Challenges for Temporal Planning with Uncertain Durations

Mausam and **Daniel S. Weld**
Dept of Computer Science and Engineering
University of Washington
Seattle, WA-98195
{mausam,weld}@cs.washington.edu

Abstract

We investigate the problem of temporal planning with concurrent actions having stochastic durations, especially in the context of extended-state-space based planners. The problem is challenging because stochastic durations lead to an explosion in the *space of possible decision-epochs*, which exacerbates the familiar challenge of growth in executable action combinations caused by concurrency.

Introduction

Recent progress in temporal planning (JAIR Special Issue 2003) raises hopes that this technology may soon apply to a wide range of real-world problems. However, concurrent actions with stochastic durations characterise many real-world domains. While both concurrency and duration uncertainty have independently received some attention by planning researchers, very few systems have addressed them in concert, and all of these systems have used an extended-state-space method (*vs.* a constraint-posting approach). In this paper we step back from specific algorithms and analyse the broader problem of concurrent temporal planning with actions having stochastic durations, especially in the context of extended-state-space planners.

We find that the problem is challenging in novel ways and opens interesting avenues for future research. The stochastic durations lead to an explosion in the *space of possible decision-epochs*, which exacerbates the familiar challenge of growth in executable action combinations caused by concurrency. The rate of decision-epoch growth increases with greater expressiveness in the action language, and we characterise the challenges along several dimensions: 1) the possible times for which action preconditions and effects may be specified, 2) the degree of uncertainty (deterministic, effects and durations controlled by independent distributions, and an expressive language allowing correlations between durations and effect outcomes), 3) continuous duration distributions, and more.

Expressiveness of Action Models

The action models handled by different temporal planners vary in complexity. Figure 1 lists different representations along two dimensions (ignoring continuous change). The simplest temporal model is used in TGP (Smith & Weld 1999). *TGP-style* actions require preconditions to be true throughout execution; the effects are guaranteed to be true only after termination; and actions may not execute concurrently if they clobber each other's preconditions or effects.

	Simple	Boundary	Metric
Deterministic duration	TGP	$PDDL_{2.1}$	Zeno
Prob. but independent	Prob. TGP	Prob. $PDDL_{2.1}$	
Joint distrib: dur ×effects			Prottle

Figure 1: *Action models* for temporal planning (ignoring continuous change). The horizontal axis varies the times at which preconditions and effects may be specified. The vertical axis varies the uncertainty in effects and its correlations with durations.

Along the horizontal axis, we vary the temporal expressiveness in the precondition and effect representations. $PDDL_{2.1}$ (Fox & Long 2003) is more expressive than TGP's representation as it can represent preconditions that are required to be true just at start, over whole action execution or just at the end. Where $PDDL_{2.1}$ allows effects to appear only at boundaries, Zeno's representation (Penberthy & Weld 1994) allows effects (preconditions) to appear at arbitrary intermediate points (and intervals).

Along the vertical axis, we vary the representation of uncertainty in the model. $PDDL_{2.1}$ doesn't support probabilistic action effects or durations. "Probabilistic $PDDL_{2.1}$" extends $PDDL_{2.1}$ along this direction, associating a distribution with each action duration; the distribution for durations is independent of that for effects. "Probabilistic TGP" extends the TGP action representation similarly. Even more expressive representations may use a single joint distribution — enabling action durations that are correlated with effects. Indeed, the representation language of Prottle (Little, Aberdeen, & Thiebaux 2005) contains all these features: effects at intermediate points, action durations correlated with probabilistic effects. Tempastic (Younes & Simmons 2004a) uses probabilistic TGP-style actions, but because it also supports exogenous events, it is at least as expressive as Prottle. The blank entries in Figure 1 denote action languages that have not yet been discussed in the literature.

Planning with TGP-style Actions

We first study TGP-style actions in the context of uncertain durations. We find that planning, even with these simplified action models, suffers additional computational blowup. All

Copyright © 2006, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

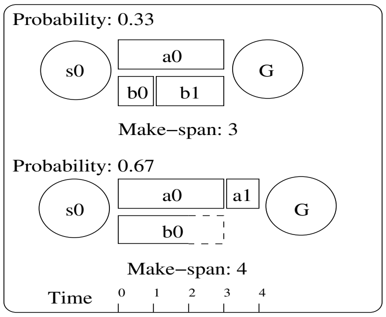

Figure 2: Planning with expected durations leads to a sub-optimal solution.

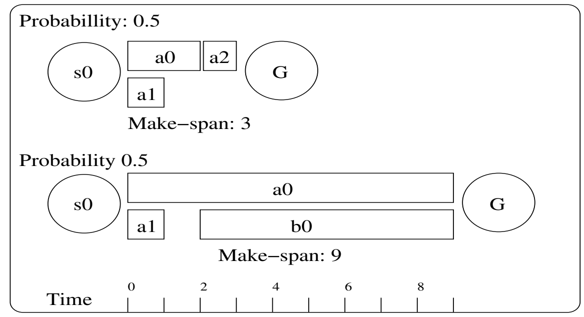

Figure 3: Intermediate decision epochs are necessary for optimal planning.

the examples in this section apply to problems regardless of whether effects are deterministic or stochastic. We investigate extensions to richer representations in the next sections.

We focus on problems whose objective is to achieve a goal state, while minimising total expected time (*make-span*), but our observations extend to cost functions that combine make-span and resource usage. This raises the question of *when* a goal counts as achieved. We require that all executing actions terminate before the goal is considered achieved.

A naive way to solve our problem is by ignoring duration distributions. We can assign each action a constant duration equal to the mean of its distribution, and then apply a deterministic-duration planner such as that of Mausam and Weld (2005). Unfortunately, this method may not produce an optimal policy as the following example illustrates.

Example: Consider the planning domain in Figure 2, in which the goal can be reached in two independent ways — executing the plan $\langle a_0; a_1 \rangle$, *i.e.*, a_0 followed by a_1, or the plan $\langle b_0; b_1 \rangle$. Let a_0, a_1 and b_1 have constant durations 3, 1, and 2 respectively. Let b_0 have a uniform distribution between lengths 1, 2 and 3. It is clear that if we disregard b_0's duration distribution and replace it by the mean 2, then both these plans have an expected cost of 4. However, the truly optimal plan has duration 3.67 — start both a_0 and b_0; if b_0 finishes at time 1 (prob. 0.33) then start b_1, otherwise (prob. 0.67) wait until a_0 finishes and execute a_1 to reach the goal. In such a policy, expected cost to reach the goal is 0.33×3 + 0.67×4 = 3.67. Thus for optimal solutions, we need to explicitly take duration uncertainty into account. □

Definition *Any time point when a new action is allowed to start execution is called a* decision epoch*. A* happening *is either 0 or a time when an action actually terminates.*

For TGP or prob. TGP-style actions with det. durations, restricting decision epochs to happenings suffices for optimal planning (Mausam & Weld 2005). Unfortunately, the same is not true for problems with duration uncertainty.

Temporal planners may be classified as having one of two architectures: constraint-posting approaches, in which the times of action execution are gradually constrained during planning (*e.g.*, Zeno and LPG (Penberthy & Weld 1994; Gerevini & Serina 2002)), and extended-state-space methods (*e.g.*, TP4 and SAPA (Haslum & Geffner 2001; Do & Kambhampati 2001)). The following example has important computational implications for state-space planners, because limiting attention to a subset of decision epochs can speed these planners.

Example: Consider Figure 3, in which the goal can be reached in two independent ways — executing both $\{a_0, a_1\}$ followed by a_2 (*i.e.* effects of both a_0 and a_1 are preconditions to a_2); or by executing action b_0. Let a_1, a_2, and b_0 have constant durations 1, 1, and 7 respectively. Suppose that a_0 finishes in 2 time units with 0.5 probability and in 9 units the other half of the time. Furthermore, b_0 is mutex with a_1, but no other pairs of actions are mutex.

In such a domain, following the first plan, *i.e.*, $\langle \{a_0, a_1\}; a_2 \rangle$, gives an expected cost of $6.5 = 0.5 \times 2 + 0.5 \times 9 + 1$. The second plan ($\langle b_0 \rangle$) costs 7. The optimal solution, however, is to first start both a_0 and a_1 concurrently. When a_1 finishes at time 1, wait until time 2. If a_0 finishes, then follow it with a_2 (total length 3). If at time 2, a_0 doesn't finish, start b_0 (total length 9). The expected cost of this policy is $6 = 0.5 \times 3 + 0.5 \times 9$. □

In this *e.g.*, notice that the optimal policy needs to start action b_0 at time 2, even when there is no happening at 2. Thus limiting the set of decision epochs to happenings does not suffice for optimal planning with uncertain durations. It is quite unfortunate that non-happenings are potentially necessary as decision epochs, because even if one assumes that time is discrete, there are many interior points during a long-running action; must a planner consider them all?

Definition *An action has* independent duration *if there is no correlation between its probabilistic effects and its duration. An action has* monotonic continuation *if the expected time until action termination is nonincreasing during execution.*

Actions without probabilistic effects have independent duration. Actions with monotonic continuations are common, *e.g.*, those with uniform, exponential, Gaussian, and many other duration distributions. However, actions with bi- or multi-modal distributions don't have monotonic continuations.

We believe that if all actions have independent duration and monotonic continuation, then the set of decision epochs may be restricted to happenings without sacrificing optimality; this idea can be exploited to build a fast planner (Mausam & Weld 2006).

Timing Preconditions & Effects

Many domains require more flexibility concerning the times when preconditions and effects are in force: different effects of actions may apply at different times within the action's

```
:action a
   :duration 4
   :condition (over all P) (at end Q)
   :effect (at end Goal)
:action b
   :duration 2
   :effect (at start Q) (at end (not P))
```

Figure 4: A domain to illustrate that an expressive action model may require arbitrary decision epochs for a solution. In this example, b needs to start at 3 units after a's execution to reach $Goal$.

execution, preconditions may be required only to hold for part of execution, and executing two actions concurrently might lead to different results than executing them sequentially. Note that the decision epoch explosion is even more pronounced for such problems. Moreover, this not only affects optimality, but also affects the completeness of the algorithms. The following example with *deterministic* durations demonstrates this further.

Example: Consider the deterministic temporal planning domain in Figure 4 that uses $\text{PDDL}_{2.1}$ notation. If the initial state is P=true and Q=false, then the only way to reach $Goal$ is to start a at time 0, and b at time 3. Clearly, no action could terminate at 3, still it is a necessary decision epoch. □

Intuitively, two actions may require a certain relative alignment within them to achieve the goal. This alignment may force an action to start somewhere in the midst of the other's execution thus requiring a lot of decision epochs to be considered.

This example clearly shows that additional complexity in planning is incurred due to a more expressive action representation. It has important repercussions on existing planners. For instance, popular planners like SAPA and Prottle will not be able to solve this simple problem, because they consider only a restricted set of decision epochs. This shows that both these planners are incomplete (*i.e.*, problems may be incorrectly deemed unsolvable). Indeed, these planners can be naively modified by considering each time point as a decision epoch to obtain a complete algorithm. Unfortunately, such a modification is bound to be ineffective in scaling to any reasonable sized problem. Intelligent sampling of decision epochs is, thus, the key to finding a good balance between the two. The exact modalities of such an algorithm is an important open research problem.

Continuous Action Durations

Previously, we assumed that an action's possible durations are taken from a discrete set. We now investigate the effects of dealing directly with continuous uncertainty. Let $f_i^T(t)dt$ be the probability of action a_i completing between times $t+T$ and $t+T+dt$, if we know that action a_i did not finish until time T. Similarly, define $F_i^T(t)$ to be the probability of the action finishing *after* time $t+T$.

Example: Consider the extended state $\langle X, \{(a_1, T)\}\rangle$, which denotes that action a_1 started T units ago in the world state X. Let a_2 be an applicable action that is started in this extended state. Define $M = \min(\Delta_M(a_1) - T, \Delta_M(a_2))$, where Δ_M denotes the maximum possible duration of execution for each action. Intuitively, M is the time by which at least one action will complete. Also, let J_n and Q_n denote the n^{th} revision to the expected cost to reach a goal starting from a state or a state-action pair respectively (Mausam & Weld 2005). Q_n may be computed as follows:

$$Q_{n+1}(\langle X, \{(a_1,T)\}\rangle, a_2) = \int_0^M f_1^T(t)F_2^0(t)\left[t + J_n(\langle X_1, \{a_2, t\}\rangle)\right]dt + \int_0^M F_1^T(t)f_2^0(t)\left[t + J_n(\langle X_2, \{a_1, t+T\}\rangle)\right]dt \quad (1)$$

Here X_1 and X_2 are world states obtained by applying the deterministic actions a_1 and a_2 respectively on X. Recall that $J_{n+1}(s) = \min_a Q_{n+1}(s,a)$. For a fixed point computation of this form, we desire that J_{n+1} and J_n have the same functional form[1]. Going by the equation above this seems very difficult to achieve, except perhaps for very specific action distributions in some special planning problems. For example, if all distributions are constant or if there is no concurrency in the domain, then these equations are easily solvable. But for anything mildly interesting, solving these equations is a challenging open question.

Non-Monotonic Duration Distributions

Dealing with continuous multi-modal distributions worsens the decision epochs explosion. We illustrate this below.

Example: Consider the domain of Figure 3 except that let action a_0 have a bi-modal distribution, the two modes being uniform between 0-1 and 9-10 respectively as shown in Figure 5(a). Also let a_1 have a very small duration. Figure 5(b) shows the expected remaining termination times if a_0 terminates at time 10. Notice that due to bi-modality, this time increases between 0 and 1. The expected time to reach the goal using plan $\langle\{a_0, a_1\}; a_2\rangle$ is shown in the third graph.

Now suppose, we have started $\{a_0, a_1\}$, and we need to choose the next decision epoch. It is easy to see that the optimal decision epoch could be any point between 0 and 1 and would depend on the alternative routes to the goal. For example, if duration of b_0 is 7.75, then the optimal time-point to start the alternative route is 0.5 (right after the expected time to reach the goal using first plan exceeds 7.75). □

We have shown that the choice of decision epochs depends on the expected durations of the alternative routes. But these values are not known in advance, in fact these are the ones being calculated in the planning phase. Therefore, choosing decision epochs ahead of time does not seem possible. This makes the optimal continuous multi-modal distribution planning problem mostly intractable for any reasonable sized problem.

Correlated Durations and Effects

When actions' durations are correlated with the effects, then failure to terminate provides additional information regarding an action's effects. For example, non-termination at a point may change the probability of its eventual effects, and this may prompt new actions to be started. Thus, these points need to be considered for decision epochs, and cannot be omitted for optimal planning, even with TGP-style actions.

[1]This idea has been exploited in order to plan with continuous resources (Feng *et al.* 2004).

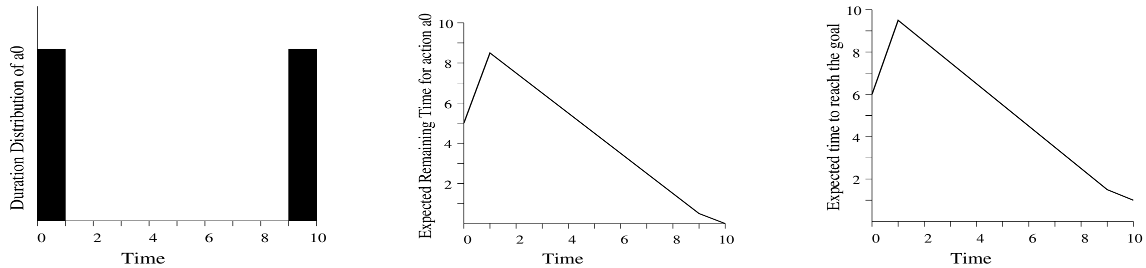

Figure 5: If durations are continuous (real-valued) rather than discrete, there may be an infinite number of potentially important decision epochs. In this domain, a crucial decision epoch could be required at any time in $(0, 1]$ — depending on the length of possible alternate plans.

Notion of Goal Satisfaction

Different problems may require slightly different notions of when a goal is reached. For example, we have assumed thus far that a goal is not "officially achieved" until all executed actions have terminated. Alternatively, one might consider a goal to be achieved if a satisfactory world state is reached, even though some actions may be in the midst of execution. There are intermediate possibilities in which a goal requires some *specific* actions to necessarily end.

Interruptible Actions

We have assumed that, once started, an action cannot be terminated. However, a richer model may allow preemptions, as well as the continuation of an interrupted action. The problems, in which all actions could be interrupted at will, have a significantly different flavour. To a large extent, planning with these is similar to finding different concurrent paths to the goal and starting all of them together, since one can always interrupt all the executing paths as soon as the goal is reached. For instance, example in Figure 3 no longer holds since b_0 can be started at time 1, and later terminated as needed to shorten the make-span.

Conclusions

This paper investigates planning problems with concurrent actions having stochastic durations, focussed primarily on extended-state-space planners. We characterise the space of such problems and identify the explosion in the number of decision epochs. No longer can a planner limit action-initiation times to points when a different action has terminated. The rate of decision-epoch growth increases with greater expressiveness in the action language, and we characterise the challenges along several dimensions.

Even with simple prob. TGP-style actions, many more decision epochs must be considered to achieve optimality. However, if all durations are unimodal and uncorrelated with effects, we conjecture that one can bound the space of decision epochs in terms of times of action terminations.

We show that for $\text{PDDL}_{2.1}$ and richer action representations, the currently-employed extended-state-space based methods are incomplete, and the straightforward ways to ensure completeness are highly inefficient. Developing an algorithm that achieves the best of both worlds is an important research question.

Additionally, we discuss the challenges posed by continuous time, observing that techniques employing piecewise constant/linear representations, which are popular in dealing with functions involving continuous variables, may be ineffective for our problem. These techniques rely on the same functional forms for successive approximations of the value function — and this does not hold in our case. Many other models form potent directions for future research: *e.g.*, multi-modal distributions, interruptible actions, and correlated action durations and effects.

Acknowledgments

We thank Sumit Sanghai for theorem proving skills and advice. We also thank Jiun-Hung Chen, Nilesh Dalvi, Maria Fox, Jeremy Frank, Subbarao Kambhampati, and Håkan Younes for helpful suggestions. Raphael Hoffman, Daniel Lowd, Tina Loucks, Alicen Smith and the anonymous reviewers gave useful comments on prior drafts. This work was supported by NSF grant IIS-0307906, ONR grants N00014-02-1-0932, N00014-06-1-0147 and the WRF / TJ Cable Professorship.

References

Do, M. B., and Kambhampati, S. 2001. Sapa: A domain-independent heuristic metric temporal planner. In *ECP'01*.

Feng, Z.; Dearden, R.; Meuleau, N.; and Washington, R. 2004. Dynamic programming for structured continuous Markov decision processes. In *UAI'04*, 154.

Fox, M., and Long, D. 2003. PDDL2.1: An extension to PDDL for expressing temporal planning domains. *JAIR Special Issue on 3rd International Planning Competition* 20:61–124.

Gerevini, A., and Serina, I. 2002. LPG: A planner based on local search for planning graphs with action graphs. In *AIPS'02*, 281.

Haslum, P., and Geffner, H. 2001. Heuristic planning with time and resources. In *ECP'01*.

2003. Special Issue on the 3rd International Planning Competition, JAIR, Volume 20.

Little, I.; Aberdeen, D.; and Thiebaux, S. 2005. Prottle: A probabilistic temporal planner. In *AAAI'05*.

Mausam, and Weld, D. 2005. Concurrent probabilistic temporal planning. In *ICAPS'05*, 120–129.

Mausam, and Weld, D. 2006. Probabilistic temporal planning with uncertain durations. *Submitted for publication*.

Penberthy, J., and Weld, D. 1994. Temporal planning with continuous change. In *AAAI'94*, 1010.

Smith, D., and Weld, D. 1999. Temporal graphplan with mutual exclusion reasoning. In *IJCAI'99*, 326–333.

Younes, H. L. S., and Simmons, R. G. 2004a. Policy generation for continuous-time stochastic domains with concurrency. In *ICAPS'04*.

Younes, H. L. S., and Simmons, R. G. 2004b. Solving generalized semi-markov decision processes using continuous phase-type distributions. In *AAAI'04*, 742.