OCCURS(\text{red-picked}, \\ \forall x\ P_{t_2}(OCCURS(\text{red-pick}\text{pick}(x), t_2, t_3)) = \\ P_{t_2}(OCCURS(\text{red-pic}\beta. \tag{9}$$

Finally by axiom P2, sentences (3) mbined to yield

$$P_{now}(\forall x\ P_{t_2}(OCCURS(\text{red-pick}\text{pick}(x), t_2, t_3)) = \\ P_{t_2}(OCCURS(\text{red-pic}\beta. \tag{10}$$

So there is at most a 1/3 chance t nce the chance of your partner picking a red ball.

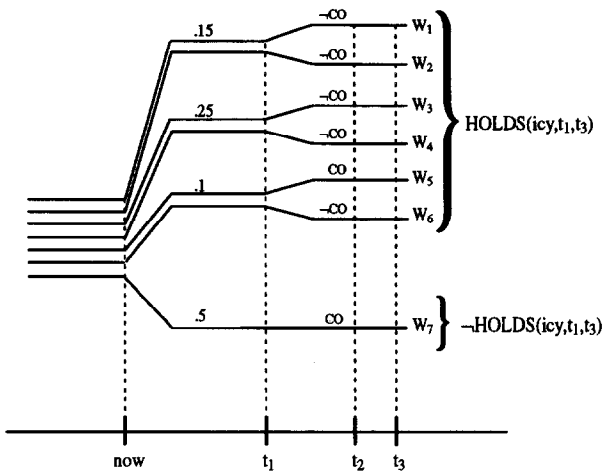

Fig. 8. Possible model for carry example.

The next example is a modified version of an example presented by Pelavin [43]. It illustrates how the ability of the logic to represent both probability and possibility can be used to reason about the chance that two actions can co-occur. Suppose that I am going shopping this evening and want to carry two grocery bags to the car simultaneously. In most cases it is not possible to carry two bags if it is icy out. There is a 50% chance that it will be icy out this evening. What is the chance that carrying both bags simultaneously will not be a possible course of action? The situation can be described by the following three sentences. In these sentences, the terms *now*, t_1, t_2, and t_3 are time constants.

$$\forall t, t' \ (now \leqslant t \leqslant t') \rightarrow P_{now}(\neg \Diamond_t [OCC(\text{carry}(b_1), t, t') \wedge OCC(\text{carry}(b_2), t, t')] \mid HOLDS(\text{icy}, t, t')) = 0.8, \tag{11}$$

$$P_{now}(HOLDS(\text{icy}, t_1, t_3)) = 0.5, \tag{12}$$

$$(now < t_1 < t_2 < t_3). \tag{13}$$

We would like to know the probability of

$$\neg \Diamond_{t_1} [OCC(\text{carry}(b_1), t_1, t_2) \wedge OCC(\text{carry}(b_2), t_1, t_2)].$$

One possible model in which the sentences are satisfied in every world is shown in Fig. 8. The labels "CO" and "¬CO" designate the co-occurrence and non-co-occurrence of the two actions, respectively. Note that in worlds w_1–w_4 we have

$$\neg \Diamond_{t_1} [OCC(\text{carry}(b_1), t_1, t_2) \wedge OCC(\text{carry}(b_2), t_1, t_2)]$$

and in w_5 and w_6 we have

$$\Diamond_{t_1} [OCC(\text{carry}(b_1), t_1, t_2) \wedge OCC(\text{carry}(b_2), t_1, t_2)].$$

The property that facts have higher chance of holding over their subintervals (Theorem 14) together with sentences (12) and (13) entail that

$$P_{now}(HOLDS(\text{icy}, t_1, t_2)) \geqslant 0.5. \tag{14}$$

By axioms FOL5 and MoPo it follows from sentences (11) and (13) that

$$P_{now}(\neg \Diamond_{t_1} [OCC(\text{carry}(b_1), t_1, t_2) \wedge OCC(\text{carry}(b_2), t_1, t_2)] \mid HOLDS(\text{icy}, t_1, t_2)) = 0.8. \tag{15}$$

By axiom P2 it follows from sentences (14) and (15) that

$$P_{now}(\neg \Diamond_{t_1} [OCC(\text{carry}(b_1), t_1, t_2) \wedge OCC(\text{carry}(b_2), t_1, t_2)]) \geqslant 0.4.$$

So there is at least a 40% chance that carrying both bags simultaneously will not be possible.

Furthermore, an upper bound on the current probability of the co-occurrence of the two actions can be calculated. Since inevitability implies certainty (axiom IP1),

$$P_{now}(P_{t_1}(OCC(\text{carry}(b_1), t_1, t_2)) \wedge OCC(\text{carry}(b_2), t_1, t_2)) = 0) \geqslant 0.4$$

and since chance is the expected value of future chance (Theorem 12),

$$P_{now}(OCC(\text{carry}(b_1), t_1, t_2) \wedge OCC(\text{carry}(b_2), t_1, t_2)) \leqslant 0.6.$$

5. Describing properties of actions

The representation of actions contained in our logic is highly expressive. We can describe actions in terms of whether they can be attempted, whether their attempt will bring about their occurrence, and what effects their attempts and occurrences will have. This next section details how actions can be described.

If an action can be attempted we say it is feasible. If an action occurs when attempted, we say that the action is executable. If an action influences the chance of certain conditions, we call these its effects. An action will be described in terms of its feasibility, its executability, and its effects.

Conditions in the world may influence the chance that an action can be attempted. For example, I can only attempt to start my car if I am at the location of the car. Such conditions are called feasibility conditions. Further conditions may influence the chance that the action will occur. For example, if I attempt to start my car I will succeed in starting it only if there is gas in the tank. Such conditions are called executability conditions. And finally conditions may influence the chance that an action will achieve certain effects. For example, if I start my car with the garage door closed there is a good chance that I will be asphyxiated. Such conditions are called ramification conditions. The following three sections discuss how $\mathcal{L}_{\text{tca}}$ can be used to describe these different aspects of actions.

5.1. Feasibility

Sometimes it may not be possible to attempt an action. This is the case if an action's generative event cannot always occur. If it is not possible to attempt an action, we say that the action is *infeasible*. Three types of infeasibility can be distinguished.

(1) It may not be possible to attempt a single action under certain conditions: I cannot attempt to go from home to the office if I am not at home.
(2) It may not be possible to attempt a compound action consisting to two concurrent actions under certain conditions:[7] I cannot attempt to drive north and drive in the direction of the wind at the same time unless the wind is blowing north.
(3) A compound action consisting to two concurrent actions may not be possible to attempt under any circumstances: I cannot attempt to raise and lower my left arm simultaneously under any conditions.

If an action can be attempted we say that it is *feasible*.

The attempt of an action is something the agent chooses to do. As long as an action's generative event is consistent with the state of affairs of the world, the agent can choose to attempt the action and, hence, the action is feasible. *We make the simplifying assumption that an action token's generative event token is instantaneous.*[8,9] Since the agent may or may not choose to attempt an action, for an action to be feasible under this assumption it suffices that there be some action token whose generative event token is possible at an infinitesimal instant before the time of the event token. Since possibility persists into the past (Theorem 16) this can be captured by defining feasibility as possibility at all times before the time of the attempt:

Definition 29. An action A is feasible at a time t_A, written $FEAS(A, t_A)$ iff

$$\forall(t < t_A)\ \Diamond_t\, ATT(A, t_A).$$

This definition of feasibility seems a bit complex and two simpler alternative definitions come to mind more immediately. First, one might be tempted to define feasibility

[7] One could make the assumption that individual actions are always feasible but, as this example shows, once we compose actions into plans we run into the problem that the plans may not be feasible. So we may as well be completely general and not assume individual actions to be feasible.

[8] We can think of a generative event token that spans time t_A t'_A as a set of sequential instantaneous actions A_i. An action is then feasible if each part of the action attempt is feasible in the context of the earlier parts of the action attempt. The chance of feasibility can be represented as

$$\int_{t_A}^{t'_A} \mathrm{d}P_t(\forall(t' < x)\ \Diamond_{t'}\, ATT(A_x, x) \mid ATT(A_{x-t_A}, t_A))\,\mathrm{d}x,$$

where $ATT(A_{x-t_A}, t_A)$ designates the occurrence of the portion of the generative event token from time t_A up to time x.

[9] Notice that had we only represented action occurrences and not attempts, we would have been forced to apply this assumption to occurrences with the result that we could not reason about actions that span time.

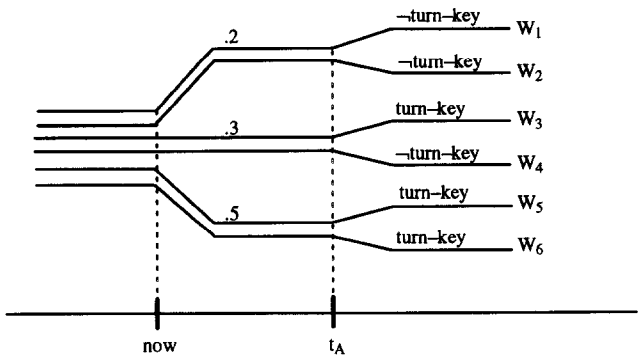

Fig. 9. Chance of feasibility.

more simply as the possibility of the attempt at the time of the attempt: $\Diamond_{t_A} ATT(A, t_A)$. But because the past and present are inevitable (axiom IT6) it follows from this that $\Box_{t_A} ATT(A, t_A)$ and this is far too strong a condition for feasibility. Second, within a probabilistic framework, a more natural definition of feasibility might seem to be $\forall(t < t_A)\ P_t(ATT(A, t_A)) > 0$. But this definition is too restrictive. In particular, it would rule out models which contain a uniform distribution over an uncountably infinite number of worlds and in which the action is attempted in only a finite number of the worlds. In such models, the probability of attempting the action would be zero, yet it would be possible to attempt the action.

In general, there will be a certain chance that an action is feasible—the chance that conditions in the world are consistent with the occurrence of the generative event. The chance at time t that an action A is feasible at time t_A is $P_t(FEAS(A, t_A))$.

Consider again the action of starting my car and suppose that the chance that it is feasible is 0.8, i.e., $P_{now}(FEAS(\text{start-car}, t_A)) = 0.8$. One possible model for this is shown in Fig. 9. In neither world w_1 nor world w_2 does the generative event of turning the key occur. So in both w_1 and w_2 we have $\neg FEAS(\text{start-car}, t_A)$. But in worlds w_3–w_6 the attempt is possible at time t_A. So the chance that it will be within my power to choose to attempt the action is 80%.

5.2. Executability

Once an agent attempts an action, whether or not the action occurs is no longer directly within the agent's control—it is a function of chance.

Definition 30. The chance that an action A is executable at time t_A is the chance that it occurs given that it is attempted:

$$P_t(OCC(A, t_A, t'_A) \mid ATT(A, t_A)).$$

5.3. Effects

One of the objectives of the present work is to capture some of the natural relations that exist between actions and effects. We show how various kinds of effects can be described and how some natural properties of effects follow directly from the models of chance and possibility.

An effect is a condition, the chance of which is influenced by an action. If an action increases the chance of a condition, it is called a *positive effect* and if the action decreases the chance, it is called a *negative effect*. Depending on the nature of the planning problem, we may wish to describe effects in one of two ways. If we are not concerned about the occurrence of actions, we may simply describe effects of action attempts. In this case, the chance of the effect will be represented as

$$P_t(EFF \mid ATT(A, t_A)).$$

More expressive power can be gained by describing effects of action occurrences. In this case, we distinguish between effects of a successful action and effects of a failed action. The effects of an action's occurrence are represented by an expression of the form

$$P_t(EFF \mid OCC(A, t_A, t'_A))$$

and the effects of a failed action are represented as

$$P_t(EFF \mid ATT(A, t_A) \wedge \neg \exists t'_A OCC(A, t_A, t'_A)).$$

Three conditions are *necessary* for *EFF* to be a positive[10] effect of the occurrence of an act A:[11]

(i) Since actions can only influence the future, *EFF* cannot temporally precede the occurrence of A. But *EFF* need not necessarily completely succeed the occurrence of A. Fig. 10 shows the necessary temporal relation between action and effect. The possible temporal relations between actions and their effects that are allowed by this constraint are shown in Fig. 11. An effect may begin before the time of an action: the action of saying the word "seven" at the right time contributes to the effect of reciting the Gettysburg address. An effect may persist after the time of an action: the effect of the hockey puck sliding across the ice persists after the action of pushing it terminates. An effect may begin exactly once the action is over: the effect of the glass being empty begins exactly once all the water is poured out. An effect may end before the action is over: the effect of the box sliding across the floor ends once it hits the wall, even though the push action is still occurring. Finally, an effect may begin some time after an action is over: the effect of the bomb exploding occurs some time after the action of setting the timer.

(ii) If an action has no chance of occurring at a particular time, it cannot have effects at that time. So A must have a positive chance of occurring.

[10] Negative effects are described similarly.

[11] Note that these three conditions are similar to Suppes' [54] prima facie causality conditions.

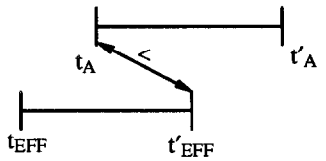

Fig. 10. Relative times of an action and its effect.

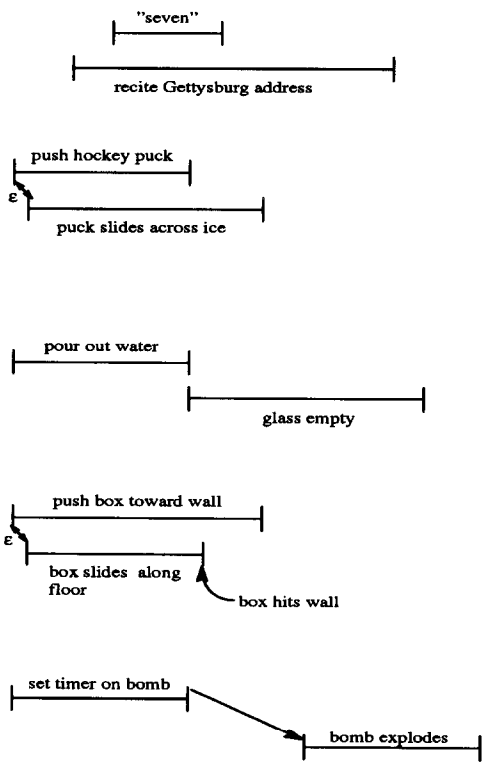

Fig. 11. Possible temporal relations between actions and effects.

(iii) An effect is a condition influenced by an action, so the occurrence of A must positively influence the chance of EFF.

These conditions are fundamental properties of effects of any kind—effects of actions, effects of events, even effects of facts. They will arise again in the discussion of conditions that influence actions since the feasibility, executability, and effects of actions are in a sense effects of such conditions.

Suppose act A occurs in the interval t_A to $t_{A'}$ and let EFF be a fact, event, or action occurrence: $HOLDS(FA, t_{EFF}, t'_{EFF})$, $OCCURS(EV, t_{EFF}, t'_{EFF})$, or $OCC(A, t_{EFF}, t'_{EFF})$. Then the conditions (i)–(iii) can be stated in the logic as:

(1) temporal non-succession: $t_A < t'_{EFF}$,
(2) positive chance of occurrence: $P_{t_A}(OCC(A, t_A, t'_A)) > 0$,[12]
(3) *positive influence*: $P_{t_A}(EFF \mid OCC(A, t_A, t_{A'})) > P_{t_A}(EFF)$.

Due to the way the semantics of $\mathcal{L}_{\text{tca}}$ has been defined, condition (3) entails conditions (1) and (2). If condition (3) is expanded out into its proper form in the logic it becomes:

$$P_{t_A}(EFF \wedge OCC(A, t_A, t'_A)) > P_{t_A}(EFF) \cdot P_{t_A}(OCC(A, t_A, t'_A)).$$

If condition (2) is false, i.e., $P_{t_A}(OCC(A, t_A, t'_A)) = 0$, then the above sentence is false since both sides of the inequality are zero. So if (3) holds, (2) must hold. Next, if (1) is false then $t'_{EFF} \leqslant t_A$. Since the present and past are certain (Theorem 20),

$$P_{t_A}(EFF) = 0 \vee P_{t_A}(EFF) = 1.$$

Either case contradicts condition (3). So if (3) holds, (1) must also hold.

This result shows that the models have captured the natural temporal relation between actions and effects—actions can only affect future conditions. As a consequence of this result, if we use condition (3) to define what is necessary for a plan to achieve a goal then actions after the time of the goal cannot contribute to achieving the goal.

The ability of the logic to represent and distinguish between truth, probability, and possibility allows us to distinguish between potential effects and actual effects. We can define actual effects as potential effects that actually occur:

$$EFF \wedge P_{t_A}(EFF \mid OCC(A, t_A, t'_A)) > P_{t_A}(EFF).$$

5.4. Conditions that influence actions

Conditions in the world may influence the feasibility, executability, and effects of actions. Such conditions are called feasibility conditions, executability conditions, and ramification conditions, respectively. In this section we present conditions, similar to those for effects, that are necessary for something to be a feasibility, executability, or ramification condition of an action.

5.4.1. Feasibility conditions

Certain conditions may influence the chance that an action will be feasible. If the conditions have a positive influence they are called *positive feasibility conditions* and otherwise they are called *negative feasibility conditions*. We discuss positive feasibility conditions here; negative feasibility conditions are similar. Let FC be a fact, event, or action occurrence: $HOLDS(FA, t_{FC}, t'_{FC})$, $OCCURS(EV, t_{FC}, t'_{FC})$, or $OCC(A, t_{FC}, t'_{FC})$. As in our discussion of effects, three conditions are *necessary* for FC to be a positive feasibility condition for action A at time t_A:

[12] Note that this is equivalent to saying that the action has a positive chance of being executable, $P_{t_A}(OCC(A, t_A, t'_A) \mid ATT(A, t_A)) > 0$. The equivalence follows from the fact that $P_{t_A}(OCC(A, t_A, t'_A)) > 0$ entails that $P_{t_A}(ATT(A, t_A)) > 0$.

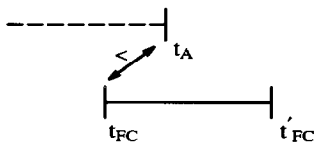

Fig. 12. Relative times of feasibility condition and action attempt.

(1) temporal non-succession: $t_{FC} < t_A$,
(2) positive chance of the condition: $P_{t_{FC}}(FC) > 0$,
(3) *positive influence*: $P_{t_{FC}}(FEAS(A, t_A) \mid FC) > P_{t_{FC}}(FEAS(A, t_A))$.

The temporal non-succession condition $t_{FC} < t_A$ is depicted in Fig. 12.

As was the case for action effects, condition (3) entails both conditions (1) and (2). The argument that (3) entails (2) is exactly the same as for effects. The proof that (3) entails (1) is slightly more complicated. Suppose that condition (1) is false:

1. $(t_A \leqslant t_{FC}) \rightarrow$
   $[\Diamond_{t_{FC}} FEAS(A, t_A) \rightarrow \Box_{t_{FC}} FEAS(A, t_A)]$ Theorem 28
2. $P_{t_{FC}}(FEAS(A, t_A)) > 0 \rightarrow \Diamond_{t_{FC}} FEAS(A, t_A)$ IP1
3. $\Box_{t_{FC}} FEAS(A, t_A) \rightarrow P_{t_{FC}}(FEAS(A, t_A)) = 1$ IP1
4. $(t_A \leqslant t_{FC}) \rightarrow [P_{t_{FC}}(FEAS(A, t_A)) > 0 \rightarrow$
   $P_{t_{FC}}(FEAS(A, t_A)) = 1]$ Theorem 4: 1–3
5. $(t_A \leqslant t_{FC}) \rightarrow [P_{t_{FC}}(FEAS(A, t_A)) = 0 \vee$
   $P_{t_{FC}}(FEAS(A, t_A)) = 1]$ definition of $\rightarrow$.

Both $P_{t_{FC}}(FEAS(A, t_A)) = 0$ and $P_{t_{FC}}(FEAS(A, t_A)) = 1$ contradict condition (3). So if (3) holds, (1) must hold. So again the models have captured a fundamental temporal relation: a condition that succeeds an action's attempt cannot influence the feasibility of that action.

5.4.2. *Executability conditions*

Certain conditions may influence the chance that an action is executable. Conditions with a positive influence are called *positive executability conditions* and conditions with a negative influence are called *negative executability conditions*. Let EC be a fact, event, or action occurrence with associated interval t_{EC}, t'_{EC}. Three conditions are *necessary* for EC to be a positive executability condition for action A at time t_A:

(1) temporal non-succession: $t_{EC} < t'_A$,
(2) positive chance of attempt and condition: $P_{t_{EC}}(ATT(A, t_A) \wedge EC) > 0$,
(3) *positive influence*:

$$P_{t_{EC}}(OCC(A, t_A, t'_A) \mid ATT(A, t_A) \wedge EC) >$$
$$P_{t_{EC}}(OCC(A, t_A, t'_A) \mid ATT(A, t_A)).$$

Negative executability conditions are represented by negating the occurrence of A. Once again, condition (3) entails both conditions (1) and (2). The proofs are similar to the

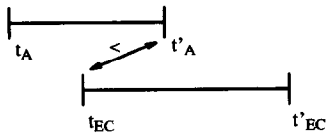

Fig. 13. Relative times of executability condition and action occurrence.

earlier proofs showing that the effects of an action cannot occur prior to the time of the action.

The temporal non-succession condition $t_{EC} < t'_A$ is depicted in Fig. 13. It just says that in order for EC to influence the success of A, it cannot be a constraint on the state of the world after the time of the action. This makes good intuitive sense because executability conditions should not just provide evidence for the executability of an action but they should influence the executability. Hence, they should be worth bringing about. For example, consider a condition after the time of the action that increases the chance now that the action will be executable. Suppose that typically 15 minutes after I start my car the engine is warm:

$$P_{now}(OCC(\text{start}(\text{car}), t_1, t_2) \mid ATT(\text{start}(\text{car}), t_1) \wedge \exists t\, HOLDS(\text{warm}(\text{engine}(\text{car})), t_2 + 15\text{min}, t)) > P_{now}(OCC(\text{start}(\text{car}), t_1, t_2) \mid ATT(\text{start}(\text{car}), t_1)).$$

We would not want to call $\exists t\, HOLDS(\text{warm}(\text{engine}), t_2 + 15\text{min}, t)$ an executability condition for starting my car and to generate a plan to warm up my car in the future in order to make my car start now.

5.4.3. Ramification conditions

Certain conditions in the world may influence the chance of an action's positive or negative effects. If the conditions have a positive influence on the effect, they are called *positive ramification conditions* and otherwise they are called *negative ramification conditions*. In some cases a positive ramification condition may even turn a negative effect into a positive effect, and conversely for negative ramification conditions. Let RAM be a fact, event, or action occurrence with associated interval $\langle t_{RAM}, t'_{RAM} \rangle$. Since we will primarily be concerned with the effects of the occurrences of successful actions, we will describe ramification conditions within this context. The definitions for failed actions and for simple action attempts are similar. Three conditions are necessary for RAM to be a positive ramification condition with respect to effect EFF for action A at time t_A:

(1) Temporal non-succession: $t_{RAM} < t'_{EFF}$,

(2) Positive chance of occurrence and condition: $P_{t_{RAM}}(OCC(A, t_A, t'_A) \wedge RAM) > 0$,

(3) *Positive influence*:

$$P_{t_{RAM}}(EFF \mid OCC(A, t_A, t'_A) \wedge RAM) > P_{t_{RAM}}(EFF \mid OCC(A, t_A, t'_A)).$$

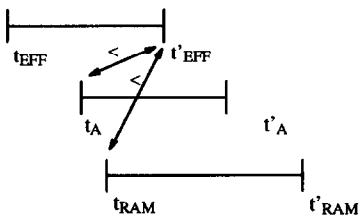

Fig. 14. Relative times an action, its effect, and ramification condition.

Negative ramification conditions are represented by negating *EFF*. Once again, condition (3) entails both conditions (1) and (2).

The temporal relations between the action, the effect, and the ramification condition are depicted in Fig. 14. The temporal non-succession condition just says that in order for *RAM* to influence *EFF*, it cannot be a constraint on the state of the world after the time of the effect. Note, however, that *RAM* can occur after the time of the action because the effect of the action can be delayed.

5.5. *Properties of plans*

In general we will be interested in reasoning about plans rather than single actions. A *plan* is simply a set of actions attempted at particular times. By specifying the times of the attempts within the plan, we can represent plans containing both sequential and concurrent actions in a homogeneous manner. To reason about plans, we will simply reason about the set of actions composing them.

5.5.1. *Feasibility of plans*

Plan feasibility is a more complex concept than action feasibility. Attempting a plan means attempting all the actions composing the plan, so a plan attempt is the conjunction of the individual action attempts. A plan is feasible if it can be attempted, i.e., if all the actions composing it can be attempted. We first examine two superficially appealing but incorrect ways of defining the chance of plan feasibility and then present a definition that has the desired properties. Consider the plan $ATT(A_1, t_{A_1}) \wedge ATT(A_2, t_{A_2})$, with any ordering of the attempts. First, we might be tempted to define the chance that this plan is feasible simply as the chance that each action is feasible:

$$P_t(FEAS(A_1, t_{A_1}) \wedge FEAS(A_2, t_{A_2})).$$

But this expression gives us too high a chance for plan feasibility because the statement $FEAS(A_1, t_{A_1}) \wedge FEAS(A_2, t_{A_2})$ could be satisfied by a model in which each action is attempted in some world but there is no world where they are both attempted together. Second, by analogy to action feasibility we might be tempted to define the chance that the plan is feasible as

$$P_t(\forall(t < t_{A_1})\ \Diamond_t (ATT(A_1, t_{A_1}) \wedge ATT(A_2, t_{A_2}))).$$

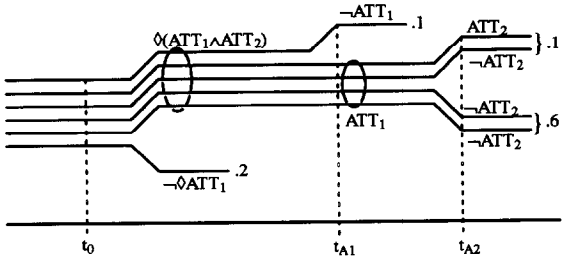

Fig. 15. Example of plan feasibility.

If $t_{A_1} = t_{A_2}$ this is the correct expression for feasibility since the joint attempt of A_1 and A_2 is equivalent to the attempt of the single more complex action. But suppose that $t_{A_1} < t_{A_2}$ and consider the model in Fig. 15. The chance of the above sentence is 0.8, but the chance that action A_1 is feasible is 0.8 while the chance that A_2 is feasible given that A_1 is attempted is only 0.14. So the chance that the actions composing the plan can both be attempted is only $0.8 \cdot 0.14 = 0.11$. Hence this expression also gives us too high a value for plan feasibility.

Each action in a plan must be feasible in the context of the attempts of the earlier actions in the plan since we wish to attempt the entire plan. So the chance that the above plan is feasible is

$$P_t(FEAS(A_2, t_{A_2}) \mid ATT(A_1, t_{A_1})) \cdot P_t(FEAS(A_1, t_{A_1})).$$

If $t_{A_1} = t_{A_2}$ this is also a valid expression for plan feasibility. So when $t_{A_1} = t_{A_2}$ we have

$$P_t(FEAS(A_2, t_{A_2}) \mid ATT(A_1, t_{A_1})) \cdot P_t(FEAS(A_1, t_{A_1})) =$$
$$P_t(FEAS(\{A_1, A_2\}, t_{A_1})),$$

where

$$FEAS(\{A_1, A_2\}, t_{A_1}) \equiv$$
$$\forall t\ (t < t_{A_1}) \rightarrow \Diamond_t[ATT(A_1, t_{A_1}) \wedge ATT(A_2, t_{A_2})].$$

Finally consider the plan $ATT(A_1, t_{A_1}) \wedge ATT(A_2, t_{A_2}) \wedge ATT(A_3, t_{A_3})$, where $t_{A_1} = t_{A_2} < t_{A_3}$. The chance that this plan isfeasible is

$$P_t(FEAS(A_3, t_{A_3}) \mid ATT(A_1, t_{A_1}) \wedge ATT(A_2, t_{A_2}))$$
$$\cdot P_t(FEAS(A_2, t_{A_2}) \mid ATT(A_1, t_{A_1})) \cdot P_t(FEAS(A_1, t_{A_1}))$$

or equivalently,

$$P_t(FEAS(A_3, t_{A_3}) \mid ATT(A_1, t_{A_1}) \wedge ATT(A_2, t_{A_2}))$$
$$\cdot P_t(FEAS(\{A_1, A_2\}, t_{A_1})).$$

Definition 31. In general, the chance that the plan consisting of the set of n action attempts $ATT(A_i, t_{A_i})$ with $t_{A_1} \leqslant t_{A_2} \leqslant \cdots \leqslant t_{A_n}$ is feasible is

$$\prod_{i=1}^{n} P_{now}\left(FEAS(A_i, t_{A_i}) \mid \bigwedge_{j<i} ATT(A_j, t_{A_j})\right).$$

5.5.2. Executability of plans

The chance that a plan is executable is represented by an expression of the form

$$P_t\left(\bigwedge_i OCC(A_i, t_{Ai}, t'_{Ai}) \mid \bigwedge_i ATT(A_i, t_{Ai})\right).$$

Just as with individual actions, plans may have executability conditions.

5.5.3. Effects of plans

The chance of the effect of a successful plan is represented by an expression of the form

$$P_t\left(EFF \mid \bigwedge_i OCC(A_i, t_{Ai}, t'_{Ai})\right).$$

Effects of plan attempts and of failed plans are represented similarly. Just as with individual actions, plans may have ramification conditions.

6. Describing and reasoning about planning problems

In this section we present a general framework for describing planning problems. The purpose of the framework is to provide guidance in the use of the rich language $\mathcal{L}_{\text{tca}}$. The main purpose of describing planning problems is to use the description in the generation and evaluation of plans. So in developing the framework, we focus on descriptions that are compositional, in the sense that properties of a plan may be inferred from the properties of the component actions and the environment. We present one way in which $\mathcal{L}_{\text{tca}}$ can be used to specify such a description. Many alternatives are possible.

6.1. Representing the planning problem

In this section we show how $\mathcal{L}_{\text{tca}}$ can be used to describe planning problems so that we can reason about the chance that a plan will achieve a given goal. Since the best knowledge we have about chance is the chance now, in describing the planning problem all probabilities will be taken relative to the present time. A planning problem is described in terms of (i) the planning environment, (ii) the individual action specifications, (iii) the action interactions, and (iv) the goal description. Following Pelavin [43], we call the world within which planning takes place the planning environment. Since we are working in a temporal setting, the description of the planning environment can include chances of facts and events that hold or occur at any time. The individual action specifications describe feasibility conditions, executability conditions, and effects with associated ramification conditions for each action. Although in principle the goal

description may be any sentence of $\mathcal{L}_{tca}$, in practice we restrict the goal description to be any non-probabilistic sentence of $\mathcal{L}_{tca}$ involving only *HOLDS* and *OCCURS* predicates and temporal relations. Due to the potential complex nature of action interactions in the present framework, the description of action interactions will be specific to a given planning problem.

Throughout this section, we will be concerned with deriving a lower bound on the chance that a plan achieves a goal. So it will suffice to specify lower bounds on all probability values.

6.1.1. The planning environment

The planning environment is described in terms of the chances that facts hold and events occur:

$$P_{now}(HOLDS(FA, t_F, t'_F)) \geqslant \alpha,$$
$$P_{now}(OCCURS(EV, t_E, t'_E)) \geqslant \alpha,$$

as well as temporal constraints relating the times of facts and events:

$$(now < t_F \leqslant t_E).$$

6.1.2. Feasibility conditions

Action feasibility conditions *FC* are described by sentences of the form

$$P_{now}(FEAS(A, t_A) \mid FC) \geqslant \alpha,$$

where t_{FC} is the earliest time associated with *FC* and $t_{FC} < t_A$.

Feasibility conditions can be used to represent interference between actions. For example, the following sentence says that action A_2 is not feasible as long as A_1 is being performed.

$$P_{now}(FEAS(A_2, t_{A_2}) \mid OCC(A_1, t_{A_1}, t'_{A_1}) \wedge (t_{A_1} \leqslant t_{A_2} \leqslant t'_{A_1})) = 0.$$

6.1.3. Executability conditions

Executability conditions *EC* are described by sentences of the form

$$P_{now}(OCC(A, t_A, t'_A) \mid ATT(A, t_A) \wedge EC) \geqslant \alpha, \tag{16}$$

where the earliest time associated with *EC* is t_{EC}, and $t_{EC} < t'_A$. Executability conditions are assumed to be independent of the action with which they are associated:

IA1. $P_{now}(EC \mid ATT(A, t_A)) = P_{now}(EC)$.

So a lower bound for the chance that the action is executable can be expressed in terms of the executability specification (16) and the chance of the executability condition:

$$P_{now}(OCC(A, t_A, t'_A) \mid ATT(A, t_A)) \geqslant$$
$$P_{now}(OCC(A, t_A, t'_A) \mid ATT(A, t_A) \wedge EC) \cdot P_{now}(EC).$$

Assumption IA1 reduces the complexity of inference and is reasonable in most cases.

6.1.4. Effects and ramification conditions

We describe positive action effects by sentences of the form:

$$P_{now}(EFF \mid OCC(A, t_A, t'_A) \wedge RAM) \geqslant \alpha, \tag{17}$$

where *RAM* represents the ramification conditions for effect *EFF* of action *A*, the latest time associated with *EFF* is t_{EFF}, the earliest time associated with *RAM* is t_{RAM}, and $t_{RAM} < t'_{EFF}$. Ramification conditions are assumed to be independent of the action with which they are associated:

IA2. $P_{now}(RAM \mid OCC(A, t_A, t'_A)) = P_{now}(RAM)$.

So a lower bound for the chance of the action's positive effects can be expressed in terms of the positive effect specification (17) and the chance of the ramification condition:

$$P_{now}(EFF \mid OCC(A, t_A, t'_A)) \geqslant$$
$$P_{now}(EFF \mid OCC(A, t_A, t'_A) \wedge RAM) \cdot P_{now}(RAM).$$

6.1.5. The goal description

The goal description is any non-probabilistic sentence of $\mathcal{L}_{\text{tca}}$ involving only *HOLDS* and *OCCURS* predicates and temporal relations. For example, the goal of getting to the bank by 5:00pm may be represented as

$$\exists t_1, t_2\ (t_1 < 5{:}00) \wedge HOLDS(\text{loc(me,bank)}, t_1, t_2).$$

6.1.6. The probability of goal achievement

In order to compare alternative plans, we will be interested in inferring the chance that trying to attempt a given plan achieves a given goal. We speak of *trying* to attempt a plan since if the plan is not feasible it cannot be attempted. For a plan with n action attempts such that $t_{A_1} \leqslant t_{A_2} \leqslant \cdots \leqslant t_{A_n}$ a lower bound on the chance that trying to attempt the plan achieves goal G is the chance that the attempt of the plan brings about the goal and that the plan is feasible: [13]

$$P_{now}\left(G \mid \bigwedge_{i=1}^{n} ATT(A_i, t_{A_i})\right)$$
$$\cdot \prod_{i=1}^{n} P_{now}\left(FEAS(A_i, t_{A_i}) \mid \bigwedge_{j<i} ATT(A_j, t_{A_j})\right). \tag{18}$$

This is only a lower bound since the goal may come about even if some of the actions are not attempted.

[13] Haddawy [21] derives this expression from an expression for the expected utility of trying to attempt an action.

If the failed attempt of a plan has no chance of achieving our goal, we can provide a precise expression for the chance that trying to attempt the plan achieves the goal by focusing on the chance that trying to attempt the plan achieves both the occurrence of the plan and the goal. This is just the chance that the plan attempt brings about both the plan occurrence and the goal multiplied by the chance that the plan is feasible:

$$P_{now}\left(G \bigwedge_i OCC(A_i, t_{A_i}, t'_{A_i}) \mid \bigwedge_i ATT(A_i, t_{A_i})\right) \cdot \prod_{i=1}^{n} P_{now}\left(FEAS(A_i, t_{A_i}) \mid \bigwedge_{j<i} ATT(A_j, t_{A_j})\right). \tag{19}$$

6.1.7. Inferring the probability of a goal from action specifications

Describing the feasibility, executability, and effects of actions is only useful for planning purposes if these descriptions can be combined to infer the chance that a plan will achieve a given goal. We show here how these descriptions can be combined to infer the chance that attempting a single action will result in the action occurring and in a given goal being achieved. For plans consisting of multiple actions we need additional information concerning the interactions of the actions composing the plan. Since numerous types of action interactions are possible, we will give a specific example later of performing such inference for a more complex plan.

Suppose we wish to achieve goal G and suppose that G is an effect of action A. The chance that trying to attempt A results in A occurring and achieves G is

$$P_{now}(G \wedge OCC(A, t_A, t'_A) \mid ATT(A, t_A)) \cdot P_{now}(FEAS(A, t_A)).$$

By the definition of c-prob the first term can be written as

$$P_{now}(G \wedge OCC(A, t_A, t'_A) \mid ATT(A, t_A)) = P_{now}(G \mid OCC(A, t_A, t'_A) \wedge ATT(A, t_A)) \cdot P_{now}(OCC(A, t_A, t'_A) \mid ATT(A, t_A)).$$

By axiom ACT1,

$$P_{now}(G \mid OCC(A, t_A, t'_A) \wedge ATT(A, t_A)) = P_{now}(G \mid OCC(A, t_A, t'_A)).$$

So the chance of achieving goal G can be expressed in terms of the feasibility, executability, and effects of A:

$$P_{now}(G \wedge OCC(A, t_A, t'_A) \mid ATT(A, t_A)) \cdot P_{now}(FEAS(A, t_A)) = P_{now}(G \mid OCC(A, t_A, t'_A)) \cdot P_{now}(OCC(A, t_A, t'_A) \mid ATT(A, t_A)) \cdot P_{now}(FEAS(A, t_A)).$$

If action A has ramification, executability, and feasibility conditions associated with it then by independence assumptions IA1 and IA2 we have

$$P_{now}(G \wedge OCC(A, t_A, t'_A) \mid ATT(A, t_A)) \cdot P_{now}(FEAS(A, t_A)) \geqslant$$
$$P_{now}(G \mid OCC(A, t_A, t'_A) \wedge RAM) \cdot P_{now}(OCC(A, t_A, t'_A) \mid ATT(A, t_A) \wedge EC)$$
$$\cdot P_{now}(FEAS(A, t_A) \mid FC) \cdot P_{now}(RAM) \cdot P_{now}(EC) \cdot P_{now}(FC).$$

6.2. *Planning example*

This section presents a detailed example of the use of the representational framework in reasoning about plans. Suppose I am at home and would like to go to my favorite restaurant for dinner. The restaurant does not take reservations. Under normal circumstances, I can get a table within fifteen minutes but if a theater performance has ended in the last hour, the wait could be much longer. We would like to determine the chance that the plan consisting of starting my car and driving to the restaurant will result in having dinner without having to wait too long. A further complication is that my car is fairly unreliable and is only likely to start if the temperature is above freezing. The problem is described by specifying the feasibility, executability, and effects for the two actions, the interactions between the actions, the state of the planning environment, and the desired goal.

start(car)

Executability. I can usually start my car if the temperature is above freezing while I am trying to start it.

$$\forall t_s \ (t_s > now) \rightarrow$$
$$P_{now}(OCC(\text{start}(\text{car}), t_s, t_s + 1) \mid$$
$$ATT(\text{start}(\text{car}), t_s) \wedge HOLDS((\text{temp} > 32), t_s, t_s + 1)) \geqslant 0.95. \quad (20)$$

Feasibility. I can attempt to start my car if I have the keys and am at the same location as the car.

$$\forall t_s, t', x \ (t_s > now) \rightarrow$$
$$P_{now}(FEAS(\text{start}(\text{car}), t_s) \mid$$
$$HOLDS(\text{have}(\text{keys}), t', t_s) \wedge$$
$$HOLDS(\text{loc}(\text{me}, x), t', t_s) \wedge$$
$$HOLDS(\text{loc}(\text{car}, x), t', t_s)) = 1. \quad (21)$$

drive(home,restaurant)

Effects. There is at least an 80% chance that if I drive to the restaurant I will get a table within 15 minutes, as long as no theater performance ended within an hour of my arrival at the restaurant.

$$\forall t_d, t'_d\ P_{now}(\exists t, t'\ (t'_d \leqslant t \leqslant t'_d + 15) \wedge OCCURS(\text{get(table)}, t, t') \mid OCC(\text{drive(home,restaurant)}, t_d, t'_d) \wedge \neg\exists t_p, t'_p\ (t'_p < t'_d \leqslant t'_p + 60) \wedge OCCURS(\text{performance}, t_p, t'_p)) \geqslant 0.8. \quad (22)$$

Executability. I can successfully drive to the restaurant if I can first start my car. Notice that the executability condition is the occurrence of the action of starting my car.

$$\forall t, t_d\ (t_d > now) \rightarrow P_{now}(OCC(\text{drive(home,restaurant)}, t_d, t_d + 10) \mid ATT(\text{drive(home,restaurant)}, t_d) \wedge OCC(\text{start(car)}, t, t_d)) = 1. \quad (23)$$

Feasibility. I can attempt to drive from home to the restaurant if both I and my car are at home.

$$\forall t, t_d\ (t_d > now) \rightarrow P_{now}(FEAS(\text{drive(home,restaurant)}, t_d) \mid HOLDS(\text{loc(me,home)}, t, t_d) \wedge HOLDS(\text{loc(car,home)}, t, t_d)) = 1. \quad (24)$$

Action interactions

We assume that the actions of starting the car and driving to the restaurant do not negatively influence one another. This is represented by the following three sentences. Attempting to start the car does not negatively influence the feasibility of driving the car at a later time.

$$\forall t_d, t_s, x, y\ (t_s < t_d) \rightarrow P_{now}(FEAS(\text{drive}(x, y), t_d) \mid ATT(\text{start(car)}, t_s)) \geqslant P_{now}(FEAS(\text{drive}(x, y), t_d)). \quad (25)$$

Starting the car does not negatively influence the effects of driving to the restaurant at a later time.

$$\forall t_d, t'_d, t_s, t'_s\ (t'_s \leqslant t_d) \rightarrow P_{now}(\exists t, t'\ (t'_d \leqslant t \leqslant t'_d + 15) \wedge OCCURS(\text{get(table)}, t, t') \mid OCC(\text{drive(home,restaurant)}, t_d, t'_d) \wedge OCC(\text{start(car)}, t_s, t'_s)) \geqslant P_{now}(\exists t, t'\ (t'_d \leqslant t \leqslant t'_d + 15) \wedge OCCURS(\text{get(table)}, t, t') \mid OCC(\text{drive(home,restaurant)}, t_d, t'_d)). \quad (26)$$

Attempting to drive to the restaurant does not negatively influence the executability of starting the car at an earlier time.

$$\forall t_d, t_s, t_s' \ (t_s' \leqslant t_d) \rightarrow$$
$$P_{now}(OCC(\text{start(car)}, t_s, t_s') \mid$$
$$ATT(\text{start(car)}, t_s) \wedge ATT(\text{drive(home,restaurant)}, t_d)) \geqslant$$
$$P_{now}(OCC(\text{start(car)}, t_s, t_s') \mid ATT(\text{start(car)}, t_s)). \tag{27}$$

Planning environment

We have the following time line.

$$(now < t_0 < t_1). \tag{28}$$

There is an 80% the temperature will be above freezing this evening.

$$P_{now}(HOLDS((\text{temp} > 32), t_1, t_1 + 120)) = 0.8. \tag{29}$$

I am certain to have my keys this evening.

$$P_{now}(HOLDS(\text{have(keys)}, t_0, t_1)) = 1. \tag{30}$$

I am certain to be at home this evening. [14]

$$P_{now}(HOLDS(\text{loc(me,home)}, t_0, t_1 + 1)) = 1. \tag{31}$$

My car is likely to be at home this evening.

$$P_{now}(HOLDS(\text{loc(car,home)}, t_0, t_1 + 1)) \geqslant 0.95. \tag{32}$$

There is no theater performance this evening.

$$P_{now}(\exists t, t' \ (t' \leqslant t_1 + 120) \wedge OCCURS(\text{performance}, t, t')) = 0. \tag{33}$$

Goal

My goal is to get a table within thirty minutes of t_1.

$$\exists t_G, t_G' \ (t_1 < t_G \leqslant t_1 + 30) \wedge OCCURS(\text{get(table)}, t_G, t_G').$$

The derivation

We want to derive the chance that the plan consisting of starting my car at time t_1 and driving to the restaurant at time $t_1 + 1$ will occur and achieve the goal:

---

[14] Since my location is essentially within my control, this would more accurately be represented with a stay-at-location action:

$$\forall x, n, t \ P_{now}(OCC(\text{stay}(x, n), t, t + n) \mid ATT(\text{stay}(x, n), t)) = 1,$$
$$\forall x, n, t \ P_{now}(HOLDS(\text{loc(me}, x), t, t + n) \mid OCC(\text{stay}(x, n), t, t + n)) = 1,$$

where stay(x, n) means that I stay at location x for n time units. For simplicity of exposition we have omitted this action from the plan.

$$P_{now}(\exists t_G, t'_G\ (t_G \leqslant t_1 + 30) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \wedge OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1) \mid ATT(\text{drive(home,restaurant)}, t_1 + 1) \wedge ATT(\text{start(car)}, t_1)) \quad (34)$$

$$\cdot P_{now}(FEAS(\text{drive(home,restaurant)}, t_1 + 1) \mid ATT(\text{start(car)}, t_1)) \quad (35)$$

$$\cdot P_{now}(FEAS(\text{start(car)}, t_1)). \quad (36)$$

We calculate a lower bound on the chance of each of the terms (34), (35), and (36).

Calculation of (36)
By axiom FOL5 and (21),

$$P_{now}(FEAS(\text{start(car)}, t_1) \mid HOLDS(\text{have(keys)}, t_0, t_1) \wedge HOLDS(\text{loc(me,home)}, t_0, t_1) \wedge HOLDS(\text{loc(car,home)}, t_0, t_1)) = 1. \quad (37)$$

By Theorem 11 and (30), (31), and (32) it follows that

$$P_{now}(HOLDS(\text{have(keys)}, t_0, t_1) \wedge HOLDS(\text{loc(me,home)}, t_0, t_1) \wedge HOLDS(\text{loc(car,home)}, t_0, t_1)) \geqslant 0.95. \quad (38)$$

By the definition of c-prob and Theorem 10 applied to (37) and (38),

$$P_{now}(FEAS(\text{start(car)}, t_1)) \geqslant 0.95. \quad (39)$$

Calculation of (35)
By axiom FOL5 and (24),

$$P_{now}(FEAS(\text{drive(home,restaurant)}, t_1 + 1) \mid HOLDS(\text{loc(me,home)}, t_0, t_1 + 1) \wedge HOLDS(\text{loc(car,home)}, t_0, t_1 + 1)) = 1. \quad (40)$$

By Theorem 11 and (31) and (32),

$$P_{now}(HOLDS(\text{loc(me,home)}, t_0, t_1 + 1) \wedge HOLDS(\text{loc(car,home)}, t_0, t_1 + 1)) \geqslant 0.95. \quad (41)$$

By the definition of c-prob and Theorem 10 applied to (40) and (41),

$$P_{now}(FEAS(\text{drive(home,restaurant)}, t_1 + 1)) \geqslant 0.95. \quad (42)$$

By FOL5 and the field axioms applied to (42) and (25),

$$P_{now}(FEAS(\text{drive(home,restaurant)}, t_1 + 1) \mid ATT(\text{start(car)}, t_1)) \geqslant 0.95. \quad (43)$$

Calculation of (34)

By the definition of c-prob and axiom ACT1, term (34) may be rewritten as

$$P_{now}(\exists t_G, t'_G\ (t_G \leqslant t_1 + 30) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \wedge OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1) \mid ATT(\text{drive(home,restaurant)}, t_1 + 1) \wedge ATT(\text{start(car)}, t_1)) =$$

$$P_{now}(\exists t_G, t'_G\ (t_G \leqslant t_1 + 30) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \mid OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1)) \tag{44}$$

$$P_{now}(OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1) \mid ATT(\text{drive(home,restaurant)}, t_1 + 1) \wedge ATT(\text{start(car)}, t_1)). \tag{45}$$

Since $(t_1+11 \leqslant t_G \leqslant t_1+26) \rightarrow (t_G \leqslant t_1+30)$, by Theorem 10 we have the following inequality for term (44).

$$P_{now}(\exists t_G, t'_G\ (t_G \leqslant t_1 + 30) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \mid OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1)) \geqslant$$

$$P_{now}(\exists t_G, t'_G\ (t_1 + 11 \leqslant t_G \leqslant t_1 + 26) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \mid OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1)). \tag{46}$$

By assumption (26) and axiom FOL5, we have

$$P_{now}(\exists t_G, t'_G\ (t_1 + 11 \leqslant t_G \leqslant t_1 + 26) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \mid OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge OCC(\text{start(car)}, t_1, t_1 + 1)) \geqslant \tag{47}$$

$$P_{now}(\exists t_G, t'_G\ (t_1 + 11 \leqslant t_G \leqslant t_1 + 26) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \mid OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11)). \tag{48}$$

By the definition of c-prob, term (45) may be written as

$$P_{now}(OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \mid ATT(\text{drive(home,restaurant)}, t_1 + 1) \wedge OCC(\text{start(car)}, t_1, t_1 + 1)) \tag{49}$$

$$\cdot\ P_{now}(OCC(\text{start(car)}, t_1, t_1 + 1) \mid ATT(\text{start(car)}, t_1) \wedge ATT(\text{drive(home,restaurant)}, t_1 + 1)). \tag{50}$$

By assumption (27) and axiom FOL5, we have the following inequality for term (50).

$$P_{now}(OCC(\text{start(car)}, t_1, t_1+1) \mid ATT(\text{start(car)}, t_1) \wedge ATT(\text{drive(home,restaurant)}, t_1+1)) \geqslant P_{now}(OCC(\text{start(car)}, t_1, t_1+1) \mid ATT(\text{start(car)}, t_1)) \quad (51)$$

So we have the following inequality for term (34).

$$P_{now}(\exists t_G, t'_G\ (t_G \leqslant t_1+30) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \wedge OCC(\text{drive(home,restaurant)}, t_1+1, t_1+11) \wedge OCC(\text{start(car)}, t_1, t_1+1) \mid ATT(\text{drive(home,restaurant)}, t_1+1) \wedge ATT(\text{start(car)}, t_1)) \geqslant$$

$$P_{now}(\exists t_G, t'_G\ (t_1+11 \leqslant t_G \leqslant t_1+26) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \mid OCC(\text{drive(home,restaurant)}, t_1+1, t_1+11)) \quad (52)$$

$$P_{now}(OCC(\text{drive(home,restaurant)}, t_1+1, t_1+11) \mid ATT(\text{drive(home,restaurant)}, t_1+1) \wedge OCC(\text{start(car)}, t_1, t_1+1) \quad (53)$$

$$\cdot\ P_{now}(OCC(\text{start(car)}, t_1, t_1+1) \mid ATT(\text{start(car)}, t_1)). \quad (54)$$

Now we derive numerical bounds for the terms (52), (53), (54). First we derive a bound for term (52). By axiom FOL5 and (22),

$$P_{now}(\exists t, t'\ (t_1+11 \leqslant t \leqslant t_1+26) \wedge OCCURS(\text{get(table)}, t, t') \mid OCC(\text{drive(home,restaurant)}, t_1+1, t_1+11) \wedge \neg\exists t_p, t'_p\ (t'_p < t_1+11 \leqslant t'_p+60) \wedge OCCURS(\text{performance}, t_p, t'_p)) \geqslant 0.8. \quad (55)$$

By Theorem 10 and (33),

$$P_{now}(\neg\exists t, t'\ (t' < t_1+11 \leqslant t'+60) \wedge OCCURS(\text{performance}, t, t')) = 1. \quad (56)$$

So by the assumption that ramification conditions are independent of the action occurrence, the definition of c-prob, and (55) and (56),

$$P_{now}(\exists t, t'\ (t_1+11 \leqslant t \leqslant t_1+26) \wedge OCCURS(\text{get(table)}, t, t') \mid OCC(\text{drive(home,restaurant)}, t_1+1, t_1+11)) \geqslant 0.8. \quad (57)$$

Next we derive a value for term (53). By axiom FOL5 and (23),

$$P_{now}(OCC(\text{drive(home,restaurant)}, t_1+1, t_1+11) \mid ATT(\text{drive(home,restaurant)}, t_1+1) \wedge OCC(\text{start(car)}, t_1, t_1+1)) = 1. \quad (58)$$

Finally we derive a bound for term (54). By axiom FOL5 and (20),

$$P_{now}(OCC(\text{start(car)}, t_1, t_1 + 1) \mid ATT(\text{start(car)}, t_1) \wedge HOLDS((\text{temp} > 32), t_1, t_1 + 1)) \geqslant 0.95. \tag{59}$$

By Theorem 15 and (29),

$$P_{now}(HOLDS((\text{temp} > 32), t_1, t_1 + 1)) \geqslant 0.8. \tag{60}$$

So by the assumption that executability conditions are independent of the action attempt, the definition of c-prob, and (59) and (60),

$$P_{now}(OCC(\text{start(car)}, t_1, t_1 + 1) \mid ATT(\text{start(car)}, t_1)) \geqslant (0.8)(0.95) = 0.76. \tag{61}$$

From (56), (57), and (61) we obtain a lower bound on the probability of term (34) of

$$0.8 \cdot 1 \cdot 0.76 = 0.60. \tag{62}$$

Finally, combining the probability values from (39), (43), and (62) we obtain the lower bound on the chance that the plan achieves the goal:

$$\begin{aligned} &P_{now}(\exists t_G, t'_G\ (t_G \leqslant t_1 + 30) \wedge OCCURS(\text{get(table)}, t_G, t'_G) \wedge \\ &\qquad OCC(\text{drive(home,restaurant)}, t_1 + 1, t_1 + 11) \wedge \\ &\qquad OCC(\text{start(car)}, t_1, t_1 + 1) \mid \\ &\qquad ATT(\text{drive(home,restaurant)}, t_1 + 1) \wedge \\ &\qquad ATT(\text{start(car)}, t_1)) \\ &\cdot P_{now}(FEAS(\text{drive(home,restaurant)}, t_1 + 1) \mid ATT(\text{start(car)}, t_1)) \\ &\cdot P_{now}(FEAS(\text{start(car)}, t_1)) \geqslant \\ &0.60 \cdot 0.95 \cdot 0.95 = 0.54. \end{aligned} \tag{63}$$

6.3. *The qualification, frame, and ramification problems*

Probabilistic representations address the qualification problem by allowing one to summarize exceptions [42]. For example, we can reason about the chance that my car will start given that I turn the key, even though many events may prevent the car from starting. We simply say that the probability my car will start given that I turn the key is, say, 0.98. This probability summarizes the fact that low likelihood events, like a potato being stuck in my tailpipe, can prevent the car from starting. We can also say that the probability that my car will start given that I turn the key and that a potato is stuck in

the tailpipe is zero, without contradiction. The conditional probability appropriate to our state of knowledge would then be used at inference time.

The reader will notice that in the previous planning example we have avoided reasoning about the persistence of facts by formulating the planning problem in such a way that actions bring about conditions just when they are needed. The current work does not propose solutions to either the frame or ramification problems. In fact, the problems are magnified by the fact that actions and events may occur concurrently. So frame axioms would need to be predicated on the lack of occurrence of actions or events which could negate a condition. For example, to say that the color of my car does not change over an interval during which it is not painted, we might write

$$\forall t_0, t_1, t_2, x\ P_{now}(\exists t_3\ HOLDS(\text{color}(\text{car},x), t_2, t_3) \mid HOLDS(\text{color}(\text{car}, x), t_0, t_1) \wedge \neg OCC(\text{paint}(\text{car}), t_1, t_2)) = 1.$$

Several researchers have addressed the problem of projecting the effects of actions and plans in a probabilistic framework. Dean and Kanazawa [12] present an approach to projection using a Bayesian network model.[15] The tendency of propositions to remain true over time is modeled with conditional probability statements that indicate the probability that a proposition P is true at time t conditioned on whether or not P was true at the previous time point and whether or not an event known to make P true or false occurred in the interim. A Bayesian network model is created which contains a node for each proposition and event of interest at every one of a set of discrete time points. Projection is performed using one of the standard inference algorithms for Bayesian networks.

Hanks [28] presents an approach to reasoning about projection in which knowledge of the world is represented in terms of probabilities that propositions are true at various time points, but not necessarily at every time point as in Dean and Kanazawa's approach. The dynamics of the world is modeled with causal rules that describe the tendency of a proposition to change state as a result of an event and with persistence rules that describe the chance that a proposition changes state over an interval during which no causally relevant event is known to occur. The projector answers queries of the form "is the probability that ϕ will hold at time t greater than τ?" The algorithm is particularly distinguished by its efficiency, which it gains by searching only for past evidence relevant to the query and by making only relevant distinctions when projecting effects forward in time. In contrast, Dean and Kanazawa's approach computes the probability of every proposition at every point in time.

More recent work in this area has discussed appropriate ways to structure Bayesian networks in order to obtain a compact representation that is sufficiently expressive to reason about projection [8,9].

[15] The framework is also described by Dean and Wellman [13].

7. Related work

7.1. Theories of objective chance

Three outstanding subjective theories of objective chance are those of van Fraassen [57], Lewis [39], and Skyrms [52]. Van Fraassen's model of objective chance is more constrained than Lewis' model which is more constrained than Skyrms' model. Thus, in van Fraassen's model, chance has more inherent properties than in either Lewis' or Skyrms' models. Since van Fraassen's theory is the only one of the three that is cast in a temporal framework, it was used as the point of departure for the model of objective chance in $\mathcal{L}_{\text{tca}}$.

Van Fraassen presents a semantic theory that models subjective probability and objective chance, using a future-branching model of time points. Van Fraassen places two constraints on objective chance:

(1) The chance of a past is either 0 or 1, depending on whether or not it actually occurred.
(2) Chance at a time is completely determined by the history of the world up to that time.

From these assumptions, he shows the following relation between subjective probability and objective chance

$$P_t(X \mid Y) = E_Y[C_t(X)],$$

where P_t is the subjective probability at time t, C_t is the objective chance at time t, E_Y is the expected value given Y, and provided the truth of Y depends only on the history up to t. This relation entails a version of Miller's principle relating subjective probability and objective chance. A similar relation holds between objective chances at different times, but van Fraassen does not demonstrate this. In van Fraassen's models, objective chance can change with time but, in contrast to the present work, truth values cannot. He does not provide a logical language for his theory.

7.2. Temporal probability logics

The theory of computing literature contains several examples of logics that can represent both time and probability [27,29,38]. The focus of these logics is on reasoning about probabilistic programs and distributed systems. The logics do not attempt to model causality or to distinguish between different types of temporal objects such as facts and events; hence, they are not suitable for reasoning about actions and plans.

Kanazawa [31] presents a logic, $\mathcal{L}_{\text{cp}}$, of time and probability. The language allows quantification over time points but is propositional otherwise. The language contains a P operator for representing probability. The P operator is not temporally indexed so the language cannot represent the change of probability over time. Furthermore, the language does not allow nesting of probability operators. Like $\mathcal{L}_{\text{tca}}$, $\mathcal{L}_{\text{cp}}$ contains numeric functions for representing probability distributions. Because Kanazawa's focus is on representing the tendency of facts to persist through time, his representation of facts and events differs from ours. He describes a fact as "something that once it becomes true tends

to stay true for some time". Facts are associated with temporal intervals. Events take place instantaneously and are associated with a single time point. He distinguishes three types of events: persistence causation events are associated with a fact becoming true; persistence termination events are associated with a fact becoming false; point events are facts that stay true for only an instant. $\mathcal{L}_{cp}$ represents actions as events.

Like the models for $\mathcal{L}_{tca}$, Kanazawa's models contain a collection of world-histories. He does not impose a branching time structure on the histories. The semantics of the P operator are defined in terms of probability distributions over world-histories. Fact tokens and event tokens are equated with ⟨temporal interval, world⟩ pairs.

Dean and Wellman [13] present a propositional temporal probability logic similar to Kanazawa's. They extend Shoham's [50] propositional temporal logic by introducing a probability operator. The probability operator cannot be nested but can otherwise be combined freely with other logical operators. The probability operator is not temporally indexed. A model contains of a universe of time points, a set of possible time lines, a binary relation over time points, and a discrete probability measure over the time lines. The models differ from those for $\mathcal{L}_{tca}$ by not imposing a branching structure on the time lines nor other constraints on the elements of the models.

7.3. Temporal logics of actions and plans

Shoham [51] presents a branching time logic that formalizes the relation between time, knowledge, and action. Knowledge is defined in the standard way for modal logic and actions are defined as the ability to make choices among sets of world-histories. The model formalizes the notion that actions can only be performed under certain conditions. Shoham's logic does not capture the fact that actions cannot influence the past.

Pelavin [43,44] develops a future-branching time logic for reasoning about planning problems involving concurrent actions and external events. He starts with Allen's [1] linear temporal logic of time intervals and extends it with two modal operators, *INEV* and *IFTRIED*, to reason about future-branching time and action effects, respectively. *INEV* is exactly our $\Box$ operator. *IFTRIED* is a counterfactual operator that associates the attempt of an action with the truth of a sentence. The semantics of the operator are based on Stalnaker's and Lewis' theories of counterfactuals. *IFTRIED* captures the temporal relation of action and effect—an action cannot affect the state of the world at any time preceding its attempt.

Pelavin represents actions and plans uniformly as "plan instances". A plan instance is an ordered pair: a set of basic action instances and a set of event instances, brought about by the basic actions. An action is attempted if its basic action instances occur and it occurs if its basic action instances and its event instances all occur. Since a plan instance is a single ordered pair and the times associated with the event and basic action instances are fixed and all terms in the language are rigid designators, a plan instance cannot occur more than once in a world-history. If we want to allow two instances of the same plan to occur in a world-history, they must have different names. Pelavin creates these different names by allowing an interval to be associated with a plan instance e.g. $pi@I$. This is to be interpreted as saying that plan instance pi occurs during interval I. But he does not provide a semantic definition for the @ operator

to formalize the intuitive interpretation. Furthermore, by using different names for two occurrences of the same plan he has no way of saying that two instances of the *same* plan occur.

Conditional chance is used in this paper to represent effects. Since chance can be conditioned on any sentence, $\mathcal{L}_{tca}$ can represent the effects of a wide variety of different phenomena. In Pelavin's logic, *IFTRIED* only associates effects with action attempts, so effects of general events cannot be represented. The use of conditional chance eliminates the need for a separate counterfactual operator and its semantic counterpart: the similarity measure over worlds. Furthermore, our representation of actions has a more intuitively appealing semantics. Unlike Pelavin, we do not assume that actions are always feasible. Such an assumption is unacceptable since it is hard to imagine what it would mean for the basic action instances of two inconsistent actions, like concurrently remaining still and moving, to all occur.

Pelavin's use of a Stalnaker/Lewis counterfactual operator to represent action effects results in some undesirable inferences. Consider the sentence "If I were to attempt to stay at home all day, then if I were to attempt to walk from home to the store at noon, I would succeed in walking from home to the store at noon." This sentence can be represented in Pelavin's logic by nesting two *IFTRIED* operators. If being at home is the executability condition for going from home to the store, Pelavin points out that the sentence is true in his logic. But attempting to stay home all day and attempting to go to the store at noon are incompatible. So it seems unreasonable for this sentence to be true in the intended interpretation[16]. In $\mathcal{L}_{tca}$ attempting to stay home all day and attempting to go the store at noon are incompatible, so their conjunction would not be feasible, and hence going to the store would have zero chance of occurring given that one was staying home all day.

Skyrms provides an elegant probabilistic account of counterfactuals based on the notion of objective chance [52, Chapter IIA] and discusses the semantics of iterated probability conditionals [52, Appendix 3]. He shows that the iterated probability conditionals can lead to more intuitive inferences than iterated Stalnaker/Lewis conditionals.

7.4. Decision-theoretic planners

A number of researchers have recently been working on the problem of building decision-theoretic planning sytems. This work can be classified into two main schools. The first school attempts to take techniques from classical AI planning and to apply them within a decision-theoretic framework. The second school assumes that planning problems are describable as Markov processes and attempts to formulate efficient and flexible algorithms for solving them.

Kushmerick et al. [35, 36] present the BURIDAN planner, which is roughly a probabilistic version of the SNLP planning algorithm. They characterize the world in terms of a probability distribution over propositions and actions in terms of transitions from one

[16] One might think that the sentence should be correct because going to the store is clipping staying at home, but the same problem arises if the temporal order of the actions is reversed.

distribution to another. The BURIDAN planner generates a plan that achieves a goal with a probability no less than a user-specified threshold. Draper et al. [15] have extended the algorithm to handle information gathering and contingent actions.

Haddawy and Suwandi [25] present a decision-theoretic refinement planning system (DRIPS). The DRIPS planner finds optimal plans for problems described in terms of a probability function over world states, a probabilistic model of actions, and a temporal utility function. DRIPS can reason about both discrete and continuous attributes, as well as metric time. Action descriptions are organized into an abstraction hierarchy. DRIPS reasons efficiently by using the hierarchy to prune away suboptimal classes of plans. Haddawy and Doan [23] describe the theory of abstraction used by the planner.

Wellman's SUDO-PLANNER system [61,62] eliminates classes of suboptimal plans in domains characterized by partially satisfiable goals and actions with uncertain effects. It eliminates only those classes of plans which it can prove are dominated without resorting to reasoning about tradeoffs. Planning knowledge is represented in the form of qualitative probabilistic networks. Plans are functions from observations to actions. The planner works by cycling between the processes of model construction and dominance proving. Model construction involves constructing a qualitative probabilistic network from a more general knowledge base for the domain. Dominance proving involves using knowledge about the effects of actions and the relative desirability of outcomes to derive facts about the preferred plan.

Drummond and Bresina [16] present an algorithm for generating plans in stochastic temporal domains. They represent the dynamics of the domain as a discrete Markov process, where actions and exogenous events are characterized by transition probabilities. Goals are temporally qualified sentences. The algorithm works by first projecting a single path that satisfies the goal. In generating this path the system explores only likely outcomes, ignoring those of low probability. This produces a plan with a lower bound probability of achieving the goal. The probability of goal satisfaction can be increased by elaborating the plan with additional paths that satisfy the goal.

Dean et al. [11] build upon the work of Drummond and Bresina by providing their basic planning algorithm with theoretically sound probabilistic foundations. They present a planning algorithm in which the world is modeled as a stochastic automaton consisting of a finite set of states; actions are represented by transition probabilities between states; goals are represented by reward functions over the states; and plans are policies that specify what action to perform in each state. The algorithm starts by generating a policy for a simplified world model. It then iteratively expands the world model and generates the optimal policy for each expanded model. The expected values of generated plans improve as a function of computation time and the algorithm can be consulted at any time after the initial policy is generated to obtain the best policy so far. In a second paper, Dean et al. [10] discuss decision-theoretic methods for allocating processor time to policy generation and world model expansion in time-critical domains.

Boutilier and Dearden [14] investigate planning in time-critical domains represented as Markov decision processes. They trade optimality for inference speed by limiting search depth and using a heuristic function to estimate the values of states. Actions are executed as the plan is being constructed. In a second paper [4] they present a method

for abstracting the state space of a Markov decision process so that approximately optimal policies can be constructed. They prove bounds on the loss of optimality due to the use of the abstraction.

Cassandra et al. [6] show how to formulate a planning problem in which the agent has incomplete information about the state of the environment as a partially observable Markov decision process. They present an algorithm for solving these problems that is empirically more efficient than existing algorithms.

Thiebaux et al. [56] represent planning problems using probabilistic logic and use exhaustive search to construct Markov models of plans from this representation. They then evaluate the models to determine the expected value of each plan.

8. Conclusion

8.1. Contributions

This paper has presented a framework for representing planning problems that integrates logical and probabilistic approaches. By drawing on the strength of previous work in both areas, we have been able to create a novel synthesis that addresses limitations of the traditional planning paradigm under each approach. The contributions made by this work fall into three main areas: the development of a vocabulary for describing planning problems, the development of a semantic theory that captures desired intuitive properties, and the extension of the probabilistic framework to accommodate actions that may not be feasible.

8.1.1. Vocabulary

We have provided a vocabulary for describing planning problems involving time and chance. The language can express the chance of temporally qualified conditions in the world as well as the chance of action feasibility, executability, and effects. We can represent the extent to which conditions in the world can and cannot be influenced. This is an important capability since a planner must be able to reason about what conditions it can influence and the extent to which it can influence them. We can represent temporal aspects of plans such as concurrent actions and conditions during an action that influence its executability and effects. The language distinguishes between actions and events. This distinction is important since an agent has much more direct control over its actions than over events. The language further distinguishes between action attempts and action occurrences. This distinction allows action duration to vary as a function of conditions in the world and it allows us to define the notion of action feasibility, the importance of which will be discussed below. The language allows first-order quantification over time points, probabilities, and domain individuals. This produces a language with great representational economy. For example, by quantifying over domain individuals we can describe classes of actions and by quantifying over time points we can describe states-of-affairs that are valid at each of some range of time points. No other currently existing logical language can represent these aspects of time, chance, and action.

8.1.2. Semantic theory

We have specified a set of constraints on the model theory that assigns meaning to the logical language. A constrained logic is desirable for several reasons. First, the more constrained the model theory, the greater the predictive power of the logic. In the extreme case, we could constrain the logic to allow only one model, thus giving us complete knowledge. But we do not want to overconstrain the logic so that we eliminate models that are consistent with our conception of reality in the intended domains of applicability. Second, constraints provide guidance in assigning probabilities to sentences. This is an important function since the assigning of priors is a notoriously difficult task. So one would like a logic that is as constrained as possible yet does not produce unwarranted inferences. The constraints we have imposed on the models capture numerous intuitive properties of time and chance in such a way that natural inferences follow directly from the semantics. For example, the following properties are a consequence of the constraints.

- The property that facts have higher chance of holding over their subintervals, (Theorem 15), is a consequence of semantic constraint (C3) and the definition of probability.
- The property that the past cannot be influenced (axioms IT3–IT6 and Theorems 17–23) follows from constraints (C1) and (C2) on the accessibility relation, constraint (C5) relating the accessibility relation to facts and events, and constraints (C6) and (C7) on probability.
- Miller's principle (axiom P3), which relates chance over time is a consequence of constraints (C6) and (C7). Furthermore, as a consequence of Miller's principle we have the property that chance is the expected value of future chance (Theorem 12).
- The property that inevitability implies certainty (axiom IP1) follows from semantic constraints (C6) and (C7).

8.1.3. Feasibility

The integration of both chance and possibility in $\mathcal{L}_{tca}$ allowed us to define the notion of the chance of feasibility of an action or plan. This concept was then used to define the chance that trying to attempt an action will achieve a given goal. This definition is important in reasoning about plans because it enables us to compose individual action descriptions into descriptions of plans involving multiple actions. Since actions may interfere in such a way that their composition may not be feasible, we must be able to reason about the effects of possibly infeasible plans.

Appendix A. Soundness proofs

We prove the soundness of the more intersting and less commonly known of the axioms in section 4. The first few proofs we present in detail. The remaining proofs are presented less formally.

I1. $\Box_t\phi \rightarrow \phi$.

Proof. We prove this sentence valid by showing that it is satisfied by an arbitrary model M, world w, and assignment function $g[d/t]$. By the semantic definitions,

$$[\![\Box_t\phi \rightarrow \phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff}$$
$$[\![\neg \Box_t \phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{or} \quad [\![\phi]\!]^{M,w,g[d/t]} = \text{true}.$$

The first disjunct is true iff for some w' such that $R(d,w,w')$ $[\![\phi]\!]^{M,w',g[d/t]} = \text{false}$. If this is not the case then ϕ is true in all worlds w' such that $R(d,w,w')$. And since by (C2) $R(d,w,w)$, it follows that ϕ is true in w. □

I2. $\Box_t(\phi \rightarrow \psi) \rightarrow (\Box_t\phi \rightarrow \Box_t\psi)$.

Proof. By the semantic definitions,

$$[\![\Box_t(\phi \rightarrow \psi) \rightarrow (\Box_t\phi \rightarrow \Box_t\psi)]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff}$$
$$[\![\neg \Box_t (\phi \rightarrow \psi)]\!]^{M,w,g[d/t]} = \text{true} \quad \text{or} \quad [\![\Box_t\phi \rightarrow \Box_t\psi]\!]^{M,w,g[d/t]} = \text{true}.$$

By the semantic definitions, this is true iff

$$[\![\Box_t\phi \wedge \neg\psi]\!]^{M,w,g[d/t]} = \text{false} \quad \text{or}$$
$$[\![\Diamond_t\phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{or}$$
$$[\![\Box_t\psi]\!]^{M,w,g[d/t]} = \text{true}.$$

If the first disjunct is not the case then in every world we have $\neg\phi \vee \psi$. So in every world ϕ is false or ψ is true. If ϕ is false in some world, the second disjunct is satisfied and we're done. Otherwise, ψ must be true in every world, in which case the third disjunct is satisfied. □

I3. $\Box_t\phi \rightarrow \Box_t \Box_t \phi$.

Proof. By the semantic definitions,

$$[\![\Box_t\phi \rightarrow \Box_t \Box_t \phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff}$$
$$[\![\Box_t\phi]\!]^{M,w,g[d/t]} = \text{false} \quad \text{or} \quad [\![\Box_t \Box_t \phi]\!]^{M,w,g[d/t]} = \text{true}.$$

If the first disjunct is not the case then ϕ is true in all worlds w' such that $R(d,w,w')$. Now

$$[\![\Box_t \Box_t \phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff} \quad [\![\phi]\!]^{M,w,g[d/t]} = \text{true}$$

for all w'' such that $R(d,w',w'')$ and $R(d,w,w')$. But by (C2) this is true iff $[\![\phi]\!]^{M,w,g[d/t]} = \text{true}$ for all w' such that $R(d,w,w')$. □

I4. $\Diamond_t\phi \rightarrow \Box_t \Diamond_t \phi$.

Proof. By the semantic definitions and the definition of possibility,

$$[\![\Diamond_t\phi \rightarrow \Box_t \Diamond_t \phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff}$$
$$[\![\Box_t\neg\phi]\!]^{M,w,g[d/t]} = \text{true} \quad \text{or} \quad [\![\Box_t \Diamond_t \phi]\!]^{M,w,g[d/t]} = \text{true}.$$

If the first disjunct is not the case then ϕ is true in some world w° such that $R(d, w, w^\circ)$. Now $[\![\Box_t \Diamond_t \phi]\!]^{M,w,g[d/t]} = \text{true}$ iff for every world w' such that $R(d, w, w')$ there exists a world w'' such that $R(d, w, w'')$ and $[\![\phi]\!]^{M,w,g[d/t]} = \text{true}$. But by (C2) $R(d, w, w^\circ)$ and $R(d, w, w')$ implies $R(d, w', w^\circ)$. So w° is the required world w''. □

NEC. Rule of necessitation: from ϕ infer $\Box_t\phi$.

Proof.

$$[\![\Box_t\phi]\!]^{M,w,g} = \text{true} \quad \text{iff} \quad [\![\phi]\!]M, w', g[d/t] = \text{true}$$

for all w' such that $R(d, w, w')$. By definition, if ϕ is valid then $[\![\phi]\!]^{M,w,g}$ =true for all M, w, g. □

P3. Miller's principle: $(t_1 \preceq t_2) \rightarrow P_{t_1}(\phi \mid P_{t_2}(\phi) \geqslant \alpha) \geqslant \alpha$.

Proof. We first prove an expected value property and then use it to prove Miller's principle. Let t and t' be two time points $t \leqslant t'$ and consider the R-equivalence classes of worlds at time t'. Let the variable r range over these equivalence classes. The r form a partition of W, so the probability of a set X can be written as the integral over this partition:

$$\mu_t^w(X) = \int_{r \subset W} \mu_t^w(X \mid r)\mu_t^w(\mathrm{d}r).$$

Since the history up to time t' determines the probability at time t', this can be written as

$$\mu_t^w(X) = \int_{r \subset W} \mu_{t'}^r(X)\mu_t^w(\mathrm{d}r),$$

where $\mu_{t'}^r$ denotes the probability at time t' in equivalence class r. Since the probability at a given time is assumed to be constant over all worlds in an R-equivalence class, the probability at a given time is the expected value of the probability at any future time:

$$\mu_t^w(X) = \int_W \mu_{t'}^{w'}(X)\mu_t^w(\mathrm{d}w').$$

Next we show that Miller's principle is valid in the probability models. By the expected value property,

$$\mu_t^w(X \cap \{w' : \mu_{t'}^{w'}(X) = \alpha\}) = \int_W \mu_{t'}^{w''}(X \cap \{w' : \mu_{t'}^{w'}(X) = \alpha\})\mu_t^w(\mathrm{d}w'').$$

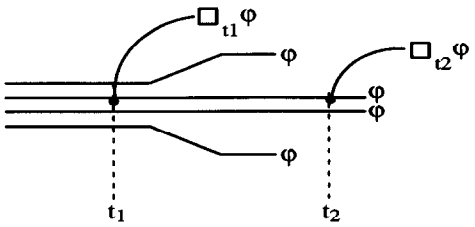

Fig. A.1. Model for the proof that inevitability persists.

Now, by semantic constraints (C6) and (C7) it follows that

$$\forall w \in \{w' : \mu_{t'}^{w'}(X) = \alpha\}, \quad \mu_{t'}^{w}(\{w' : \mu_{t'}^{w'}(X) = \alpha\}) = 1$$
$$\forall w \notin \{w' : \mu_{t'}^{w'}(X) = \alpha\}, \quad \mu_{t'}^{w}(\{w' : \mu_{t'}^{w'}(X) = \alpha\}) = 0.$$

So we can restrict the integral to the set $\{w' : \mu_{t'}^{w'}(X) = \alpha\}$:

$$= \int_{\{w' : \mu_{t'}^{w'}(X) = \alpha\}} \mu_{t'}^{w''}(X \cap \{w' : \mu_{t'}^{w'}(X) = \alpha\}) \mu_t^w(\mathrm{d}w'').$$

And by the above property again $\mu_{t'}^{w''}(X \cap \{w' : \mu_{t'}^{w'}(X) = \alpha\}) = \alpha$, so

$$= \alpha \cdot \int_{\{w' : \mu_{t'}^{w'}(X) = \alpha\}} \mu_t^w(\mathrm{d}w''),$$
$$= \alpha \cdot \mu_t^w(\{w' : \mu_{t'}^{w'}(X) = \alpha\}).$$

By the semantic definitions it follows that

$$P_t(\phi \wedge P_{t'}(\phi) = \alpha) = \alpha \cdot P_t(P_{t'}(\phi) = \alpha).$$

And by a slight generalization of the proof it follows that

$$\forall (t \preceq t') \quad P_t(\phi \wedge P_{t'}(\phi) \geqslant \alpha) \geqslant \alpha \cdot P_t(P_{t'}(\phi) \geqslant \alpha). \qquad \square$$

TL1. Facts hold over their subintervals.

$$(t_1 \preceq t_2 \preceq t_3 \preceq t_4) \wedge (t_1 \neq t_3) \wedge (t_2 \neq t_4) \rightarrow$$
$$[HOLDS(FA, t_1, t_4) \rightarrow HOLDS(FA, t_2, t_3)].$$

Proof. This follows directly from constraint (C3). □

IT1. Inevitability persists.

$$(t_1 \preceq t_2) \rightarrow (\Box_{t_1}\phi \rightarrow \Box_{t_2}\phi).$$

Proof. The model described in the proof is shown in Fig. A.1.

$$[\![\Box_{t_1}(\phi)]\!]^{M,w',g[d_1/t_1,d_2/t_2]} = \text{true}$$

iff for all w' such that $R(d_1, w, w')$

$$[\![\phi]\!]^{M,w',g[d_1/t_1,d_2/t_2]} = \text{true}.$$

Now suppose that for some w'' such that $R(d_2, w, w'')$, $[\![\phi]\!]^{M,w'',g[d_1/t_1,d_2/t_2]} = \text{false}$. Then by (C1), $R(d_1, w, w'')$. This is a contradiction.

So for all w'' such that $R(d_2, w, w'')$, $[\![\phi]\!]^{M,w'',g[d_1/t_1,d_2/t_2]} = \text{true}$. □

IT3. Past facts are inevitable.

$$(t_0 \preceq t_1 \preceq t_2) \rightarrow [\Box_{t_2} HOLDS(Q, t_0, t_1) \vee \Box_{t_2} \neg HOLDS(Q, t_0, t_1)].$$

Proof.

$$[\![\Box_{t2} HOLDS(Q, t_0, t_1) \vee \Box_{t_2} \neg HOLDS(Q, t_0, t_1)]\!]^{M,w,g[d_0/t_0,d_1/t_1,d_2/t_2]} = \text{true} \quad \text{iff}$$
$$[\![HOLDS(Q, t_0, t_1)]\!]^{M,w',g[d_0/t_0,d_1/t_1,d_2/t_2]} = \text{true}$$

for all w' such that $R(d_2, w, w')$ or

$$[\![HOLDS(Q, t_0, t_1)]\!]^{M,w',g[d_0/t_0,d_1/t_1,d_2/t_2]} = \text{false}$$

for all w' such that $R(d_2, w, w')$. This is the case iff

$$\langle\langle d_0, d_1\rangle, w'\rangle \in F(Q) \quad \text{for all } w' \text{ such that } R(d_2, w, w') \text{ or}$$
$$\langle\langle d_0, d_1\rangle, w'\rangle \notin F(Q) \quad \text{for all } w' \text{ such that } R(d_2, w, w').$$

This last statment follows directly from (C5). □

IP1. Inevitability implies certainty: $\Box_t(\phi) \rightarrow P_t(\phi) = 1$.

Proof. We prove this sentence valid by showing that it is satisfied by an arbitrary model M, world w, and assignment function $g[d/t]$. By the semantic definitions,

$$[\![\Box_t(\phi) \rightarrow P_t(\phi) = 1]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff}$$
$$[\![\Diamond_t(\neg\phi)]\!]^{M,w,g[d/t]} = \text{true} \quad \text{or} \quad [\![P_t(\phi) = 1]\!]^{M,w,g[d/t]} = \text{true}.$$

The first disjunct holds if for some w' such that $R(d, w, w')$ $[\![\phi]\!]^{M,w',g[d/t]} = \text{false}$. If this is not the case then $[\![\phi]\!]^{M,w',g[d/t]} = \text{true}$ for all w' such that $R(d, w, w')$. So $R_d^w \subseteq \{w' : [\![\phi]\!]^{M,w,g[d/t]} = \text{true}\}$. But by Meta-Theorem 2, $\mu_d^w(R_d^w) = 1$. So $\mu_d^w(\{w' : [\![\phi]\!]^{M,w',g[d/t]} = \text{true}\}) = 1$. And from the semantic definitions it follows that

$$[\![P_t(\phi) = 1]\!]^{M,w,g[d/t]} = \text{true}. \quad \square$$

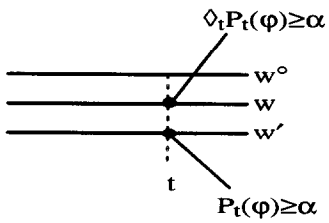

Fig. A.2. Model for the proof that current chance is inevitable.

IP2. Current chance is inevitable: $\Diamond_t P_t(\phi) \geqslant \alpha \rightarrow \Box_t P_t(\phi) \geqslant \alpha$.

Proof. The model described in the proof is shown in Fig. A.2.

$$[\![\Diamond_t P_t(\phi) \geqslant \alpha]\!]^{M,w,g[d/t]} = \text{true} \quad \text{iff} \quad [\![P_t(\phi) \geqslant \alpha]\!]^{M,w,g[d/t]} = \text{true}$$

for some w' such that $R(d,w,w')$. This is the case iff $\mu_d^{w'}(\{w'' : [\![\phi]\!]^{M,w'',g[d/t]} = \text{true}\}) \geqslant \alpha$ for some w' such that $R(d,w,w')$. Choose an arbitrary w° such that $R(d,w,w^\circ)$. By (C2), $R(d,w',w^\circ)$. Hence by (C6), $\mu_d^{w^\circ}((\{w'' : [\![\phi]\!]^{M,w'',g[d/t]} = \text{true}\}) \geqslant \alpha$. □

Acknowledgements

This paper is a greatly truncated and slightly revised version of my Ph.D. dissertation, completed in August 1991 at the University of Illinois. The dissertation also appears as a monograph [22]. In that work the logic is developed and presented in the full context of decision theory.

I would like to thank my thesis advisor Alan Frisch for generously sharing with me his time and ideas and for carefully reading through many drafts. I thank Patrick Maher for guiding me through the subtleties of decision theory and for painstakingly reading through my proofs. Any remaining errors are my responsibility. Thanks to the other members of my dissertation committee: Caroline Hayes, George Monahan, and Marianne Winslett for their many helpful comments and discussions. Thanks to an anonymous reviewer for comments on an earlier draft.

This work was partially supported by the author's Shell doctoral dissertation fellowship (University of Illinois) and by NSF grant IRI-9207262 (University of Wisconsin-Milwaukee).

References

[1] J.F. Allen, Towards a general theory of action and time, *Artif. Intell.* **23** (2) (1984) 123–154.

[2] J.F. Allen and J.A. Koomen, Planning using a temporal world model, in: *Proceedings IJCAI-83*, Karlsruhe, Germany (1983) 741–747.

[3] F. Bacchus, *Representing and Reasoning with Probabilistic Knowledge* (MIT Press, Cambridge, MA, 1990).

[4] C. Boutilier and R. Dearden, Using abstractions for decision-theoretic planning with time constraints, in: *Proceedings AAAI-94*, Seattle, WA (1994) 1016–1022.

[5] R. Carnap, *Logical Foundations of Probability* (University of Chicago Press, Chicago, IL, 1950).
[6] A.R. Cassandra, L.P. Kaelbling and M.L. Littman, Acting optimally in partially observable stochastic domains, in: *Proceedings AAAI-94*, Seattle, WA (1994) 1023–1028.
[7] D. Chapman, Planning for conjunctive goals, *Artif. Intell.* **32** (3) (1987) 333–377.
[8] A. Darwiche and M. Goldszmidt, Action networks: a framework for reasoning about actions and change under uncertainty, in: *Proceedings Tenth Conference on Uncertainty in Artificial Intelligence*, Seattle, WA (1994) 136–144.
[9] R. Davidson and M.R. Fehling, A structured, probabilistic representation of action, in: *Proceedings Tenth Conference on Uncertainty in Artificial Intelligence*, Seattle, WA (1994) 154–161.
[10] T.L. Dean, L.P. Kaelbling, J. Kirman and A. Nicholson, Deliberation scheduling for time-critical sequential decision making, in: *Proceedings Ninth Conference on Uncertainty in Artificial Intelligence*, Washington, DC (1993) 309–316.
[11] T.L. Dean, L.P. Kaelbling, J. Kirman and A. Nicholson, Planning with deadlines in stochastic domains, in: *Proceedings AAAI-93*, Washington, DC (1993) 574–579.
[12] T.L. Dean and K. Kanazawa, A model for reasoning about persistence and causation, *Comput. Intell.* **5** (3) (1989) 142–150.
[13] T.L. Dean and M.P. Wellman, *Planning and Control* (Morgan Kaufmann, San Mateo, CA, 1991).
[14] R. Dearden and C. Boutilier, Integrating planning and execution in stochastic domains, in: *Proceedings Tenth Conference on Uncertainty in Artificial Intelligence*, Seattle, WA (1994) 162–169.
[15] D. Draper, S. Hanks and D.S. Weld, Probabilistic planning with information gathering and contingent execution, in: *Proceedings Second International Conference on Artificial Intelligence Planning Systems*, Chicago, IL (1994) 31–36.
[16] M. Drummond and J. Bresina, Anytime synthetic projection: maximizing the probability of goal satisfaction, in: *Proceedings AAAI-90*, Boston, MA (1990) 138–144.
[17] R. Fagin, J.Y. Halpern and N. Megiddo, A logic for reasoning about probabilities, Tech. Rept. RJ 6190 (60900), IBM Almaden Research Center (1988).
[18] R.E. Fikes and N.J. Nilsson, STRIPS: a new approach to the application of theorem proving to problem solving, *Artif. Intell.* **2** (1971) 189–208.
[19] A.I. Goldman, *A Theory of Human Action* (Prentice Hall, Englewood Cliffs, NJ, 1970).
[20] A.R. Haas, Possible events, actual events, and robots, *Comput. Intell.* **1** (1985) 59–70.
[21] P. Haddawy, Representing plans under uncertainty: a logic of time, chance, and action, Ph.D. Thesis, Report No. UIUCDCS-R-91-1719, University of Illinois, Urbana, IL (1991).
[22] P. Haddawy, *Representing Plans under Uncertainty: A Logic of Time, Chance, and Action*, Lecture Notes in Artificial Intelligence **770** (Springer-Verlag, Berlin, 1994).
[23] P. Haddawy and A.H. Doan, Abstracting probabilistic actions, in: *Proceedings Tenth Conference on Uncertainty in Artificial Intelligence*, Seattle, WA (1994) 270–277; available via anonymous FTP from `pub/tech_reports` at `ftp.cs.uwm.edu`.
[24] P. Haddawy and A.M. Frisch, Modal logics of higher-order probability, in: R.D. Shachter T.S. Levitt, J.F. Lemmer and L.N. Kanal, eds., *Uncertainty in Artificial Intelligence* **4** (Elsevier Science Publishers, Amsterdam, 1990) 133–148.
[25] P. Haddawy and M. Suwandi, Decision-theoretic refinement planning using inheritance abstraction, in: *Proceedings Second International Conference on AI Planning Systems* (1994) 266–271; available via anonymous FTP from `pub/tech_reports` at `ftp.cs.uwm.edu`.
[26] J.Y. Halpern, An analysis of first-order logics of probability, *Artif. Intell.* **46** (1991) 311–350.
[27] J.Y. Halpern and M.R. Tuttle, Knowledge, probability, and adversaries, in: *Proceedings Eighth Annual ACM Symposium on Principles of Distributed Computing* (1989) 103–118.
[28] S. Hanks and D. McDermott, Modeling a dynamic and uncertain world I: symbolic and probabilistic reasoning about change, *Artif. Intell.* **66** (1) (1994) 1–55.
[29] S. Hart and M. Shair, Probabilistic temporal logics for finite and bounded models, in: *Proceedings Sixteenth ACM Symposium on the Theory of Computing*, Washington, DC (1984) 1–13.
[30] G.E. Hughes and M.J. Cresswell, *An Introduction to Modal Logic* (Methuen, London, 1968).
[31] K. Kanazawa, A logic and time nets for probabilistic inference, in: *Proceedings AAAI-91*, Anaheim, CA (1991) 360–365.
[32] J. Keynes, *A Treatise on Probability* (MacMillan, London, 1921).

[33] A. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea, New York, 1950).
[34] S. Kripke, Semantical considerations on modal logic, *Acta Philos. Fenn.* **16** (1963) 83–94 (Proceedings of a Colloquium on Modal and Many-Valued Logics, Helsinki, Finland, 1962).
[35] N. Kushmerick, S. Hanks and D.S. Weld, An algorithm for probabilistic planning, *Artif. Intell.* **76** (1995) 239–286.
[36] N. Kushmerick, S. Hanks and D.S. Weld, An algorithm for probabilistic least-commitment planning, in: *Proceedings of the Twelfth National Conference on Artificial Intelligence*, Seattle (1994) 1073–1078.
[37] H. Kyburg Jr and H. Smokler, eds., *Studies in Subjective Probability* (Wiley, New York, 1964).
[38] D. Lehmann and S. Shelah, Reasoning with time and chance, *Inform. Control* **53** (1982) 165–198.
[39] D. Lewis, A subjectivist's guide to objective chance, in: W. Harper, R. Stalnaker and G. Pearce, eds., *Ifs* (Reidel, Dordrecht, Netherlands, 1980) 267–298.
[40] D.V. McDermott, A temporal logic for reasoning about processes and plans, *Cognitive Sci.* **6** (1982) 101–155.
[41] D.H. Mellor, Fixed past, unfixed future, in: B.M. Taylor, ed., *Michael Dummett* (Martinus Nijhoff, Dordrecht, Netherlands, 1987) 166–186.
[42] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, San Mateo, CA, 1988).
[43] R.N. Pelavin, A formal logic for planning with concurrent actions and external events, Ph.D. Thesis, University of Rochester, Department of Computer Science, Rochester, NY (1988).
[44] R.N. Pelavin, Planning with simultaneous actions and external events, in: *Reasoning about Plans* (Morgan Kaufmann, San Mateo, CA, 1991) Chapter 3, 128–211.
[45] R.N. Pelavin and J.F. Allen, A formal logic of plans in temporally rich domains, *Proc. the IEEE* **74** (10) (1986) 1364–1382.
[46] F.P. Ramsey, Truth and probability, in: D.H. Mellor, ed., *Foundations* (Humanities Press, Atlantic Highlands, NJ, 1926) Chapter 3, 58–100 (Collection published 1978).
[47] E.D. Sacerdoti, A structure for plans and behavior, Technical Note 109, SRI, Menlo Park, CA (1975).
[48] L.J. Savage, *The Foundations of Statistics* (Wiley, New York, 1954, 2nd rev. ed., 1972).
[49] J.R. Shoenfield, *Mathematical Logic* (Addison-Wesley, Reading, MA, 1967).
[50] Y. Shoham, *Reasoning about Change* (MIT Press, Cambridge, MA, 1987).
[51] Y. Shoham, Time for action: on the relationship between time, knowledge, and action, in: *Proceedings IJCAI-89*, Detroit, MI (1989) 954–959.
[52] B. Skyrms, *Causal Necessity* (Yale University Press, New Haven, CT, 1980).
[53] B. Skyrms, Higher order degrees of belief, in: D.H. Mellor, ed., *Prospects for Pragmatism* (Cambridge University Press, Cambridge, 1980) Chapter 6, 109–137.
[54] P. Suppes, *A Probabilistic Theory of Causality*, Acta Philosophica Fennica, Fasc. **XXIV** (North-Holland, Amsterdam, 1970).
[55] G.J. Sussman, *A Computer Model of Skill Acquisition* (American Elsevier, New York, 1975).
[56] S. Thiebaux, J. Hertzberg, W. Shoaff and M. Schneider, A stochastic model of actions and plans for anytime planning under uncertainty, *Internat. J. Intelligent Systems* (1994).
[57] B.C. van Fraassen, A temporal framework for conditionals and chance, in: W. Harper, R. Stalnaker and G. Pearce, eds., *Ifs* (Reidel, Dordrecht, Netherlands, 1980) 323–340.
[58] J. Venn, *The Logic of Chance* (MacMillan, London, 1866); new paperback edition: Chelsea, New York, 1962.
[59] R. von Mises, *Probability, Statistics and Truth* (Allen and Unwin, London, 1957).
[60] D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavioral Research* (Cambridge University Press, Cambridge, 1986).
[61] M.P. Wellman, *Formulation of Tradeoffs in Planning Under Uncertainty* (Pitman, London, 1990).
[62] M.P. Wellman, Qualitative probabilistic networks for planning under uncertainty, in: G. Shafer and J. Pearl, eds., *Readings in Uncertain Reasoning* (Morgan Kaufmann, San Mateo, CA, 1990) Chapter 9, 711–722.
[63] A.N. Whitehead and B.A.W. Russell, *Principia Mathematica* (Cambridge University Press, Cambridge, 1910).