

# Robust Execution of Deterministic Plans in Non-deterministic Environments

Oscar Lima,<sup>1,2</sup> Michael Cashmore,<sup>3</sup> Daniele Magazzeni,<sup>4</sup> Andrea Micheli,<sup>5</sup> and Rodrigo Ventura<sup>1</sup>

<sup>1</sup>Institute for Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal

<sup>2</sup>DFKI German Research Center for Artificial Intelligence, Germany

<sup>3</sup>University of Strathclyde, UK

<sup>4</sup>King's College London, UK

<sup>5</sup>Fondazione Bruno Kessler, Trento, Italy

## Abstract

In order to ensure the robust execution of a deterministic plan, execution must be adaptable to unexpected situations in the world and to exogenous events. This is critical in domains in which committing to a wrong ordering of actions can cause the plan failure, even when all the actions succeed. We propose an approach for the execution of temporal plans that permits some adaptability to unexpected situations of the environment (non-determinism) while maintaining the validity of the plan through online reasoning.

Our approach computes an adaptable, partially-ordered plan from a given totally-ordered plan. The partially-ordered plan is adaptable in that it can exploit beneficial differences between the world and what was expected. The approach is general in that it can be used with any task planner that produces either a totally or a partially-ordered plan. We propose a plan execution algorithm that computes online the complete set of valid totally-ordered plans described by an adaptable partially-ordered plan together with the probability of success for each of them. This set is then used to choose the next action to execute.

## 1 Introduction

Robust task plan execution is a fundamental problem in the intersection of AI planning and Robotics: the execution of the planned course of actions in the real world may differ from what was expected at planning time. A classical example of this discrepancy is the duration of actions that often depend upon external factors, impossible to model in the planning domain. In addition, plans are usually generated under the assumption of a static world that does not change without performing an action. In reality, however, the world in which the robot operates is often dynamic and comprises exogenous events, i.e. events that are not under the control of the agent and may happen unexpectedly. Such events can interfere with the planned course of actions and therefore we need to monitor the execution by means of observations and possibly adapt the decisions to cope with such contingencies.

Nonetheless, in many situations, minor adjustments to the plan can be sufficient to retain validity with respect to the

ground truth and reach the plan objective. For example, several techniques have been devised to absorb small variations in action durations. This is not the case when the planning problem exhibits temporal deadlines, time-windows or synchronizations, because minor delays could impact the success of the plan. This is particularly relevant when concurrency is considered, e.g. in multi-robot domains, because there can be positive and negative interactions between parallel actions. Hence, ordering constraints that arise from the coordination of these actions must be considered at execution time.

In this paper, we propose a novel flow from AI Planning to action execution aiming at the following research question:

Since observations during execution may differ from what was expected at planning time (including action duration and propositional state), is the plan valid and what is the next execution choice that maximizes probability of reaching the goal?

Our approach (described in 3) starts from a totally-ordered plan by extracting an adaptable partially-ordered plan as an offline step. Differently from other approaches, we allow some causal constraints<sup>1</sup> to be violated in order to allow for a stronger run-time adaptation. Then, we define an online algorithm that, given an estimation of the probabilities of each planning variable, analyzes all the valid totally-ordered plans induced by the adaptable partially-ordered plan, associating a probability of success to each of them. In turn, this set of totally-ordered plans is used by a novel action selection policy to choose the next action to execute that maximizes the probability of achieving the planning goal during execution. This execution flow is extremely flexible because, depending on the observations, it allows for dynamic re-ordering of the planned set of actions as well as the skipping of actions that might no longer be needed.

In 4 we describe how the approach is integrated into the planning and execution framework ROSPlan (Cashmore et al. 2015) and we empirically demonstrate in simulation that our approach leads to a consistently fewer re-plannings, and

<sup>1</sup>A causal constraint between actions  $a_1$  and  $a_2$  with respect to a proposition  $p$  is established when an effect  $p$  asserted by  $a_1$  is a precondition for  $a_2$ .

results in fewer actions executed. Moreover, we show how, despite being theoretically demanding in terms of performance, the whole technique can be implemented to be a fast action selection policy for practical use-cases.

## Related Work

There has been considerable work in the literature concerning the robust execution of plans. Some authors proposed ways to increase the flexibility of temporal plans to cope with more situations at runtime (Bäckström 1998; Frank and Morris 2007; Do and Kambhampati 2003); others devised techniques to synthesize correct-by-construction flexible plans (Ghallab and Laruelle 1994; Frank and Jónsson 2003; Cesta et al. 2009; Umbrico et al. 2018). In this paper, we relax a fundamental constraint that has been at the base of these previous works: we break the causal structure of the plan by discarding causal constraints in order to allow for more run-time adaptability. The obtained plan admits executions that are invalid for the planning model, therefore we employ a runtime action selection policy that dynamically selects actions that are causally-valid and are likely to reach the goal. This effectively moves the causal reasoning online instead of limiting the executor to blindly follow the causal structure prescribed by the plan.

The authors of (Levine and Williams 2018) show that a Temporal Plan Network Under Uncertainty is an encoding of a set of many different candidate STN (Dechter, Meiri, and Pearl 1991) sub-plans. They define a *correct execution* as an ordering of activities that is causally complete (each event’s conditions are satisfied) and temporally consistent with respect to the STN. Our approach is strongly related to this idea: we solve the problem of generating the complete set of *valid executions* for an adaptable partially-ordered plan. A valid execution is a *correct execution* which also achieves the goal.

This problem is also described in (Kim, Williams, and Abramson 2001) in the form of selecting an execution of a Temporal Plan Network compiled from RMPL, and in (Cimatti, Micheli, and Roveri 2016) in the form of synthesizing a dynamically controllable strategy for a disjunctive temporal network with uncertainty. During execution, we tackle a similar problem, without considering decision nodes, and accounting for the selected execution’s probability of success. In addition, our approach allows the plan execution to adapt to some unexpected observations.

Uncontrollable temporal durations can be also addressed via strong controllability (Karpas et al. 2015). Instead, (Cashmore et al. 2019) proposes a technique that given a plan parameterized with  $n$  temporal durations and domain constants automatically generates an  $n$ -dimensional region corresponding to a valid execution. The latter work generates a region over real-valued parameters for which the plan remains valid. Our work differs because we generate the set of total orderings for which the plan is valid and attach to each of these a probability of success; this copes with discrepancies in the discrete state that are not covered by (Cashmore et al. 2019).

## 2 Background

We start by formalizing our definition of a planning problem and a plan using the PDDL2.1 formalism (Fox and Long 2003).

**Definition 1.** A *planning problem* is a tuple  $\langle P, V, A, I, G \rangle$  where  $P$  is a set of propositions;  $V$  is a set of real variables;  $A$  is a set of durative and instantaneous actions;  $I$  is the total function describing the initial state of the propositions and real variables;  $G$  is the function indicating the goal condition.

**Definition 2.** A *durative action*  $a$  is a tuple  $\langle pre(a), eff(a), dur(a) \rangle$  where  $pre(a)$  is a set of conditions partitioned into *at-start*, *over-all* and *at-end* conditions;  $eff(a)$  is the set of action effects; and  $dur(a)$  is the duration constraint. An *instantaneous action* is a tuple  $\langle pre(a), eff(a) \rangle$  where  $pre(a)$  is the set of preconditions; and  $eff(a)$  is the set of action effects.

A durative (resp. instantaneous) action is *applicable* in a state  $S = P \cup V$  if the at-start condition (resp. precondition) of the action is satisfied by  $P \cup V$ . We also say that the end of a durative action is applicable if the action is currently executing in the state, the action duration constraint is satisfied, and the at-end condition of the action is satisfied by  $P \cup V$ . Applying an instantaneous action, durative action start, or durative action end  $a$  to the state  $S$  produces a resultant state  $S(a)$ .

We formalize plans of actions as networks, similar to Temporal Plan Networks (Kim, Williams, and Abramson 2001):

**Definition 3.** A *plan*  $\Pi$  for a planning problem  $\langle P, V, A, I, G \rangle$  is the graph  $\langle N, C \rangle$  where each node  $n \in N$  represents the plan start, an instantaneous action, or the start or end of a durative action; and each edge  $c \in C$  represents a temporal relation:  $x < time(n_1) - time(n_2) < y$  for  $n_1, n_2 \in N$  and  $x, y \in \mathbb{R}$ . Each edge  $c$  is labelled as either *causal*, *interference*, or *action duration*. Action duration edges express the temporal constraints between the start and end of durative actions. Causal edges express temporal relationships inferred from the causal support between actions (causal links). Similarly, interference edges express the temporal relationships inferred from the interference between actions.

A plan is *totally-ordered* if there exists only one total ordering of nodes that can satisfy all of the temporal relations.

Similarly to the definitions presented in (Howey, Long, and Fox 2004) we say that a totally-ordered plan is *executable* if the plan can be simulated by applying each action in order, and all of the prescribed actions are applicable. The plan is *valid* if the final state satisfies the goal condition  $G$ .

## 3 Robust Plan Execution

PDDL2.1 planners output is represented as *time-triggered* plans (e.g. (Coles et al. 2010; Rankooh and Ghassem-Sani 2015)). A time-triggered plan is a set of tuples  $\langle t, a, d \rangle$ , where  $a$  is an action,  $t$  is the time at which the action should be executed and  $d$  is the prescribed duration.

As a running example, consider a scenario where a fleet of robots can move on a known map and are tasked to retrieve specific items produced from a pool of machines located at known positions. When a robot is at a machine location, it can turn the machine on. Once the machine is on, two robots are required to be at the machine location to produce an item. Clearly, the navigation actions of the two robots are independent until they synchronize to be at a specific machine together. A concrete instance is as follows. Two robots  $r0$ , initially in location  $wp1$ , and  $r1$ , initially in location  $wp0$ , must reach location  $m0$  where a machine is located, switch on the machine and produce an item that is finally delivered in location  $wp1$ . A valid time-triggered plan  $\Pi_{ex}^{tt}$  for the problem is reported below, indicating the starting time for each action to be executed (before the colon) and the expected duration (in brackets)<sup>2</sup>.

```

0.000: (goto r0 wp1 m0) [14.000]
0.000: (goto r1 wp0 m0) [ 9.000]
14.001: (switch_on r0 m0) [ 5.000]
19.002: (load_at_machine r1 r0 m0) [15.000]
34.002: (goto r1 m0 wp1) [14.000]
48.002: (ask_unload r1 wp1) [ 5.000]
53.003: (wait_unload r1 wp1) [15.000]

```

In order to allow for temporal flexibility, time-triggered plans must be converted into partially-ordered plans as per 3. In addition, we further relax the plan into an *adaptable* partially-ordered plan by removing the ordering relations that represent the causal support between actions. This allows for unexpected, but beneficial events in the environment to achieve the preconditions of actions. An adaptable partially-ordered plan can then be executed using the online procedure defined in 3.

Our approach is essentially composed of three phases, depicted in 1. The first phase is performed offline and is responsible for converting a totally-ordered input plan into an adaptable partially-ordered plan (this corresponds to the upper part of 1 and we provide full details in 3). The adaptable partially-ordered plan is passed to the second, online phase (3) that generates the set of valid totally-ordered plans described by the adaptable partially-ordered plan. This phase is indicated as “Total-Order Extractor” in 1. Finally, the produced set of valid totally-ordered plans is the input for the action dispatcher (3) that chooses the next action to execute by reasoning over the set of valid totally-ordered plans.

### Phase 1: Generating Adaptable Partially-Ordered Plans

Once a time-triggered plan ( $\Pi^{tt}$  in 1) is generated by an off-the-shelf planner, it needs to be converted into an ordinary partially-ordered plan  $\Pi$  by generating a node for each instantaneous action and each durative action start and end. The relations are generated as follows. For each node  $n_1$  that supports the condition of a node  $n_2$ , the temporal relation  $0 < time(n_2) - time(n_1) < \infty$  is generated. For each pair of nodes  $\langle n_1, n_2 \rangle$  representing the start and end of a durative action  $a$ , a relation representing the constraints in  $dur(a)$  is generated. Finally, for each pair of actions in the

plan  $a_1$  and  $a_2$  that interfere, an interference relation is generated. Two actions interfere if they have conflicting effects, or the effects of one action conflict with the conditions of the other (see (Fox and Long 2003) Definition 12). As durative actions are represented by two nodes (*at-start* and *at-end*), the relation is generated between the nodes that contain the interfering effects or conditions. In the case of an interfering *over-all* condition a relation is generated for both nodes.

For our running example, this gives the plan  $\Pi_{ex}$ , whose graph representation is reported in 2. It is easy to see that, while multiple ordering of the events are possible, the plan has a well-defined causal structure that ensures a sequence of pre-defined actions.

Before execution begins, the plan is relaxed into an adaptable partially-ordered plan. Given a plan  $\Pi$  the adaptable partially-ordered plan  $\Pi'$  is defined as  $\Pi' = \langle N', C' \rangle$ , where  $N' = N$  and  $C' \subseteq C$  contains only the edges of  $C$  labelled interference and action duration. By removing the causal support edges, the execution algorithm will be given more flexibility in action selection, possibly skipping actions whose effects have already been achieved by exogenous events. This is achieved using the algorithm in 3.

In our running example, we obtain the adaptable partial order plan  $\Pi'_{ex}$  reported in 3. This graph allows for many more executions orderings with respect to 2, but causal structure is no longer guaranteed for all of them and we need to restore this structure using runtime reasoning.

### Phase 2: Extracting the Set of Totally-Ordered Plans

Each adaptable partially-ordered plan generated from a valid totally-ordered plan encodes a superset of valid totally-ordered plans. Plan execution can be thought of as selecting one such totally-ordered plan to execute. In addition to this, our algorithm considers a notion of uncertainty in the environment, selecting the total-orderings that are most likely to succeed, given a known probability over propositions.

To perform this reasoning, we assign a probability to each proposition in  $P$  by means of a map  $\rho : P \rightarrow [0; 1]$ . These probabilities model the belief of the agent in the truthness value of the corresponding propositions. For instance, in a robot, propositions in the initial state may have their values assigned by a perception system based on its confidence on their values. Another example are probabilities in the effects of an action, their success probabilities.

Then, we consider a set of states of the form  $S = \langle P \cup V, \rho \rangle$ . Moreover, for each action  $a$  we use the map  $\psi_a : eff(a) \rightarrow [0; 1]$  to represent the probability of setting a proposition  $p \in eff(a)$  to true after executing the action. For a deterministic action,  $\psi_a(p) = 1$  if  $a$  adds  $p$ , and  $\psi_a(p) = 0$  if  $a$  deletes  $p$ . The operation of applying an action  $a$  in a state  $S$  (i.e.  $S' = S(a)$ ) is extended to update the set  $\rho$  by assigning

$$\rho(p) = \psi_a(p)$$

for all  $p \in eff(a)$ , while the  $P \cup V$  part of the state is updated as per 2. As for the transitions,  $S'$  can be reached from state  $S$  with the joint probability of  $a$ 's preconditions.

The algorithm requires as input an adaptable partially-ordered plan  $\Pi$  and the current state  $S_0$  (where we have

<sup>2</sup>This is the usual syntax PDDL planners use for temporal plans.

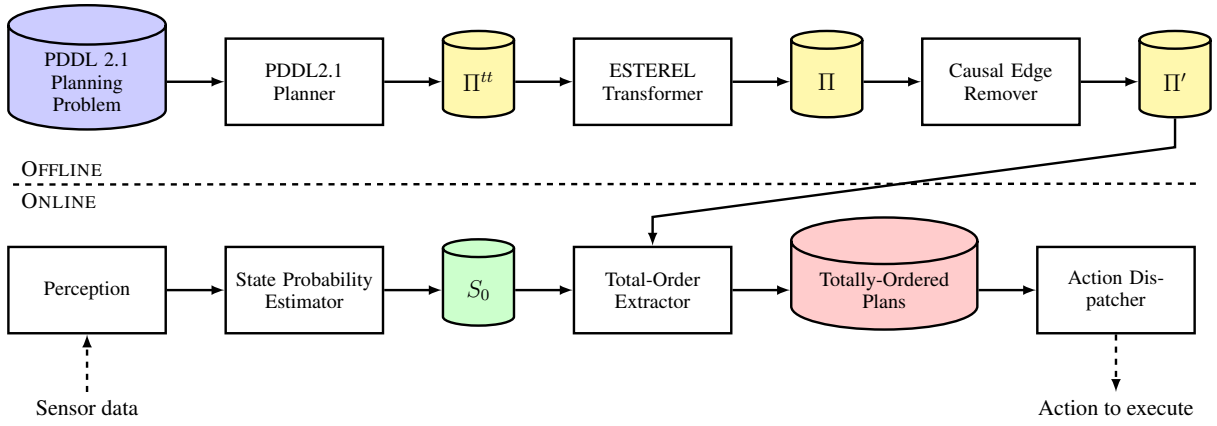


Figure 1: High-level overview of the proposed flow from planning to execution.

---

**Algorithm 1** Generate Totally-Ordered Plans

---

```

1: function GENERATEPLANS( $S_0, \Pi$ )
2:    $O \leftarrow \{n \mid \forall n \in \Pi\}$            ▷ initialize open list
3:    $C \leftarrow \{c \mid \forall c \in \Pi\}$        ▷ set of temporal constraints
4:    $F \leftarrow \emptyset$                  ▷ initialize set of ordered nodes
5:    $R \leftarrow \emptyset$                  ▷ return set of totally-ordered plans
6:   ORDERNODES( $O, C, F, S_0, R, 1$ )
7:   return  $R$ 

8: function ORDERNODES( $O, C, F, S, R, Q$ )
9:   if  $\neg$  CHECKTEMPORALCONSTRAINTS( $F, C$ ) then
10:    return                               ▷ the ordering is not executable
11:   if  $S \models G$  then
12:     $R \leftarrow R \cup \{\langle Q, F \rangle\}$    ▷ add new valid ordering to  $R$ 
13:    return
14:    $\Phi \leftarrow \text{VALIDNODES}(O, S)$ 
15:   if  $\Phi = \emptyset$  then
16:    return                               ▷ the ordering is not valid
17:   for each  $\langle q, a \rangle \in \Phi$  do
18:     $K \leftarrow \{b \in O \mid b \prec a\}$    ▷ all nodes  $b$  ordered before  $a$ 
19:     $F' \leftarrow F \cup \{a\}$          ▷ update current ordering
20:     $O' \leftarrow O \setminus (K \cup \{a\})$  ▷ remove skipped nodes from  $O$ 
21:     $S' \leftarrow S(a)$                ▷ apply  $a$  to  $S$ 
22:    ORDERNODES( $O', C, F', S', R, Q \times q$ ) ▷ recurse

```

---

probabilities assignments in the map  $\rho$ . The complete set of totally-ordered plans are then generated as described in 1. The algorithm can be run before execution, in which case the plan  $\Pi$  is the whole adaptable partially-ordered plan and the state  $S_0$  is the initial state before starting the execution. If, instead, the algorithm is run online,  $S_0$  is the current state with the probabilities derived from observations and the plan  $\Pi$  is the adaptable partially-ordered plan where nodes representing actions that have been executed in the past are removed. In either case, the algorithm will return the set of valid totally-ordered plans that can be executed from the state  $S_0$ .

The algorithm describes a search through the nodes contained in the plan. A representation of the state,  $S = \langle P \cup V, \rho \rangle$ , is used throughout the search to check action conditions and simulate their effects. At each step of the search,

the VALIDNODES procedure is called to return a set of tuples  $\langle q, a \rangle$ , where  $a$  is an applicable next node, and  $q$  is the joint probability of that action's preconditions (line 14). The search branches on each element of  $\Phi$ , applying the action  $a$  to  $S$  (line 21). The probability of reaching a node is stored in  $Q$  (initially 1), which is  $\prod_{a \in F} \rho(\text{pre}(a))$  where  $\rho(\text{pre}(a))$  is the probability of the action's preconditions in the state in which  $a$  was applied. For example, consider an action  $a_0$  with precondition  $p_0$  and positive effects  $p_1, \dots, p_5$ , with  $\psi_a(\text{eff}) = 1$  for all effects. Action  $a_1$  has precondition  $p_1 \wedge \dots \wedge p_5$ . The probability in the initial state is  $\rho(p_i) = 0$  for all  $i > 0$  and  $\rho(p_0) = 0.5$ .  $a_0$  is added to  $F$  (line 19) and the state is updated (line 21). The effects of  $a_0$  set  $p_i = 1$  in  $S'$  for all  $i = 1 \dots 5$ . The recursion is then called updating the probability of the next node to  $Q = 0.5$  (line 22). The joint probability of action  $a_1$  from the new search node is 1. Thus, the probability of being able to apply the sequence  $[a_0, a_1]$  from the initial state is 0.5.

The search has three *backtracking* conditions. First, if adding the new node violates a temporal constraint, including both interference and action duration relations, then the current total ordering is discarded (lines 9-10). Second, if the goal is achieved then the current total ordering is saved (line 12). Third, if the goal is not achieved and there are no more applicable nodes, then the current total ordering is discarded (lines 15-16). In the case of any of these three backtracking conditions, the search resets  $S$  and tries the next element of  $\Phi$  (line 17), so that every *executable* ordering of nodes is explored, and all *valid* orderings are saved. The returned result is the a of tuples:  $\langle Q, \Pi \rangle$  where  $\Pi$  is a totally-ordered plan and  $Q$  is that plan's probability of success.

The VALIDNODES procedure forms the expansion step of the search by returning a set of tuples  $\langle q, a \rangle$ , where  $a$  is an applicable next node, and  $q$  is the probability that the node is applicable. An action start node is applicable if the *at-start* and *over-all* conditions of  $\text{pre}(a)$  are true in the state, while an action end node is applicable if its *at-end* conditions are true. The probability  $q$  of the preconditions of node  $a$  (indicated as  $\text{pre}(a)$ ) in  $S$  is computed from  $\rho$  as the joint probability of the preconditions of  $a$ . For example, a node  $n$  with precondition  $\text{pre}(a) = p_1 \wedge \neg p_2$ , where  $p_1, p_2 \in P$  is

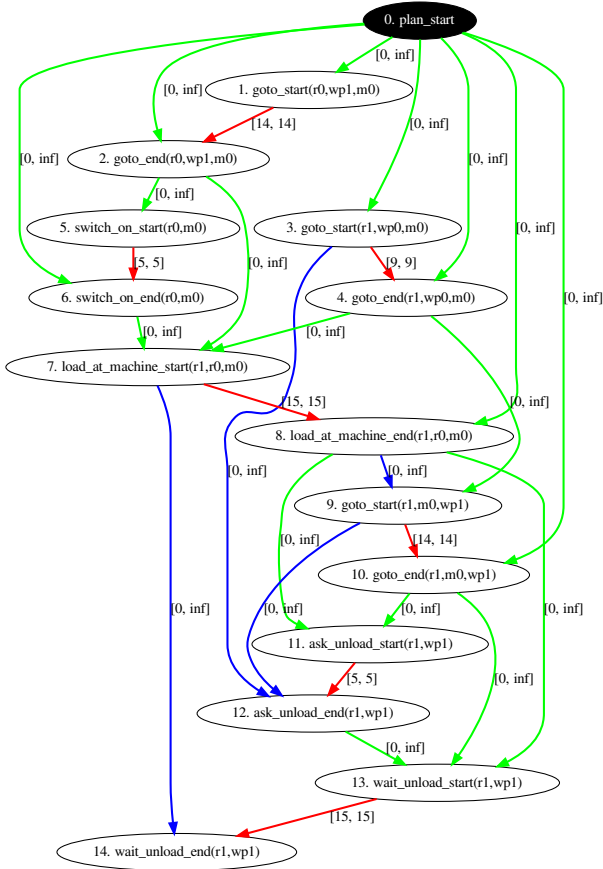


Figure 2: Graph of the de-ordered plan  $\Pi_{ex}$  for the running example. Nodes represent instantaneous actions or startings or endings of durative actions. Edges represent temporal relationships in the form  $[min, max]$ , and illustrate causal support (green) action duration (red) and interference (blue) constraints.

$q = \rho(p_1) * (1 - \rho(p_2))$ . If  $q$  is greater than some threshold (in our implementation this is 0), then the tuple  $\langle q, a \rangle$  is added to the return set of `VALIDNODES`, otherwise it is discarded.

Theoretically, the number of possible total orders induced by an adaptable partially ordered plan is factorial in the number of nodes: it suffices to consider a plan with no constraints where every permutation of the nodes is a valid total order. This is the dominant complexity cost, hence the algorithm runs in  $O(n!)$ . Nonetheless, this case is practically never encountered for meaningful, practical plans, because interference and duration constraints dynamically prune executions that are causally impossible given the current observations. In 4 we show that this approach exhibits very good empirical performance.

### Phase 3: Action Selection Policy

During execution, the executor is tasked with selecting the next node, whether this is the start of an action, the end of

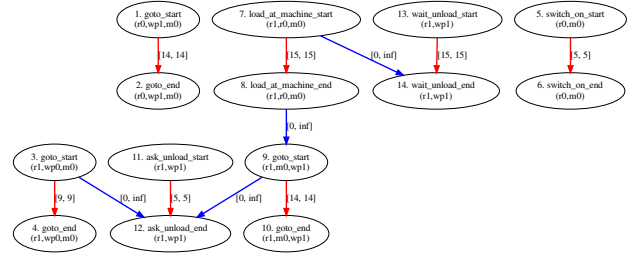


Figure 3: Graph of the adaptable partial-order plan  $\Pi'$  of the running example, obtained from  $\Pi$  by removing causal edges.

an action, or to await some other external timed event.

Given a set of totally-ordered plans with probabilities,  $\langle Q, F \rangle$ , our executor takes the ordering with maximum  $Q$  and select the node that is first in the order (whether that is to dispatch an action start, wait for an action end). For example, given the set:  $\{ \langle 0.5, [a, b, c] \rangle, \langle 0.3, [b, a, c] \rangle, \langle 0.3, [b, c] \rangle \}$  The executor would choose node  $a$  to be executed first. Note that in this example there are two orderings beginning with node  $b$  both with probability 0.3, but their probabilities are not necessarily independent. For example, it could be that the probability of being able to apply  $a$  and  $c$  is always 1, and that the probability of applying  $b$  is 0.5 after  $a$ , and 0.3 otherwise. For this reason, we cannot combine the probabilities of success of the orderings starting with  $b$  (that would be 0.6) and we decide to execute  $a$  instead.

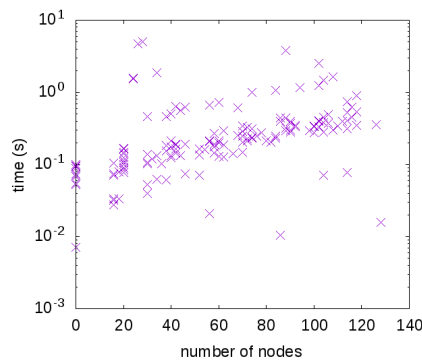
Given the plan  $\Pi'_{ex}$  for our running example situation, I can extract all the possible valid total orderings. Among these, the procedure is clearly able to reconstruct the original, totally-ordered plan  $\Pi'_{ex}$ , but other orderings are also possible, depending on the observed probabilities. Let us consider an example situation to clarify the possibilities opened by our approach. Suppose that the machine  $m0$  is found to be already switched on (with high probability) upon  $r0$  arrival. By running 1, the orderings having  $(switch\_on\ r0\ m0)$  as first action to execute will have very low probability, while the ones having  $(load\_at\_machine\ r1\ r0\ m0)$  will have high probability. For this reason, our action selection policy, exploiting the result from 1, will choose to execute  $(load\_at\_machine\ r1\ r0\ m0)$  and a subsequent successful continuation of the plan will effectively skip the  $(switch\_on\ r0\ m0)$  action. This is because the system will reach the goal without ever executing such an action.

## 4 Implementation and Evaluation

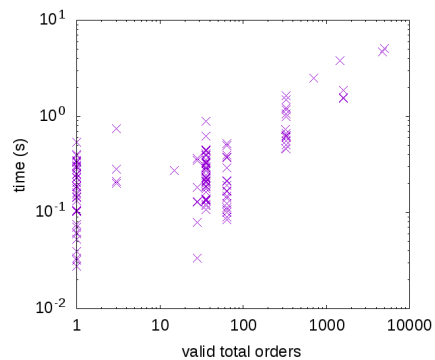
The approach has been implemented in ROS and integrated as an alternative execution algorithm in ROSPlan (Cashmore et al. 2015). The architecture of the resulting system is illustrated in 5. A PDDL2.1 domain and problem file are passed to the system at launch, thereafter a new planning problem is automatically produced during each planning episode.

	#orders	Coverage		Avg. Replans		Avg. #Actions	
		2	3	2	3	2	3
Deadline-Free	RO	98%	94%	0.9	0.9	12.5	17.2
	RP	98%	97%	1.6	2.0	13.7	19.3
With Deadlines	RO	91%		0.7		17.3	
	RP	97%		2.9		19.6	

(a) Results for deadline-free tasks (top) and for problems with deadlines (bottom), showing coverage, average number of replans across successful tasks, and average number of actions executed.



(b) Time taken to produce valid orders for plans of varying size.



(c) The number of valid orders generated, and time taken in generation.

Figure 4: Experimental results.

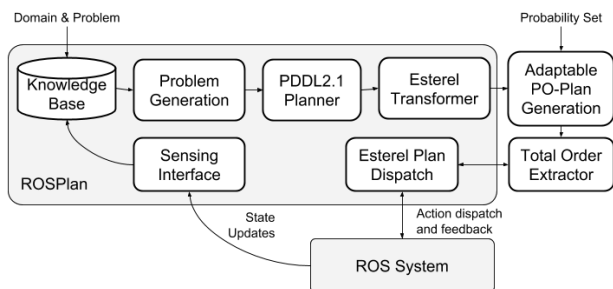


Figure 5: System Architecture for the evaluation. The proposed approach has been integrated into the ROSPlan system.

We use the planner POPF (Coles et al. 2010), as it is a PDDL2.1 planner that produces time-triggered plans and is already available with ROSPlan. The planner output is parsed into a partially-ordered plan within ROSPlan, and this output is subscribed to by our node, which publishes an adaptable partially-ordered plan following the procedure described in 3. This plan is then used to produce a set of valid totally-ordered plans, through an implementation of 1.

The plan dispatcher selects the first node of the plan with highest probability, and executes that node as described in 3. After each node is executed, the totally-ordered plan generation is run again (1).

## Experiment Description

We use the tasks in the Robot Delivery domain to investigate our approach. The domain is a much simplified version of the domain used in the Planning and Execution Competition for Logistics Robots in Simulation (Niemueller, Lakemeyer, and Ferrein 2015) comprising a fleet of small robots that can navigate in an euclidean graph. These robots are tasked to pick and deliver orders within a deadline. Collecting orders requires two robots present at a machine. We randomly generated a total of 30 initial states with 3 robots, 3 to 5 machines, 4 to 8 delivery locations, and the goal to deliver 2

or 3 orders. For 9 of these problems, the optimal plan duration was calculated, and a deadline was added to each order equal to 1.5 times the optimal plan duration. If this deadline is passed, the order cannot be delivered and the task is failed. Thus, the problem set contains 39 tasks overall. 9 tasks with deadlines, and 30 tasks without.

Each task was run in simulation using our re-ordering approach (RO) and a replanning (RP) approach. The RP approach attempts to directly execute the partially ordered plan  $\Pi$ , and replans when the execution fails. In contrast, RO follows the procedure described above, generating new total orders online to select the next action, and replanning only when no valid total ordering can be found by 1. Replanning takes place when (1) an action reports failure, (2) an action is to be executed, but its preconditions are not true, and (3) a temporal constraint is violated.

We use a non-physics simulation that includes a probability of action failure and non-deterministic action duration. In addition, the ground truth of the simulation is not static. For example, the proposition (`machine_on m0`) may be true in the initial state, but later change due to exogenous events. In this case, the plan execution may fail due to a mismatch between the planner’s model and the ground truth, leading to replanning condition (2). Each task and system was run in simulation 10 times to account for random events, action failure, and non-deterministic action duration.

For each run we recorded whether or not the task succeeded in delivering all of the orders. For the tasks that succeeded, we also recorded the number of times the plan execution failed and the system had to replan, and the number of actions that were actually executed in order to achieve the goals.

## Experimental Results

The results obtained are shown in Table 4a. The results demonstrate our hypothesis that in problems with and without deadlines the re-reordering approach will result in fewer replans, and that fewer actions will be executed overall. The coverage of RP is marginally higher in the problems with deadlines, where the overhead of RO can cause fail-



ures. However, the reduction in the number of replans is significant while impact on coverage is minimal. In domains with dead-ends (from which recovery through replanning is impossible) or in platforms with insufficient computational power to perform online planning, the results show that RO is a viable plan repair.

As discussed in 3, the worst case complexity of our algorithm can be factorial in the size of the plan. To provide some empirical insight in terms of performance, 160 problems were generated and solved, producing STN plans with up to 128 nodes. The time taken to generate the adaptable partially-ordered plan and to extract all valid total orders was recorded, as well as the number of total orders. 4b plots the time against the number of nodes, while 4c shows the time against the number of valid total orders produced. The algorithm takes less than 10 seconds in all cases, hence it is definitely suitable to be used online.

## 5 Conclusions

In this paper, we proposed a novel approach for improving the robustness of the execution of automatically-generated task plans with respect to unforeseen circumstances. The approach consists in relaxing the causal structure of the generated plan, allowing for run-time adaptation of the ordering of the actions and, possibly, for the skipping of some action in situations where their execution is no longer needed. The approach is proven effective on a simulated use-case, where the number of re-planning attempts was reduced.

There are several directions for future work. First, we plan to integrate this approach with (Cashmore et al. 2019) to natively handle discrepancies in continuous dimensions of the problem, such as time or resources. Second, we would like to try to use our approach with multiple plans: instead of generating the total orders from a single plan, we can use several, diverse plans to allow for more variability in the action selection policy. Finally, while in this work it is assumed that all fluents in the state are independent random variables (a simplification that was made), we leave a more rigorous approach for future work.

## Acknowledgments

This work was supported by the FCT projects [PTDC/EEI-SII/4698/2014], [UID/EEA/50009/2013], and [UIDB/50009/2020].

The work presented is part of the project TransFIT which is funded by the German Space Agency (DLR Agentur) with federal funds of the Federal Ministry for Economic Affairs and Energy under grant no. FKZ 50RA1701, FKZ 50RA1702 und FKZ 50RA1703.

## References

- Bäckström, C. 1998. Computational aspects of reordering plans. *Journal of Artificial Intelligence Research* 9:99–137.
- Cashmore, M.; Fox, M.; Long, D.; Magazzeni, D.; Ridder, B.; Carrera, A.; Palomerias, N.; Hurtos, N.; and Carreras, M. 2015. Rosplan: Planning in the robot operating system. In *ICAPS*.
- Cashmore, M.; Cimatti, A.; Magazzeni, D.; ; Micheli, A.; and Zehtabi, P. 2019. Robustness envelopes for temporal plans. In *AAAI*.
- Cesta, A.; Cortellessa, G.; Fratini, S.; Oddi, A.; and Rasconi, R. 2009. The APSI Framework: a Planning and Scheduling Software Development Environment. In *ICAPS (Application Showcase)*.
- Cimatti, A.; Micheli, A.; and Roveri, M. 2016. Dynamic controllability of disjunctive temporal networks: Validation and synthesis of executable strategies. In *AAAI*.
- Coles, A.; Coles, A.; Fox, M.; and Long, D. 2010. Forward-chaining partial-order planning. In *ICAPS*, 42–49.
- Dechter, R.; Meiri, I.; and Pearl, J. 1991. Temporal constraint networks. *Artificial intelligence* 49(1-3):61–95.
- Do, M., and Kambhampati, S. 2003. Improving temporal flexibility of position constrained metric temporal plans. In *ICAPS*, 42–51.
- Fox, M., and Long, D. 2003. PDDL2.1: An extension to pddl for expressing temporal planning domains. *Journal of Artificial Intelligence Research* 20:61–124.
- Frank, J., and Jónsson, A. 2003. Constraint-based Attribute and Interval Planning. *Constraints* 8(4):339–364.
- Frank, J., and Morris, P. 2007. Bounding the resource availability of activities with linear resource impact. In *ICAPS*, 136–143.
- Ghallab, M., and Laruelle, H. 1994. Representation and control in IxTeT, a temporal planner. In *AIPS*, 61–67.
- Howey, R.; Long, D.; and Fox, M. 2004. Val: Automatic plan validation, continuous effects and mixed initiative planning using pddl. In *ICTAI*, 294–301.
- Karpas, E.; Levine, S. J.; Yu, P.; and Williams, B. C. 2015. Robust execution of plans for human-robot teams. In *ICAPS*.
- Kim, P.; Williams, B. C.; and Abramson, M. 2001. Executing reactive, model-based programs through graph-based temporal planning. In *IJCAI*, 487–493.
- Levine, S., and Williams, B. 2018. Watching and acting together: Concurrent plan recognition and adaptation for human-robot teams. *Journal of Artificial Intelligence Research* 63.
- Niemueller, T.; Lakemeyer, G.; and Ferrein, A. 2015. The robocup logistics league as a benchmark for planning in robotics. *Planning and Robotics (PlanRob-15)* 63.
- Rankooh, M. F., and Ghassem-Sani, G. 2015. Itsat: An efficient sat-based temporal planner. *Journal of Artificial Intelligence Research* 53(1):541–632.
- Umbrico, A.; Cesta, A.; Mayer, M. C.; and Orlandini, A. 2018. Integrating resource management and timeline-based planning. In *ICAPS*, 264–272.