

Bachelorarbeit

# Malleable TOHTN Planning using CrowdHTN and Mallob

Niko Wilhelm

Abgabedatum: 19.10.2012

Betreuer: Prof. Dr. Peter Sanders  
M.Sc. Dominik Schreiber

Institut für Theoretische Informatik, Algorithmik  
Fakultät für Informatik  
Karlsruher Institut für Technologie

---

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen, als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Instituts für Technologie zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet habe.

Karlsruhe, den October 15, 2022

## **Zusammenfassung**

Hier die deutsche Zusammenfassung.

Ich bin Blindtext. Von Geburt an. Es hat lange gedauert, bis ich begriffen habe, was es bedeutet, ein blinder Text zu sein: Man macht keinen Sinn. Man wirkt hier und da aus dem Zusammenhang gerissen. Oft wird man gar nicht erst gelesen. Aber bin ich deshalb ein schlechter Text? Ich weiß, dass ich nie die Chance haben werde im Stern zu erscheinen. Aber bin ich darum weniger wichtig? Ich bin blind! Aber ich bin gerne Text. Und sollten Sie mich jetzt tatsächlich zu Ende lesen, dann habe ich etwas geschafft, was den meisten „normalen“ Texten nicht gelingt.

Ich bin Blindtext. Von Geburt an. Es hat lange gedauert, bis ich begriffen habe, was es bedeutet, ein blinder Text zu sein: Man macht keinen Sinn. Man wirkt hier und da aus dem Zusammenhang gerissen. Oft wird man gar nicht erst gelesen. Aber bin ich deshalb ein schlechter Text? Ich weiß, dass ich nie die Chance haben werde im Stern zu erscheinen. Aber bin ich darum weniger wichtig? Ich bin blind! Aber ich bin gerne Text.

## **Abstract**

And here an English translation of the German abstract.

I'm blind text. From birth. It took a long time until I realized what it means to be random text: You make no sense. You stand here and there out of context. Frequently, they do not even read. But I have a bad copy? I know that I will never have the chance of appearing in the. But I'm any less important? I'm blind! But I like to text. And you should see me now actually over, then I have accomplished something that is not possible in most "normal" copies.

I'm blind text. From birth. It took a long time until I realized what it means to be random text: You make no sense. You stand here and there out of context. Frequently, they do not even read. But I have a bad copy? I know that I will never have the chance of appearing in the. But I'm any less important? I'm blind! But I like to text.

## **Danksagungen**

Thanks to i10pc135 which suffered much to make the experimental evaluation possible.

# Contents

<b>1</b>	<b>Introduction</b>	<b>8</b>
1.1	Motivation . . . . .	8
1.2	Research Goal . . . . .	8
1.3	Thesis Overview . . . . .	8
<b>2</b>	<b>Preliminaries</b>	<b>8</b>
2.1	Planner Properties . . . . .	8
2.2	TOHTN Formalism . . . . .	9
2.2.1	Defining TOHTN Planning Problems . . . . .	9
2.2.2	Complexity of (TO)HTN planning . . . . .	10
2.2.3	Differences from other Kinds of Planning . . . . .	11
2.3	Techniques to solve TOHTN planning problems . . . . .	12
2.4	Translation-based . . . . .	12
2.4.1	Search-based . . . . .	12
2.4.2	Lifted and Ground HTN Planning . . . . .	13
2.5	Malleability . . . . .	15
2.6	The CrowdHTN Planner . . . . .	16
2.7	The Mallob Scheduler . . . . .	18
<b>3</b>	<b>TOHTN Metadata</b>	<b>19</b>
<b>4</b>	<b>Implementation</b>	<b>19</b>
4.1	Adapting CrowdHTN for Use in Mallob . . . . .	20
<b>5</b>	<b>Improvements to CrowdHTN</b>	<b>21</b>
5.1	Reducing Memory Consumption . . . . .	21
5.1.1	Efficiently Storing the Preceding Plan . . . . .	21
5.1.2	Reducing Copies of the World State . . . . .	21
5.1.3	Only Saving 'Potentially Interesting' Nodes . . . . .	21
5.2	Efficiently Hashing Nodes of the Search Graph . . . . .	21
5.3	Preceding Plan . . . . .	22
5.4	Lazy Instantiation of Child Nodes . . . . .	22
5.5	Using Domain Meta-Information . . . . .	23
5.6	Global and Memory Efficient Loop Detection . . . . .	23
5.6.1	Restarts Under Loop Detection . . . . .	24
5.6.2	Reaching Consensus on Search Version . . . . .	24
5.6.3	Completeness Under Loop Detection . . . . .	24
<b>6</b>	<b>Theoretical Improvements to the Crowd Planner</b>	<b>25</b>
6.1	Search Algorithms Used in CrowdHTN . . . . .	26
6.1.1	Random Depth-First Search . . . . .	26
6.1.2	Random Breadth-First Search . . . . .	26
6.1.3	Heuristic Search . . . . .	26
6.1.4	Completeness of different Search Algorithms . . . . .	28
6.2	Loop Detection . . . . .	31
6.2.1	Loop Detection in Other HTN Planners . . . . .	31
6.2.2	Assumptions in Loop Detection for CrowdHTN . . . . .	31
6.2.3	Perfect Loop Detection . . . . .	32

6.2.4	Statistical Loop Detection . . . . .	32
6.2.5	Distributed Loop Detection . . . . .	33
<b>7</b>	<b>Grounding and Pruning</b>	<b>34</b>
<b>8</b>	<b>Malleability in the Context of CrowdHTN</b>	<b>35</b>
<b>9</b>	<b>Experimental Evaluation</b>	<b>39</b>
<b>10</b>	<b>Future Work</b>	<b>39</b>
10.1	A Common (TO)HTN Interface . . . . .	40
10.2	Combine Pruning Approaches . . . . .	40
10.3	Investigate the Completeness of Planners . . . . .	40
10.4	Lifted Parallel Search . . . . .	40
10.5	Improvements Unlocked by a Shared Memory Implementation . . . . .	40

## List of Figures

1	A pathological case in our new HTN heuristic . . . . .	30
2	Example diagram of a lost message . . . . .	37

## List of Tables

1	Completeness of the different search algorithms in CrowdHTN . . . . .	29
---	---	----

## Algorithmenverzeichnis

1	Classical Progression Search for HTN as introduced in [13] . . . . .	13
2	The parallel CrowdHTN algorithm . . . . .	17
3	GBFS heuristic calculation . . . . .	28

# 1 Introduction

## 1.1 Motivation

## 1.2 Research Goal

- Provide a performant parallel TOHTN planner by improving upon the Crowd planner - Provide integration of TOHTN into the Mallob malleable load balancer - Compare performance of parallel to malleable TOHTN planning -

## 1.3 Thesis Overview

# 2 Preliminaries

## 2.1 Planner Properties

- soundness: all solutions we find are correct - completeness: we always find a plan if it exists
- correctness: kinda obvious - systematicity ([13]): for search we explore each search node at most once (Kambhampati, Knoblock, & Yang, 1995) -> not a problem for TOHTN? check p.18 in holler2020 again - optimality: shortest plan (few planners, Lilotane maybe?)



## 2.2 TOHTN Formalism

In this section we first define what HTN and TOHTN problems are from a formal perspective 2.2.1. Afterwards we take a short look at the algorithmic worst case complexity of HTN and TOHTN planning 2.2.2.

### 2.2.1 Defining TOHTN Planning Problems

Both HTN and TOHTN planning are based on the idea of decomposing a list of initial tasks down into smaller subtasks until those subtasks can be achieved by simple actions.

Multiple definitions for HTN planning exist. In this work we build on the definition introduced in [10].

**Definition 1.** A **predicate** consists of two parts. Firstly a predicate symbol  $p \in \mathcal{P}$  where  $\mathcal{P}$  is the finite set of predicate symbols. Secondly of a list of terms  $\tau_1, \dots, \tau_k$  where each term  $\tau_i$  is either a constant symbol  $c \in \mathcal{C}$ , with  $\mathcal{C}$  being the finite set of constant symbols, or a variable symbol  $v \in \mathcal{V}$ , where  $\mathcal{V}$  is the infinite set of variable symbols.

The set of all predicates is called  $\mathcal{Q}$ .

With the definition of a predicate in place, we can then define a grounding as well as our world state.

**Definition 2.** A **ground predicate** is a predicate where the terms contain no variable symbols or, in other words, a predicate that contains only constant symbols.

**Definition 3.** A **state**  $s \in 2^{\mathcal{Q}}$  is a set of ground predicates for which we make the closed-world-assumption. Under the closed-world-assumption, only positive predicates are explicitly represented in  $s$ . All predicates not in  $s$  are implicitly negative.

**Definition 4.** With  $T_p$  the set of primitive task symbols, a **primitive task**  $t_p$  is defined as a triple  $t_p(\tilde{t}_p(a_1, \dots, a_k), pre(t_p), eff(t_p))$ .  $\tilde{t}_p \in T_p$  is the task symbol,  $a_1, \dots, a_k \in \mathcal{C} \cup \mathcal{V}$  are the task arguments,  $pre(t_p) \in 2^{\mathcal{P}}$  the preconditions and  $eff(t_p) \in 2^{\mathcal{P}}$  the effects of the primitive task  $t_p$ . We further define the positive and negative preconditions of  $t_p$  as  $pre^+(t_p) := \{p \in pre(t_p) : p \text{ is positive}\}$  and  $pre^-(t_p) := \{p \in pre(t_p) : p \text{ is negative}\}$ . We define  $eff^+(t_p)$  and  $eff^-(t_p)$  analogously.

We call a fully ground primitive task an **action**.

As preconditions and effects may not be concerned with the whole world state the closed-world assumption does not apply to them. To any HTN instance we could create an equivalent one where each precondition and effect cares about the whole world state. This would be achieved by instantiating all the "don't care" terms in preconditions and effects with all possible combinations of predicates. Doing this would, however, come at the price of a huge blowup of our planning problem.

**Definition 5.** An action  $t_p$  is **applicable** in state  $s$  if  $pre^+(t_p) \subseteq s$  and  $pre^-(t_p) \cap s = \emptyset$ . The **application** of  $t_p$  in state  $s$  results in the new state  $s' = (s \setminus eff^-(t_p)) \cup eff^+(t_p)$ .

**Definition 6.** We define a **compound task** as  $t_c = \tilde{t}_c(a_1, \dots, a_k)$ , where  $\tilde{t}_c \in T_c$  is the task symbol from the finite set of compound task symbols  $T_c$  and  $a_1, \dots, a_k$  are the task arguments.

quote m  
formalis

Primitive and compound tasks together form task networks. In places where both can be used, we will refer to them simply as tasks  $t \in T$ .

**Definition 7.** Let  $T = T_p \cup T_c$  be a set of primitive and compound tasks. A task network is a tuple  $\tau = (T, \psi)$  consisting of tasks  $T$  and constraints  $\psi$  between those tasks.

**Definition 8.** Let  $M$  be a finite set of method symbols and  $T = T_p \cup T_c$  a set of primitive and compound tasks. A **method**  $m = (\tilde{m}(a_1, \dots, a_k), t_c, \text{pre}(m), \text{subtasks}(m), \text{constraints}(m))$  is a tuple consisting of the method symbol  $\tilde{m}$ , the method arguments  $a_1, \dots, a_k$ , the associated compound task  $t_c \in T_c$  the method refers to, a set of preconditions  $\text{pre}(m) \in 2^P$ , a set of tasks  $\text{subtasks}(m) = \{t_1, \dots, t_l\}, t_i \in T$  and a set of ordering constraints  $c_1, \dots, c_m$  defining relationships between the subtasks. Any arguments appearing in  $t_c, \text{pre}(m), \text{subtasks}(m)$  must also appear in  $a_1, \dots, a_k$ .

In TOHTN planning,  $\text{constraints}(m)$  is implicitly set s.t. the subtasks  $t_1, \dots, t_l$  are totally ordered.

We call a fully ground method a **reduction**.

Each method  $m$  has exactly one associated compound task  $t_c$ . However, multiple methods  $m_1, \dots, m_k$  may be associated with a single compound task  $t_c$ . Additionally, while any arguments of  $t_c$  must be present in  $m$ , the contrary is not true and  $m$  may have arguments not present in  $t_c$ , i.e.,  $m$  is not fully determined by  $t_c$ . As a result methods present choice points both in the choice of method itself as well as through the argument instantiation.

**Definition 9.** Let  $\tau = (T, \psi)$  be a task network,  $s$  a state,  $m = (\tilde{m}(a_1, \dots, a_k), t_c, \text{pre}(m), \text{subtasks}(m), \text{constraints}(m))$  be a method.  $m$  **resolves**  $\tau$  iff  $t_c \in T$ , the constraints in  $\psi$  allow for  $t_c$  to be resolved,  $\text{pre}^+(m) \in s$  and  $\text{pre}^-(m) \cap s = \emptyset$ .

Resolving a compound task  $t \in T$  results in a new task network  $\tau' = ((T \setminus t) \cup \{t : t \in \text{subtasks}(m)\}, \psi \cup \text{constraints}(m))$  and state  $s$ .

Applying a primitive task results in a new task network  $\tau' = (T \setminus t, \psi)$  in state  $s'$  where the effects of  $t$  have been applied to  $s$ .

**Definition 10.** An **HTN domain** is a tuple  $D = (V, C, P, T, M)$  consisting of finite sets variables  $V$ , constants  $C$ , predicates  $P$ , tasks  $T$  and methods  $M$ . An **HTN problem**  $\Pi = (D, s_0, \tau_0)$  consists of a domain  $D$ , an initial state  $s_0$  and an initial task network  $\tau_0$ .

If  $\text{subtasks}(m)$  has a total order for all  $m \in M$  and the tasks in  $\tau_0$  are totally ordered, we speak of a **TOHTN domain** and **TOHTN problem**.

It is possible to translate any HTN problem with initial task network  $\tau_0$  into an equivalent HTN problem with initial task network  $\tau'_0$  s.t.  $\tau'_0$  consists of only a single task.

It is possible to simplify the model s.t.  $\tau_0$  always consists of only a single task with no constraints. We do this by inserting a new initial task  $t_0$  and method  $m_0$  with no arguments s.t. resolving  $t_0$  via  $m_0$  results in  $\tau_0$ .

### 2.2.2 Complexity of (TO)HTN planning

The complexity of HTN and TOHTN planning has been studied in many papers. Here the problem PLANEXIST describes, whether for any given (TO)HTN instance a plan exists at all. It is not concerned with optimality.

Early on it was shown by [7] and [8] that the complexity of hierarchical planning formalisms

depends on things such as the existence and ordering of non-primitive tasks, whether a total order between tasks is imposed and whether variables are allowed. The combination of arbitrary non-primitive tasks, no total order imposed and allowing variables is what we talk about with HTN planning, the same combination but with a total order is what we mean with TOHTN planning. They showed that HTN planning is semi-decidable whereas TOHTN planning is decidable in D-EXPTIME while being EXPSPACE-hard.

We can see what D-EXPTIME means in practise when we consider the maximum size of a task network we need to consider. From [4] we know that if a solution to an HTN instance exists, it can be found within a maximum depth of  $|T_c| \cdot (2^{|\mathcal{Q}|})^2$ . Similarly, we see that a task network can have exponential width in it's depth. Consider for this an instance constructed such that each compound task has exactly two children and where primitive tasks are only occuring at the bottom most layer. Now each layer will be twice as wide as the one before, giving us exponential width.

Regarding the general relationship of hierarchical planning to complexity theory, [7] and [8] showed early on that HTN instances can be used to simulate context-free languages. This was extended by [12] who showed that TOHTN instances correspond exactly to context-free grammars.

In addition to planning itself, the problem of plan verification was studied. Here, [3] showed that plan verification is NP-complete, even under the assumption that not only the plan but also the decompositions leading to it are provided.

### 2.2.3 Differences from other Kinds of Planning

[17] creates a classification of planners into domain-specific, domain-independent and domain-configurable planners. They argue that HTN planning falls under domain-configurable with the decompositions providing advice to the planner to gain efficiency.

[13] argue that HTN-planning is not simply a domain-configurable version of classical planning and argue on the basis that [7, 8] showed that HTN-planning is strictly more powerful compared to classical planning which is PSPACE-complete.

While we agree with [13], one can still use HTN planning without using the full complexity of the model, using it instead to provide more efficient and guided versions of classical planning problems.

[2] mentions 4 ways to compute per bound on the task network size?

## 2.3 Techniques to solve TOHTN planning problems

In this section, we will give an overview over the different techniques with which HTN problems can be solved. The HTN planners produced by researchers can be classified along two main axes:

- the planning algorithm
- lifted vs grounded approaches

For the algorithms, the two main variations are translation-based algorithms that take an HTN instance and translate it into a problem in a simpler complexity class such as classical planning or SAT and search-based algorithms, that utilize techniques such as plan-space search and progression search. We will focus on progression search here, as it is the technique employed in our own planner, CrowdHTN.

After that we will have a short discussion on lifted versus grounded approaches which is largely independent of the search algorithm.

## 2.4 Translation-based

instead  
AT-  
l"?

One of the main techniques employed in HTN planning is to find an efficient encoding into a simpler problem. Two such problems are classical planning ([1]) and propositional logic (SAT). While translation to SAT was already proposed in 1998 ([15]), the first complete encodings without assumptions about the instance were publicized in 2018 ([4]). In recent years, SAT seems to be the most popular problem to translate an HTN instance into, utilized by planners such as totSAT ([4], [5]), Tree-REX ([20]) and Lilotane ([19]).

As we have seen in the previous section 2.2.2 on (TO)HTN complexity, (TO)HTN problems are in D-EXPTIME and undecidable respectively. Both classical planning and SAT are less powerful. As a result, HTN problems cannot be encoded and even for TOHTN problems we would suffer a blowup in the size of the instance. Instead, as noted in [20], SAT-based planners tend to explore the set of potential hierarchies layer by layer, increasing the encoding size as they go. As a result, those SAT-based planners tend to have a BFS-like characteristic to their search.

### 2.4.1 Search-based

The second main category of techniques to solve HTN planning problems are search-based algorithms, such as plan space search and progression search. Both of these are described in [13]. Plan space search, as the name says, searches the space of partial plans, aiming to fix flaws - take definitions from [13] unless noted otherwise - explain where we differ from HTN progression search (holler2020htn also does this)

Progression search on the other hand generates plans in a forward way. What is meant by this is that it always chooses an open task that is currently unconstrained, i.e. has no unresolved predecessors under the ordering constraints, and resolves it. This allows the planner to update the world state as it goes along, as the sequence of actions from the start to the current point is known at each step of the search. In case of TOHTN planning, the choice of the next unconstrained task becomes trivial, as there is always exactly one such task. Knowing the full world state gives progression search two main advantages over plan space search. First, it allows the planner to prune parts of the search space by immediately validating action and reduction preconditions against this world state. Second, it gives us maximum information to be used in heuristics that guide our search.

The progression search algorithm is given in pseudocode in algorithm 1. As mentioned, line

6 becomes trivial for TOHTN planning and the loop from lines 7 to 16 is no loop, as there is always exactly one unconstrained task. Additionally, the location of our goal test can be moved around, depending on need. Performing the goal test upon popping a node is useful if we want to find optimal plans and our fringe data structure - and thus popping order - have a notion of node cost. Performing the goal test upon node creation allows us to terminate earlier.

Notable search-based planners are SHOP ([16]), HyperTension ([14]), PANDA ([13]) and our own planner CrowdHTN which will be presented in a later section 2.6.

---

**Algorithm 1:** Classical Progression Search for HTN as introduced in [13]

---

```

1 fringe  $\leftarrow \{(s_0, tn_I, \epsilon)\}$ 
2 while fringe  $\neq \emptyset$  do
3   n  $\leftarrow$  fringe.pop()
4   if n.isgoal then
5     return n
6   U  $\leftarrow$  n.unconstrainedNodes
7   for t  $\in$  U do
8     if isPrimitive(t) then
9       if isApplicable(t) then
10        n'  $\leftarrow$  n.apply(t)
11        fringe.add(n')
12      else
13        for m  $\in$  t.methods do
14          n'  $\leftarrow$  n.decompose(t, m)
15          fringe.add(n')

```

---

### 2.4.2 Lifted and Ground HTN Planning

As mentioned in section 2.2.1, (TO)HTN instances are normally given in a lifted representation and can be ground, i.e. all variables are filled with all possible parameter combinations. Specifying the instance in a lifted fashion is done for ease of use, as it is a more compact representation and allows domains to be reused for different problems [6].

The efficient grounding and pruning of HTN instances is an active field of research ([18], [6]). While it is an easier problem than (TO)HTN planning itself, the ground instance may still be exponential in size compared to the lifted instance [6].

Planners may choose to operate on either lifted or ground instances. A discussion on the trade-offs involved is found in [19]. We will reiterate the main advantage of each approach here. Grounded representations have more information available for pruning. As an example, while some parameter combinations in reductions may lead to a contradiction later on and can thus be pruned, not all such combinations may be invalid and thus the corresponding lifted method may not be prunable. Lifted representations on the other hand may be a lot more compact in practice. For example, our (TO)HTN instance may want us to choose any of  $N$  trucks to transport a package from  $A$  to  $B$  where in practice the choice might not matter. Whereas a grounded representation will have to instantiate all operators concerning a truck  $N$  times, a lifted operation will avoid this and only choose a truck ad-hoc.

The choice of grounded vs lifted representation is independent of the choice of planning algorithm. We have examples of grounded translation-based planners (totSat [4], Tree-REX [20]),

lifted translation-based planners (Lilotane [19]) and also search-based planners that work on both lifted (HyperTensioN [14]) and ground representations (PANDA [13]).

Our own planner, CrowdHTN, walks a middle ground . It performs its search on a ground representation to allow detailed pruning according to the world state. However, it does not front-load the cost of a grounding procedure and instead grounds tasks and methods as needed.

## 2.5 Malleability

## 2.6 The CrowdHTN Planner

The CrowdHTN (Cooperative randomized work stealing for Distributed HTN) planner is implemented as a parallel state machine that uses work stealing for work balancing purposes. According to the definition of [9] we introduced in section 2.5, CrowdHTN is a moldable task, as the number of workers is arbitrary but fixed during execution.

Each local worker of CrowdHTN owns its own queue of search nodes, tuples of *open tasks* and *world state* and performs progression search on these nodes as explained in section 2.4.1. The basic CrowdHTN parallel search algorithm is shown in algorithm 2.

In the initial state, only the root worker has any search nodes. All other workers start empty. To perform load balancing, randomized work stealing is used. The work package exchange is implemented as a three step protocol

- (i) work request
- (ii) response
- (iii) ack (if response was positive)

Upon sending a positive work response, a worker increments its local tracker of outgoing work packages. When receiving an ack, the local tracker of outgoing work packages is decremented. This ensures that there is always at least one node that acknowledges the existence of each search node. This is used to enable CrowdHTN to determine a global UNPLAN. To do this each worker reports whether it has any work left. A worker reports true if it has a non-empty fringe or at least one outgoing work package.

This capability is especially helpful for small instances where it is plausible to explore the whole search space. As we saw in the earlier section on complexity (2.2.2), TOHTN planning is in D-EXPTIME making it infeasible to explore the whole search space for big instances.

The specifics of which node to explore for a work step and which node to split off to send a positive work response depend on the fringe implementation. In general, we perform work at the back end of the fringe, in the case of depth-first-search these are the last nodes to have been inserted. To split off work, we take from the front end of the fringe. This is done as a heuristic to split off a node which is still far from a plan and thus forms a relatively larger work package, leading to less communication. In case of depth-first search, this reduction of communication volume is doubly true. Nodes which are close to the beginning of our search path generally have fewer open tasks, reducing the size of the message when sending it off.



---

**Algorithm 2:** The parallel CrowdHTN algorithm

---

```
1 while true do
2   work_step()
3   if fringe.empty and not has_active_work_request then
4     r ← random worker id
5     send work request(r)
6     has_active_work_request ← true
7   for (message, source) ∈ incoming messages do
8     if message is work request then
9       if fringe.has_work() then
10        send positive work response(fringe.get_work(), source)
11        outgoing work messages += 1
12      else
13        send negative work response(source)
14    if message is work response then
15      if response is positive then
16        fringe.add(work response)
17        send work ack(source)
18      has_active_work_request ← false
19    if message is work ack then
20      outgoing work messages -= 1
```

---

## 2.7 The Mallob Scheduler

all from [21] - Mallob - **M**alleable **L**oad **B**alancer or **M**ulti-tasking **A**gile **L**ogic **B**lackbox - Mallob contains a parallel SAT solver - is able to solve multiple instances in parallel, adjusting resources per job on a dynamic basis - is a decentralized malleable job scheduler - there is no shared RAM, communication via message passing - jobs  $j$  can arrive at arbitrary times - for SAT: a logic formula in Conjunctive Normal Form (CNF) - each job  $j$  has a constant priority  $pi_j \in (0, 1)$  - each job  $j$  has a variable (over time) resource demand  $d_j \in \mathcal{N}$  - simple case:  $d_j$  is the number of available PEs -  $d_j$  allows a job how many resources it is able to utilize at most - Mallob expects the number of active jobs to be lower than the number of workers, i.e. that  $d_j > 1$  - each PE works on only one job at a time - the first PE assigned to a job  $j$  is its root and is fixed over the lifetime of the job - the volume of a job is proportional to  $d_j p_j / \sum_{j'} d_{j'} p_{j'}$  - PEs of a job are arranged as a binary tree which is complete except for potentially the lowest layer (nodes from the right of the last layer may be missing) - new nodes in a job are preferably taken from the list of suspended nodes of this job

### 3 TOHTN Metadata

- researchers in TOHTN planning have collected a number of test instances for TOHTN planning
- provide some analysis of those instances in the context of modelling TOHTN planning as graph search
- provide a foundation to discuss the effects of changes and improvements in the crowd planning framework

### 4 Implementation

## 4.1 Adapting CrowdHTN for Use in Mallob

- CrowdHTN is no longer a standalone process - one way would be to launch CrowdHTN as a new process from within Mallob - instead re-architect CrowdHTN to prepare it as a library for external use - the abstractions of CrowdWorker and CooperativeCrowdWorker - the TohtnMultiJob (representing a crowd worker) contains a thread which executes CrowdHTN internally - CrowdHTN no longer worries about how messages are sent and received (completely agnostic to the mechanism) - on the Mallob side, messages are received and written to a buffer protected by mutex - the crowd thread takes all available messages and forwards them to CrowdHTN - similarly, Crowd generates messages, hands them to Mallob and stops caring
- communication between Mallob and CrowdHTN happens via atomic boolean flags (except for the mutex'd message buffers) - this ensures that all Mallob Job functions respond fast - even initialization of CrowdHTN (parsing, starting the work thread) happens within it's own thread to return fast

## 5 Improvements to CrowdHTN

### 5.1 Reducing Memory Consumption

#### 5.1.1 Efficiently Storing the Preceding Plan

- no longer use a vector of USignature - instead use an `ImmutableStack<USignature>` - this way we can go from linear overhead to constant memory per search node! - actually, with A star or loop detection via hashset it gets like, really complicated - argh - still, lots of sharing! - lets protocol the thing (memory logging overall)

#### 5.1.2 Reducing Copies of the World State

- if the next open task is an action, there is only a single possible way to resolve it - in other words, a search node whose next open task is an action always has zero or one children - this is not a very interesting node, either the preconditions hold or they do not - such nodes are no longer represented explicitly - instead, apply all actions from the sequence of open task up until the first abstract task is encountered - then also resolve that abstract task

#### 5.1.3 Only Saving 'Potentially Interesting' Nodes

### 5.2 Efficiently Hashing Nodes of the Search Graph

Overview:

- the hash of a node consists of two parts
- 1: the hash of the open tasks
- 2: the hash of the world state
- We do not care about the hash of the preceding plan. Nodes with open tasks and world state equivalent have equivalent plans leading to a goal (somewhat similar to the Nerode relation) and we do not care about optimality. How he reach this point with equivalent remaining options is thus not of interest to us
- The hash of the world state can be large, but the number of elements is ultimately bound for any particular instance
- The length of the preceding plan is unbounded

Open tasks:

- Care needs to be taken to use an order independent hash function for the world state, as the underlying representation as a set does not guarantee us any particular iteration order, especially between nodes (might depend on the order of things inserted into the set!)
- Alternatively we could have chosen some other ordered representation, e.g. maintain the world state as a sorted list of predicates with a defined order. This would imply extra work we are unwilling to do.
- For open tasks, we save not only the task itself but additionally the hash of all open tasks from first to current one
- The order in which we hash open tasks is from oldest to newest one
- On applying a Reduction, we push the new open tasks onto the open task stack and compute each of their hashes combined with their predecessors. Each of these hashing operations is completed in  $O(1)$

Add gra  
showing  
many fe  
nodes a  
represen

Time is  
actually  
constan  
world st  
-> what  
the exp  
time?

- Effectively, this means that each task is hashed exactly once
- As we already have to push each task onto the open tasks, inserting an additional  $O(1)$  operation on pushing does not change the asymptotic runtime

World state:

- In section XXX we discussed how each copy of the world state can be shared by many search nodes to reduce the memory footprint
- Instead of hashing the world state each time on demand, we can store a shared tuple of world state and its hash
- This way, we only hash the world state once, reusing the computation
- Is this actually a time saving?
- Future work: the hash of the world state is order independent (sum of squares of individual hashes), to not worry about forcing any fixed iteration order. Utilize this and wrapping maths to hash only the differential of

### 5.3 Preceding Plan

In the initial implementation each node stored the full preceding plan as a sequence of all reductions that were applied so far. This leads to roughly quadratic overhead (sum over  $1..n$ , only roughly as not each step increases the length. Wait, is it roughly, then? Probably, as the fraction of steps that increment the preceding plan should be kinda constant) This duplication was not needed. The newer implementation instead stores an optional<reduction> in each node. I.e., the preceding reduction is stored if one exists, nothing if there isn't one. When the preceding plan is needed, either for communication or to extract a plan, the current search path is iterated and all reductions are accumulated.

### 5.4 Lazy Instantiation of Child Nodes

lazy instantiation works on the basis of finding all free variables of a method and creating Reductions based on all possible combinations

Initial implementation: instantiate all possible reductions, filter out any with not fulfilled pre-conditions, then shuffle them

Problems: this spends both time and memory instantiating reductions that might never be needed for the rest of the search We effectively save not only the current path, but also follow all possible branches to a distance of 1

Solution: lazily create reductions as needed How to do this (first way): adapt the argiterator from Lilotane into the CrowdHTN code base. Adapt it to only substitute the arguments that are not already determined by the corresponding task (arguments)

To achieve randomization: each domain is iterated to create the substitutions Each time we build such an iterator, we randomize the order within the domains for this specific iterator This will lead to different orders

Further ideas: each time domains  $1..k$  have been fully iterated, increment  $k+1$  by 1, then shuffle the order of domains  $1..k$

For  $n$  total domains, each time domains  $1..n-1$  have been iterated, remove the current value from domain  $n$ , then shuffle all domain orders

Compare the runtimes of eager and lazy instantiation, check at which point it is worth it to incur the (potential) additional overhead of lazy instantiation Compare on multiple domains?

Check different metrics for comparison (size of domains, number of parameters, number of potential children (product of sizes))

Potential problems: We need quite a bit of state (domains, current index into each domain) to perform lazy instantiation The order is not truly random. We iterate some domains faster/more often than others. What if the important change is in a domain which is iterated slowly? More random order makes this easier

A potential solution: space filling curves Advantages: little state (can just be incremented), iterates all dimensions equally Disadvantage: fixed order. With shuffling within each domain might be random again

Space filling curves come with the restriction of being strictly continuous We do not need this property. All we are interested in is an easy to compute fixed order in which the whole space is iterated where each permutation is hit exactly once We want only self-avoiding curves, to not hit any instantiation twice (could loop detection just fix that? But it would be a bad fix)

## 5.5 Using Domain Meta-Information

inspiration taken from HyperTension

taking preconditions of tasks and lifting them up into methods Only do this if the precondition is guaranteed to also apply to the method (cannot be achieved before the respective task?) This allows us to stop exploring paths earlier

## 5.6 Global and Memory Efficient Loop Detection

So far: each worker has a hash set of each node ever encountered (note: we define node equality through the sequence of open tasks and the set of predicates that is the world state. Depth in the graph and preceding plan are ignored) Advantages: allows for perfect local loop detection, as we can always fall back to equality checks in case of hash collisions. Problems: This approach leads to a massive memory overhead, as the world state is duplicated countless times. This also leads to increased run times just to manage the growing hash set. This is also not suited for global loop detection. Global loop detection would need some mechanism to communicate seen nodes. Nodes are too large to communicate all of them, though. If we only communicate some nodes instead of all we run the risk of a node "going past" that barrier of known nodes and still exploring the swathes of known nodes beyond the "barrier" TODO: what about open tasks (memory consumption)? Or was that an ImmutableStack with little overhead?

Idea: drop the precondition of perfect loop detection with no false positives. For loop detection use a bloom filter with a set of hash functions for search nodes. This comes with a fixed, configurable memory overhead per worker. As a result, overall memory consumption should - drop - be more predictable To communicate nodes we can just communicate the bit array that makes up our filter and combine them via bitwise OR. This ensures that communication volume is also fixed. It is independent of both size of nodes (i.e. length of open tasks sequence) and number of explored nodes (since last round of communication). In addition to communicating the bloom filter we communicate the number of newly added nodes. Then each worker knows an upper bound for the total number of nodes that are part of it's bloom filter (might be an overestimation if two workers insert identical nodes as the node would be counted twice). As bloom filter performance/ false positives degrade with the number of elements inside it, we can restart our work depending on the number of elements in the bloom filter.

- loop detection information is shared at most every second (or however long it takes for the communication to complete) - as a result, when the root node increases it's version, a broadcast

is started that communicates the new version to everyone else - the version in normal messages is still kept. In case a message arrives before the version broadcast, the version is increased even earlier. New nodes likely get their version with the first work package.

### 5.6.1 Restarts Under Loop Detection

- bloom filters are of fixed size. If we propose a maximum false-positive rate, they will fill up at some point - this necessitates the need for restarts - bloom filter architecture: for local loop detection, use an expanding bloom filter (citation!) - this allows us to not incur any restarts due to local state - for global loop detection, we use a fixed-size bloom filter - if a node is encountered twice for local loop detection, add it to the global loop detector, to communicate
- we know that in mallob, the tree of workers is always a 'full' tree, except for the lowest level - i.e. we may be missing some nodes to the right end of the lowest level - especially, the root node always survives - as a result, the root node always receives all the messages about shared loop detection data - i.e., if the global loop detection bloom filter on any node is full, it follows that the global loop detection bloom filter on the root node is also full - the opposite does not hold (i.e. a new node enters just as the root node bloom filter fills up) - as a result, we can simply delegate the question of whether to restart to the root node and have it done in a centralized way, simplifying our communication patterns

### 5.6.2 Reaching Consensus on Search Version

- each search version corresponds to a restart with subsequent doubling in size of the global loop detector - keeping versions in sync is of importance - we do not want to loose any work - we do not want to insert search nodes from a wrong version into our bloom filter (especially old node into new filter, other way around is discarded either way) - each message between workers is extended by their internal version - if the version of a received message is higher than the internal version, the internal version is increased and then acted accordingly - new workers get updated to the current version the first time they ask for a work package - work requests and responses do not suffice as a version updating mechanism, as work packages may be arbitrarily large and messages arbitrarily rare -

### 5.6.3 Completeness Under Loop Detection

- any single iteration may be incomplete, e.g. if initial node and goal node hash to the same values - with restarts, we have uniform hashes, compute the chance of colliding each time - this necessitates changing seeds with each restart - soon, we will not collide - yay, plan can be found!
- we do loose termination - so far, termination could be known if the search tree gets too deep (cite limits known from other TOHTN paper) - however, this is infeasible, as RAM is actually finite (refer to known branching factors!) - i.e., we do not know the hash of a goal node, as a result we cannot exclude the possibility of it simply being cut off each time so far - we also do not know the hash of all nodes just before a goal node, etc (the path could always only be a single node wide) - as a result, we cannot terminate at any point, as we cannot exclude the plan hiding in some always-cut-off part of the search space - losing termination is not a big deal anyways, I guess?



## **6 Theoretical Improvements to the Crowd Planner**

## 6.1 Search Algorithms Used in CrowdHTN

As part of the re-engineering of CrowdHTN, we changed the implementation of the search algorithms to be based on a fringe. As mentioned in [13] we can simply switch out the underlying fringe data structure to emulate different search algorithms without making any changes to our core planner. Enabled by this change, we have implemented four search algorithms and will discuss them in the following section:

- Random depth-first search
- Random breadth-first search
- Greedy best-first search
- A-star like search

### 6.1.1 Random Depth-First Search

Randomized depth-first search is the only search algorithm that was already present in the previous implementation of CrowdHTN. It is implemented using a Last-In-First-Out queue as our fringe. Whenever new search nodes are inserted into the fringe we randomize their order to keep the random properties of the original CrowdHTN. This is done to avoid any pathological cases where e.g. the first method applicable to a task leads to an endless loop. We do not expect any differences in behavior or performance compared to the previous implementation.

### 6.1.2 Random Breadth-First Search

Randomized breadth-first search is the first new search algorithm that we implemented. It is done by using a First-In-First-Out queue, allowing us to explore all the potential task hierarchies layer by layer. The insertion order of new nodes is randomized as in the depth-first-search. We do this as the number of search nodes may be exponential in the depth, e.g. if each task can be resolved by at least two reductions. As such, the order in which we explore the nodes of any layer may still have a big impact and in this way we avoid pathological behavior.

In general, we expect a higher memory footprint compared to depth-first search and assume that the planner will struggle with domains where either plans are only found in deep layers or where the branching factor is very high as both will lead to a blowup in the size of our queue. At the same time, we expect the performance of breadth-first search to be a lot more consistent than for depth-first search, as the layer at which we find a plan stays fixed for any single instance.

Overall, we do not expect high performance of our breadth-first search. It may however prove useful on some domains and help us understand and validate assumptions about the behavior of TOHTN problems.

### 6.1.3 Heuristic Search

Both depth-first and breadth-first search are unguided and do not depend the order of search node exploration on any information contained in those nodes. Other planners, such as PANDA, use heuristics to guide their search. We will describe general information about heuristics and their use in PANDA according to [13] and then describe how we try and adapt the use of heuristics for malleable TOHTN planning.

**Heuristics in HTN in general and PANDA specifically** The general idea of heuristics is to avoid exploring all of our search space. They achieve this by guiding the search to the most promising search nodes first. In TOHTN planning, our choices during search are restricted by both the hierarchy of tasks as well as the world state. As a result, the best heuristics should make use of both pieces of information for the best results.

One avenue to deriving heuristics for HTN planning would be to adapt classical planning heuristics. This proves difficult, however, as these heuristics do not know about the hierarchy and may assume a state-based goal which HTN planning often does not have. To avoid such issues, the PANDA planner goes the other way around. PANDA computes a classical planning problem which is a relaxation of the HTN problem at hand. Then a solution to this relaxed model is approximated with the help of classical planning heuristics and the result is used to guide the initial HTN planning procedure. The computation of the classical model is possible in polynomial time and only done fully in the beginning, afterwards the model is only updated for the current state of planning. As a result, the heuristic takes into account both hierarchy and world state while remaining relatively efficient.

**Problems with the PANDA heuristic for malleable HTN planning** While PANDA has managed to make great use of heuristics in HTN planning, we cannot simply adopt the same heuristics for malleable CrowdHTN. The reason for this lies in the assumption of PANDA that a ground problem instance is already available. Grounding is an expensive operation, though, as discussed in [6]. A full grounding may be exponential in size compared to the input and runtimes of grounding and can be accordingly high.

While such a grounding is already available in PANDA, CrowdHTN does not perform explicit grounding before planning. Doing so may prove interesting for a parallel planner where the grounding would only be performed once. In a malleable environment without shared memory we can expect this grounding to take place every time a new worker is added, adding a high startup cost. Worse, a short-lived worker may be interrupted while still grounding, never getting around to any planning work. This would interfere with the efficient usage of available resources. For this reason we have decided against using the PANDA heuristics in CrowdHTN and instead tried to design a simpler heuristic to be used in malleable TOHTN planning.

**A Heuristic for malleable HTN Planning** To counter the startup cost of the PANDA heuristic, we have devised a simpler heuristic which is cheap to precompute and can easily be used in malleable planning. The goal is to have little startup overhead and to retain the efficient evaluation at each planning step. While this necessarily limits any precomputation to the lifted instance and the precision of our heuristic comes with problems as discussed in the previous paragraphs, we hope to still find some performance gains on at least some problem instances. As heuristic value, we simply use a lower bound on the number of reductions we still need to perform to fully resolve our list of open tasks. When computing this lower bound we ignore preconditions and effects, searching the shortest possible way through the hierarchy. We precompute this value for each task as described in algorithm 3. The remaining depth for each action is initialized to zero. The remaining depth for each abstract task is initially unknown. In each step we loop over all tasks  $t$  whose remaining depth is unknown. For each method  $m$  of task  $t$  we check the depth of all subtasks. If all depths are known, the method depth is set to the sum of all subtask depths plus one. For each task we choose the minimum value over all available reductions.

If in any iteration we do not get a new depth for at least one task, we stop the computation. Any task which does not have an assigned depth at this point is not resolvable at all and can be ignored.

Let  $t_a$  be the set number of abstract tasks,  $t_p$  be the number of primitive tasks and  $m$  the number of methods. A single iteration takes time in  $\mathcal{O}(t_a + m)$ . The number of overall iterations required is bound by  $t_a$ , as in any iteration at least one abstract task will be assigned it's final score, as we do not have any negative cycles. This gives us an overall runtime bound of  $\mathcal{O}(t_p + t_a \cdot (t_a + m))$ . As all those numbers are relating to the lifted instance we expect the overall runtime to be small in practise. To evaluate our heuristic while planning, we now

---

**Algorithm 3:** GBFS heuristic calculation

---

```

1 task depths  $\leftarrow \{(t_c, 0) | t_c \in \text{concrete tasks}\}$ 
2 while task depths changed do
3   for  $t_c \in \text{compound tasks}$  do
4     reduction depths =  $\emptyset$ 
5     for  $r \in \text{reductions for } t_c$  do
6       if depths of all subtasks are known then
7         reduction depths  $\cup = 1 + \sum\{d | d \text{ is depth of a subtask of } r\}$ 
8     if reduction depths  $\neq \emptyset$  then
9       task depths  $\cup = \{(t_c, \min(\text{reduction depths}))\}$ 

```

---

need to look not only at one but at the whole sequence of open tasks and calculate the sum of our heuristic over those tasks. While the naive solution gives us linear runtime in the number of open tasks, we can stretch the computation over task instantiation and reuse parts of it to perform heuristic evaluation in  $\mathcal{O}(1)$  at runtime. [?] Additionally, while we only use this heuristic to guide TOHTN planning, it ignores any orderings between open tasks. As such, it can be applied to HTN planning as well.

**Using the new heuristic in TOHTN planning** With the new heuristic presented in the previous paragraph, we implemented two new search algorithms for CrowdHTN, those being a guided depth-first search and an A-star like search.

The implementation of depth-first search used so far performs the search in a uniformly random order. Without knowing anything about the domain, this is a reasonable choice. While it is far from optimal search, it also avoids any pathological cases that may arise from a fixed order of exploration. As an alternative to this random order, we used the heuristic to guide our depth-first search. We have to note that this is still not necessarily deterministic. Randomness comes into play both when two reductions have the same heuristic score - which happens for different instantiations of the same method - and when performing work stealing in a parallel setting.

In addition to the guided depth-first search, we also implemented an A-star like search where a node's value is the sum of the heuristic value and the number of applied reductions. We differ from A-star in that we terminate the search as soon as a plan has been found instead of continuing on the search for an optimal plan. The hope is that the heuristics guide us towards a plan while giving weight to the number of applied reductions forces us to turn back and explore other parts of the search space without getting lost in pathological cases.

#### 6.1.4 Completeness of different Search Algorithms

In the previous paragraphs we discussed a number of different search algorithms that we implemented for CrowdHTN and how we expect them to effect the performance characteristics

Table 1: Completeness of the different search algorithms in CrowdHTN

Algorithm	Completeness Property
Random DFS	Not complete. Positive probability to find a plan if it exists
Random BFS	Complete
Heuristic DFS	Incomplete
A-star like	Complete

of our planner. The search algorithm has a more fundamental impact than that, however, and may effect the property of completeness. In this section we want to give a short overview over the completeness of each of the algorithms. A short overview can be seen in table 1. Note that this discussion is only about the algorithms without any modifications. In section 6.2 we discuss loop detection mechanisms to improve those algorithms which are so far not listed as complete.

Depth-first search is not complete as it may enter an endless loop and, even if it still explores side-tracks from this loop, will never be able to backtrack to before the start of this loop, cutting of parts of the search space. There is however always a chance to find a plan if it exists. I.e., there is a non-zero chance that the random choices all happen to be done correctly and the goal is found even if other paths may lead into endless loops.

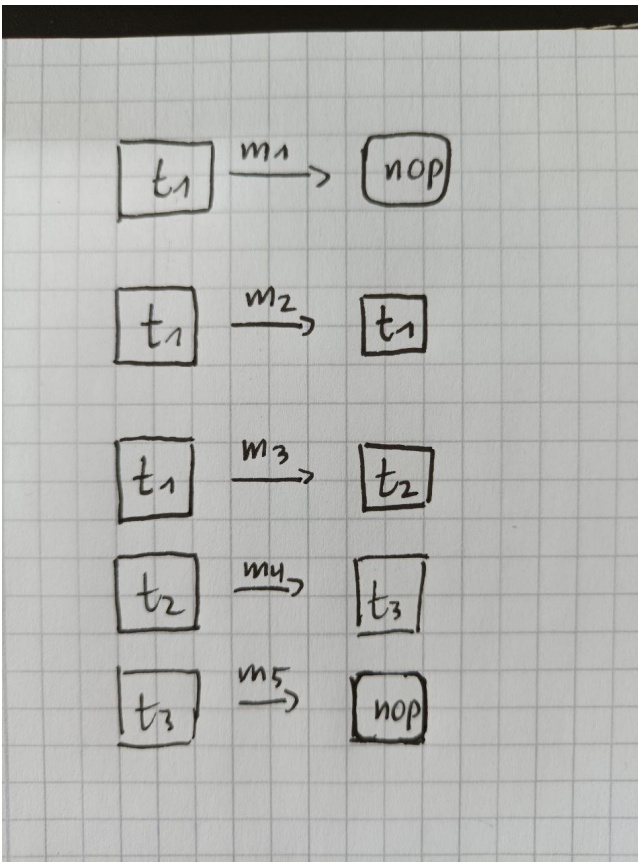
Breadth-first search on the other hand is trivially complete. While the exploration order within each layer is random we can provide an upper bound on the number of steps required to explore a given search node  $n$  on layer  $i$ . One such bound is the sum of all layer sizes from 0 up to and including  $i$ .

Next in our list is heuristic depth-first search. Similar to random depth-first search it may run into an endless loop. Compared to depth-first search, however, heuristic depth first search may do so in a deterministic fashion if the domain triggers a pathological case in the heuristic. One example domain which would trigger such a case in our heuristic is visualized in figure 1. Here rectangles stand for tasks, rounded rectangles for actions and arrows to reductions. Assume we want to resolve task  $t_1$  where the preconditions of  $m_1$  do not hold while those of  $m_2, m_3, m_4, m_5$  are fulfilled. Our heuristic will want us to resolve  $t_1$  via  $m_1$  which does not work. Then the next shortest path to a full resolution is through  $m_2, m_1$ .  $m_2$  will be applied successfully to recurse,  $m_1$  will fail again and the cycle continues like this forever. The path through  $m_3, m_4, m_5$  has length 3 and will thus never be taken.

Lastly, we implemented our heuristic search. While it does reuse the same heuristic, this algorithm achieves completeness by also valuing the number of applied methods so far. Let  $n$  be a node with heuristic value  $h$  and  $r$  previously applied reductions to reach  $n$ . Then  $n$  is guaranteed to be explored once all nodes with at most  $h + r$  previously applied reductions have been explored.

All of this discussion so far has assumed sequential planners. We can easily extend the incompleteness of random and heuristic depth-first search to the parallel case. If we have an upper bound  $p$  on the number of parallel processes we can build a domain as in figure 1 which does not induce 1 but  $p$  endless looping opportunities that either trigger our heuristic or may simply be taken due to random chance. In this way, using parallelism does not get us to completeness but it does improve the our chances of at least 1 process taking the correct path.

Figure 1: A pathological case in our new HTN heuristic



## 6.2 Loop Detection

This section will discuss loop detection as it is used in (TO)HTN planning in general. It will start with an overview over loop detection in other HTN planners in section 6.2.1. This is followed by a discussion of the simplifying assumptions we can make for TOHTN planning in section 6.2.2. - distributed loop detection (communication and merge operations become important!) - perfect loop detection - probabilistic loop detection (approximate membership query)

- [14] domains can have introduce infinite loops, without some kind of loop detection we loose completeness (and performance)

### 6.2.1 Loop Detection in Other HTN Planners

Loop detection in HTN planning is a recent phenomenon and was introduced in 2020 by the HyperTensioN planner ([14]) with the so-called 'Dejavu' technique. Dejavu works by extending the planning problem, introducing primitive tasks and predicates that track and identify when a particular recursive compound task is being decomposed. These new primitive tasks are invisible to the user. Information about recursive tasks is stored externally to the search as to not loose it during backtracking. Dejavu comes with performance advantages and protects against infinite loops. However, as Dejavu only concerns itself with information about the task network but ignores the world state it may have false positives. This was also noted by [11] and means that HyperTensioN is not complete. [11] further notes that the loop detection is limited in that it only finds loops in a single search path but cannot detect if multiple paths lead to equivalent states.

In response to [14], loop detection was introduced to the PANDA planner in [11]. Similar to HyperTensioN, PANDA keeps its loop detection information separate in a list of visited states,  $\mathcal{V}$ . Search nodes, identified by a tuple  $(s, tn)$  of world state  $s$  and task network  $tn$ , are only added to the fringe if they are not contained in  $\mathcal{V}$ . To speed up comparisons,  $\mathcal{V}$  is separated into buckets according to a hash of  $s$ . To then identify whether  $tn \in \mathcal{V}[s]$  multiple algorithms are proposed. In the sub-case of TOHTN planning, both an exact comparison of the sequence of open tasks as well as an order-independent hash of the open tasks, called *taskhash* are used. Similar to HyperTensioN, using a hash to identify equal task networks can lead to false positives and an incomplete planner. Both hashing-based and direct comparison as used in PANDA have a performance cost linear in the size of  $tn$ . The loop detection in PANDA improves upon the one in HyperTensioN insofar as it is able to detect loops where equivalent states were reached independently.

### 6.2.2 Assumptions in Loop Detection for CrowdHTN

To design the loop detection in CrowdHTN, we have both simplifying and complicating assumptions that we will discuss here.

While both PANDA and HyperTensioN are HTN planners, CrowdHTN concerns itself only with TOHTN planning. As a result, the remaining task network can be represented as a sequence of open tasks with the ordering constraints implicit in how the sequence is stored.

Crowd identifies search states as a tuple of  $(s, tn)$  of world state  $s$  and task network  $s$ , similar to PANDA and uses hashing to efficiently identify duplicate search states. As tasks of equivalent task networks are always in the same order, we can incorporate that order into our hash of  $tn$  to reduce the number of collisions compared to PANDA's *taskhash*. This will increase performance where we fall back to comparisons in case of collisions and reduce our false-positive rate in case

Make su  
to alway  
talk abo  
isomor-  
phic tas  
network  
instead  
equivale  
Quote [  
Define i  
morphis

quote al  
loop de-  
tection  
in graph  
search?

quote al  
distribu  
loop de-  
tection  
in graph  
search?

refer to  
impleme  
tation p  
where t  
open ta  
are repr  
sented a  
sequenc



we use probabilistic loop detection.

Both PANDA and HyperTension are sequential planners whereas CrowdHTN is highly parallel. This adds an additional design constraint to our loop detection. If we want to efficiently share information about visited states using the full state information becomes infeasible as those full states would have to be communicated. If we perform loop detection only locally, we predict to suffer from decreased performance the higher the degree of parallelism. I.e., if two branches of our search tree contain the same search node, the chance that those two branches are encountered on different processors is higher the more processors we have.

Refer to implementation about how we reduce hashing to O(1) time

### 6.2.3 Perfect Loop Detection

One simple way to perform loop detection which is similarly used in PANDA ([11]) is to use a hashset of visited states. The implementation in Crowd is slightly different from PANDA in that we use one combined hash for both world state and open tasks. Other than PANDA, CrowdHTN does incorporate the order of tasks into the hash, which should reduce collisions and makes the two levels of hashing less needed.

Using hashes combined with a full comparison provides us a perfect loop detection, i.e., neither false positives nor false negatives exist. This makes it a useful technique to benchmark other loop detection methods. However, both in the sequential and in the distributed case this technique suffers from performance problems.

In case of hash collisions, we have to fall back to a full comparison of world state  $s$  and open tasks  $tn$ . While  $s$  is bound in size by the total number of predicates, the size of  $tn$  is effectively unbound. Additionally, we have to keep both  $s$  and  $tn$  around for all nodes ever encountered, increasing the memory footprint of our program. In case of distributed loop detection, communicating full states would lead to a large communication overhead. Especially, we can expect each node to be larger than those sent as work packages, as those are optimized to have a small  $tn$  which would not hold for nodes encountered in our loop detection.

about exponential!

### 6.2.4 Statistical Loop Detection

**Approximate membership queries** - efficient merging is possible - communication is size efficient (especially does not depend on the number of workers!)

**Completeness in the face of false positives** - we lose completeness

- go back to something like HyperTension, accepting some false positive rate - approximate membership query - we get either 'definitely contained' - or 'maybe contained' - examples are bloom filter and quotient filter

Quotient filter: - do we need the original elements to re-insert them? - do we need to communicate whole hashes to combine filters? (worse communication size!)

Advantages: - communication takes less effort (a lot lower size of data!) - hashing in  $O(1)$ ? Or at least in a lot easier - we can pre-hash the open tasks (can be of exponential depth (requires a good argument, as the path down the task network could be of width 1?)), thus turning hashing into  $O(1)$  - in case of hash collisions we need to walk the whole sequence of open tasks - in practice this is even worse, as our open tasks are saved in a tree-like manner to allow sharing of the tasks between search nodes - this means we wildly jump through memory for hashing, adding another constant factor - we could save on this constant factor by duplicating the open tasks for each node and saving them sequentially, but leading to a lot higher memory footprint

approximate membership queries

about a quotient filter at the end! (or at least why is not in)

the same



(might be quadratic compared to what it is already (if any fixed fraction of nodes have at least 2 children, maybe? )) - compromise between performance and false-positive rate (better for correctness than just comparing a single hash) some m

### 6.2.5 Distributed Loop Detection

- two new considerations - memory footprint becomes more important for communication (effectively an all-to-all operation -> quote Sanders' book!) - efficient merge of loop detection data is needed - communicating everything might still be inefficient for bloom filters ()

## 7 Grounding and Pruning

[12] - a ground HTN instance has (worst-case) a size that is exponential in the arity of task names and predicates

## 8 Malleability in the Context of CrowdHTN

communicating the problem at hand the problem itself: - communicate the problem and domain files

- upgrade from moldable to malleable task according to [9] as introduced in section 2.5 - we can no longer assume that messages always arrive and get a response - we want to achieve completeness - when sending a message, we always want to send it to another As discussed in section 2.6, CrowdHTN in its original form is a moldable program, with the number of parallel workers fixed for each single execution. To adapt CrowdHTN to work as a malleable program within Mallob, we have to address three key concerns.

- (i) Messages sent to workers who are suspended or terminated before it arrives
- (ii) Integrating new processes to work efficiently
- (iii) Dealing with workers dying without losing completeness

To help understanding, this section explicitly discerns between Mallob workers and Crowd workers.

### Messages sent to dead workers

**Receiving worker stays dead** As Mallob runs each worker in a different processes communicating via message passing, we never have a full view of the current global state of our process, i.e., which other Mallob workers are assigned to the same job. Instead we only ever get information about this that might be outdated. As a result, when a message such as a work request is sent to another worker, we cannot be sure that the receiving worker is in a position to actually respond. Luckily, Mallob does provide us with a mechanism to detect such messages. Each Mallob worker knows which job it currently works on and all messages are tagged with their job as well. If a message belonging to job  $J_i$  is received by a worker working on job  $J_k, k \neq i$ , the message is simply tagged with a *returned to sender* flag and sent back. Now we can simply adapt CrowdHTN to deal with each message both if it is received normally and if it is received as a return message. On normal messages, nothing changes. On return messages we do the following:

- **Work request:** we treat this like a negative response. This means we set the *has active work request* flag to false and simply send out another work request at the next chance.
- **Positive work response:** this is the node we sent out ourselves. As to not lose any information, we simply re-insert the node at the front of our local fringe. While this may slightly change the order of the nodes at the front, we expect the number of outgoing work packages at any point to be small and the disturbance to be limited. In case of random searches, it should not negatively affect our algorithm.
- **Negative work response:** we do not have to care about this and can simply ignore it.
- **Ack:** we do not have to care about this and can simply ignore it.

While this mechanism should deal with most cases of workers dying, we can always construct pathological cases in which the return message mechanism fails. One such exchange can be seen in graphic 2. One way of dealing with this would be to change the way Mallob handles messages whose receiver is no longer valid. E.g., if a return message was delivered to a worker who changed jobs in the meantime, one could instead find out the root worker of the accompanying job and send the message there instead. This way no information would be lost until we decide to terminate the whole job at which point this would be unavoidable and okay by the user. To avoid having to change Mallob, we instead elected to make our adaptation of CrowdHTN resistant to such information loss.

**Receiving worker gets replaced by a new worker** - we have no way to detect this case - some messages we can absorb - a lone ack: we don't care, instead of decreasing the number of

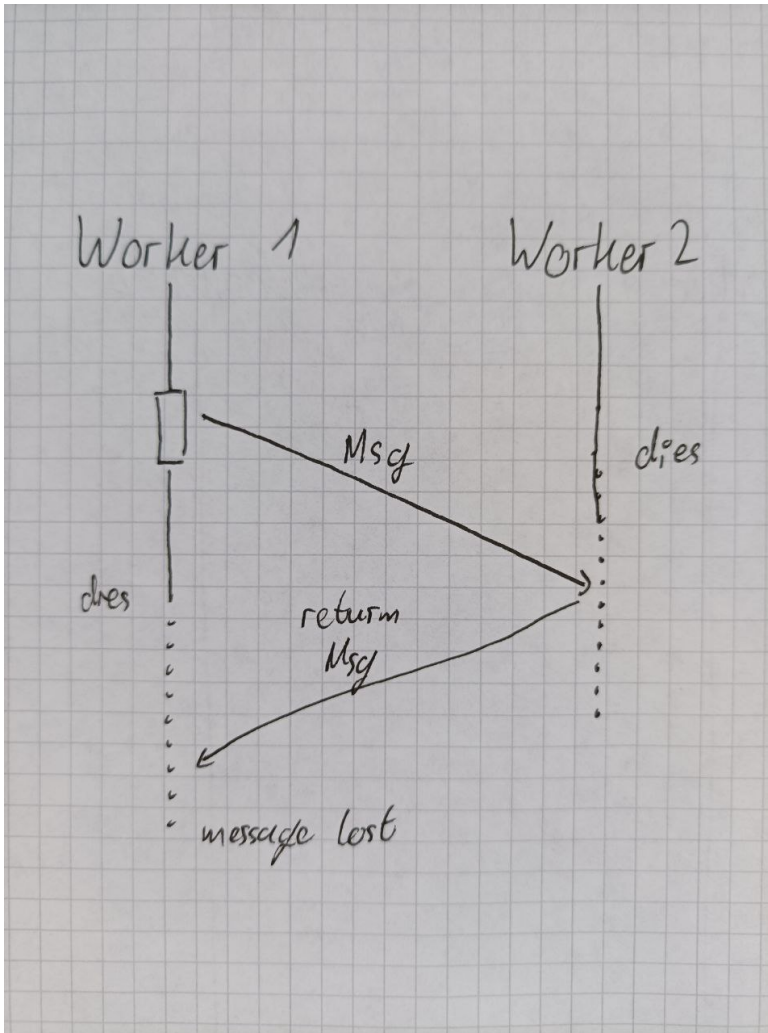


Figure 2: Example diagram of a lost message

outgoing packages we set the count to  $\max(0, \text{count} - 1)$  - a work request: can be responded to as if it was meant for us - work responses pose a problem: a node may send out a work request, die, a new node starts, starts its own work request and receives the response afterwards - dafuq, this is complicated - buuut, it can happen - what if we just accept it? Worst case we get two positive responses, take both of them and deal with it - number can decrease again if two negative responses arrive before sending off a new request - cycle will interrupt anyways upon the first positive response (yay!) - everything is fine dot jay pee gee

**Integrating new workers** As CrowdHTN uses work stealing to distribute work and perform load balancing, new workers can be seamlessly integrated into a running job. Upon construction all workers are treated the same, whether they are part of the initial batch or appear at a later point. The root worker is initialized with the initial search node, all other nodes are empty. To get new work, a worker sends a work request to a random other node and then goes from there. The very same process can be used to integrate a new worker.

**Dealing with the information loss of dying workers** As mentioned before, dying workers can lead to a loss of information. A previous paragraph discussed what this means for messages sent between workers. There is a second part of dying workers that we need to adjust to - the loss of the local fringe. There are three main ways in which we could deal with this loss.

- Communicate the local fringe to the parent or root worker
- Communicate the root search node of the local search to the parent or root worker
- Accept the loss of information and thus of parts of our search space

We will discuss these strategies one by one.

**Communicate the local fringe** - cleanest solution in a way - no information is ever lost - problems: - time needed for encoding the fringe is unbounded - an efficient encoding of the shared open tasks and world state would complicate the implementation - we need to ensure that the message containing our local fringe is not lost - i.e. it has to be sent to the root worker as it is guaranteed to live as long as anyone else - this imposes additional memory pressure on our root if many nodes terminate

**Communicate the root back** - this is guaranteed to not lose any information - however, we will have to re-do some work - communication of only a single node is rather easy, we already routinely do this - the pressure on the root worker is not too bad either, we could even re-integrate the node into the local fringe without much trouble - poses question regarding global loop detection - to not lose any part of our search space, we have to remove any node from global loop detection which originated from the dying worker, i.e. remove conservatively to not lose completeness - i.e., we can use this design efficiently but will have to re-do a lot of work

**Accept the information loss** - the simplest thing to do - we do not disturb the randomness of our work stealing by imposing additional work onto the root worker - performance is no problem at all - however, we lose parts of the search space - in combination with distributed bloom filters this may be the best solution either way - we simply wait for a future restart in which we do not lose the specific work

## **9 Experimental Evaluation**

## **10 Future Work**

## 10.1 A Common (TO)HTN Interface

- be able to mix grounders, pruners, planners - maybe even loop detection! - be able to combine different pruners - use a lifted pruner early and then plug in a grounder and grounded pruner afterwards - PANDA seems to try to provide this - quite successful regarding the input format for HTN planning which finally makes planners easily comparable, advancing the state of the art - sadly not as successful for the rest of planning - we believe their to be great potential in

## 10.2 Combine Pruning Approaches

- HyperTension and Panda base their pruning on different paradigms - HyperTension's pruning could be extended to work on grounded instances where it would have even more information available - both these approaches are structurally different and could thus be used to compliment each other - whether the better pruning would offset the increased runtime cost remains to be seen

## 10.3 Investigate the Completeness of Planners

- some planners are known to not be complete - some planners can easily be shown to be complete - we make a case in section that heuristic planners may not be complete - to catch these cases planners may have to get more intelligent about which parts of the search space to cut off - else we may have to adapt our search algorithms to enforce completeness at another level - similarly we may achieve completeness by simply cutting off the search once we get too deep - however, this may be theoretically complete but also hits us with the full expense of an EXPSPACE-hard problem, i.e. we may never realistically reach this condition (except on very small problems) - similarly it is not clear whether an algorithm like heuristic + plan length search gives us completeness in feasible time or more theoretically - HyperTension, a decidedly non-complete planner is also the best-rated planner in the IPC 2020 - maybe we need to have a full conversation on how much completeness we actually want/ need (or two separate categories of planners between 'HTN as a way to provide advice into planning domains' and 'HTN for the full power it provides')

## 10.4 Lifted Parallel Search

- the Crowd way of performing relatively simple search does not work out in the end - we can still see some scaling for the parallel case - lifted planning could massively shrink our search space (by an exponential factor in the number of nodes!) - lilotane provides much intelligence on how to perform lifted search - same as HyperTension - a lilotane-like behavior may be the best from a practical perspective, as it is more consistent - a search-based formulation may be easier to both parallelize and adapt to a malleable context - any search may be plugged in where CrowdHTN currently resides

## 10.5 Improvements Unlocked by a Shared Memory Implementation

- we have decided to stay with CrowdHTN's way of performing grounding on-demand and just in time, as it lends itself particularly well to a malleable environment - specifically, having to perform a big chunk of work that is not yet possible to lead to a plan would impose a problem on the efficiency of new and short-lived workers - in a shared-memory environment, this could be



done once by the root and then re-used by other workers, making better pruning and heuristics such as in PANDA available

## References

- [1] ALFORD, RON, GREGOR BEHNKE, DANIEL HÖLLER, PASCAL BERCHER, SUSANNE BIUNDO and DAVID W AHA: *Bound to plan: Exploiting classical heuristics via automatic translations of tail-recursive HTN problems*. In *Twenty-Sixth International Conference on Automated Planning and Scheduling*, 2016.
- [2] BEHNKE, GREGOR, DANIEL HÖLLER and SUSANNE BIUNDO: *Finding Optimal Solutions in HTN Planning-A SAT-based Approach*.
- [3] BEHNKE, GREGOR, DANIEL HÖLLER and SUSANNE BIUNDO: *On the complexity of HTN plan verification and its implications for plan recognition*. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 25, pages 25–33, 2015.
- [4] BEHNKE, GREGOR, DANIEL HÖLLER and SUSANNE BIUNDO: *totSAT-Totally-ordered hierarchical planning through SAT*. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.
- [5] BEHNKE, GREGOR, DANIEL HÖLLER and SUSANNE BIUNDO: *Tracking branches in trees-A propositional encoding for solving partially-ordered HTN planning problems*. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 73–80. IEEE, 2018.
- [6] BEHNKE, GREGOR, DANIEL HÖLLER, ALEXANDER SCHMID, PASCAL BERCHER and SUSANNE BIUNDO: *On succinct groundings of HTN planning problems*. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 9775–9784, 2020.
- [7] EROL, KUTLUHAN, JAMES HENDLER and DANA S NAU: *HTN planning: Complexity and expressivity*. In *AAAI*, volume 94, pages 1123–1128, 1994.
- [8] EROL, KUTLUHAN, JAMES HENDLER and DANA S NAU: *Complexity results for HTN planning*. *Annals of Mathematics and Artificial Intelligence*, 18(1):69–93, 1996.
- [9] FEITELSON, DROR G: *Job scheduling in multiprogrammed parallel systems*. 1997.
- [10] GEORGIEVSKI, ILCHE and MARCO AIELLO: *HTN planning: Overview, comparison, and beyond*. *Artificial Intelligence*, 222:124–156, 2015.
- [11] HÖLLER, DANIEL and GREGOR BEHNKE: *Loop Detection in the PANDA Planning System*. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 31, pages 168–173, 2021.
- [12] HÖLLER, DANIEL, GREGOR BEHNKE, PASCAL BERCHER and SUSANNE BIUNDO: *Language Classification of Hierarchical Planning Problems*. In *ECAI*, pages 447–452, 2014.
- [13] HÖLLER, DANIEL, PASCAL BERCHER, GREGOR BEHNKE and SUSANNE BIUNDO: *HTN planning as heuristic progression search*. *Journal of Artificial Intelligence Research*, 67:835–880, 2020.
- [14] MAGNAGUAGNO, MAURÍCIO C, FELIPE RECH MENEGUZZI and LAVINDRA DE SILVA: *HyperTension: A three-stage compiler for planning*. In *Proceedings of the 30th International Conference on Automated Planning and Scheduling (ICAPS)*, 2020, França., 2020.
- [15] MALI, AMOL DATTATRAYA and SUBBARAO KAMBHAMPATI: *Encoding HTN Planning in Propositional Logic*. In *AIPS*, pages 190–198, 1998.
- [16] NAU, DANA, YUE CAO, AMNON LOTEM and HECTOR MUNOZ-AVILA: *SHOP: Simple hierarchical ordered planner*. In *Proceedings of the 16th international joint conference on Artificial intelligence-Volume 2*, pages 968–973, 1999.
- [17] NAU, DANA S: *Current trends in automated planning*. *AI magazine*, 28(4):43–43, 2007.

- [18] RAMOUL, ABDELJALIL, DAMIEN PELLIER, HUMBERT FIORINO and SYLVIE PESTY: *Grounding of HTN planning domain*. International Journal on Artificial Intelligence Tools, 26(05):1760021, 2017.
- [19] SCHREIBER, DOMINIK: *Lilotane: A lifted SAT-based approach to hierarchical planning*. Journal of Artificial Intelligence Research, 70:1117–1181, 2021.
- [20] SCHREIBER, DOMINIK, DAMIEN PELLIER, HUMBERT FIORINO et al.: *Tree-REX: SAT-based tree exploration for efficient and high-quality HTN planning*. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 29, pages 382–390, 2019.
- [21] SCHREIBER, DOMINIK and PETER SANDERS: *Scalable SAT solving in the cloud*. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 518–534. Springer, 2021.