PARIS: Planning Algorithms for Reconfiguring Independent Sets

Abstract
Combinatorial reconfiguration studies how one solution of a combinatorial problem can be transformed into another. The transformation can only make small local changes and may not leave the solution space. An important example is the independent set reconfiguration (ISR) problem, where an independent set of a graph (a subset of its vertices without edges between them) has to be transformed into another one by a sequence of modifications that remove a vertex or add another that is not adjacent to any vertex in the set. The 1st Combinatorial Reconfiguration Challenge (CoRe Challenge 2022) was a competition focused on the ISR problem. Our team participated with two solvers that model the ISR problem as a planning problem and employ different planning techniques for solving it. They successfully competed in the challenge and were awarded 4 first, 3 second, and 3 third places across 9 tracks. In this work, we present the ISR problem as a new problem to the planning community and describe the planning techniques used in our solvers. We re-ran the entire competition under equal computational conditions to allow for a fair comparison. Besides showcasing the success of planning technology, we hope that this work will create a cross-fertilization of the two research fields.

Keywords
classical planning, independent set, combinatorial reconfiguration

1. Introduction

Combinatorial reconfiguration studies the space of solutions for combinatorial problems. The task is to transform one solution of a combinatorial problem into a different one, without leaving the space of solutions. Each transformation can only make a small local change to the current solution. The term was coined by Ito et al. [1] who show that there is a host of problems derived from NP-complete (combinatorial) problems that fall into the category of combinatorial reconfiguration problems and that they are PSPACE-complete. Two prominent examples for reconfiguration tasks are propositional satisfiability [2] and graph k-coloring [3]. But probably the most well-studied representative of combinatorial reconfiguration tasks is the *independent set reconfiguration* (ISR) problem [4].

An independent set of a graph is a subset of its vertices such that no two vertices of the subset share an edge. Reconfiguring an independent set means replacing one vertex in the subset with another one such that the new subset is still an independent set. The ISR problem is to find a sequence of such reconfiguration steps to reach a given target configuration from a given start configuration. The problem is PSPACE-complete [5], which means it is as hard as automated planning [6].

The *1st Combinatorial Reconfiguration Challenge* (CoRe 2022)[1] is a competition that compares practical combinatorial reconfiguration algorithms. Its first instantiation targeted the ISR problem, featuring different tracks. We participated in the competition using two solvers that model ISR problems as planning tasks and use various planning techniques for solving them. Among the seven teams that participated, our solvers achieved 4 first, 3 second, and 3 third places across all tracks, winning the majority of awards.

In this work, we present the ISR problem and explain how we can model it as a planning problem. We describe the technology used in our solvers, which is mostly based on planning techniques, including a technique for detecting unsolvable problems which we believe to be useful for unsolvability planning in general. For space considerations, we focus only on those tracks (i.e., the "solver tracks") where our approach relied heavily on planning technology. Furthermore, since competitors of the competition ran their solvers themselves using different hardware and resource limits, we re-ran all of them under equal computational conditions and report the results in this work. Besides showcasing the success of planning technology, we also introduce a problem that is new to our community. We believe this will lead more planning researchers to develop ideas for the ISR problem and create a cross-fertilization of the fields.

2. Background

A *graph* is a pair $G = \langle V, E \rangle$, where V is a set of *vertices* and $E \subseteq \{\{u, v\} \mid u, v \in V, v \neq u\}$ is a set of *edges* between the vertices. An *independent (vertex) set* of a graph G is a subset of vertices $I \subseteq V$ such that no two vertices in the subset I are edges of G, i.e., for all $v, u \in I$ it holds that $\{v, u\} \notin E$.

CEUR Workshop Proceedings (CEUR-WS.org)
[1] https://core-challenge.github.io/2022

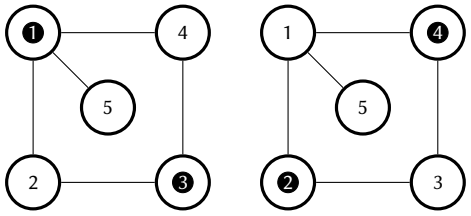

Figure 1: Visualization of the independent set reconfiguration problem described in Example 1 with a graph consisting of five nodes, two tokens depicted in black, the start independent set I_s (left), and the target independent set I_t (right).

2.1. Independent Set Reconfiguration

Similar to Kaminski et al. (2012), we consider an independent set as a set of tokens placed on the vertices of a graph G, called *token configuration*, such that no two tokens are adjacent. The *token jump* reconfiguration rule describes how to transform one token configuration into another, moving a token from one vertex to any other unoccupied vertex, so that the resulting configuration again describes an independent set. Note that the token can jump, i.e., it does not have to move along an edge. Given the reconfiguration rule, we define a reconfiguration sequence $\rho = \langle I_0, \dots, I_n \rangle$ as a sequence of non-repeating independent sets, where each set I_i with $1 \leq i \leq n$ results from a single token jump from the previous set I_{i-1}. The length $|\rho|$ of a reconfiguration sequence $\rho = \langle I_0, \dots, I_n \rangle$ is the number of token jumps inducing the sequence, i.e., $|\rho| = n$. The Independent Set Reconfiguration decision problem [4] is defined as follows.

Definition 1 (Independent Set Reconfiguration). *Given a graph G and two independent sets I_s and I_t, the* independent set reconfiguration (ISR) decision problem *is to determine whether there exists a reconfiguration sequence $\rho = \langle I_s, \dots, I_t \rangle$.*

The ISR problem is one of the most prominent representatives of combinatorial reconfiguration. It is known to be PSPACE-complete for general input graphs [4, 5] and formed the central problem of CoRe 2022.

Example 1. *Figure 1 shows an ISR problem with the start set $I_s = \{1, 3\}$ and the target set $I_t = \{2, 4\}$. A solution to this problem is the reconfiguration sequence $\rho = \langle \{1, 3\}, \{3, 5\}, \{2, 5\}, \{2, 4\} \rangle$, where first the token at node 1 is moved to node 5, then the token at node 3 is moved to node 2, and finally the token at node 5 is moved to node 4. This sequence has a length of $|\rho| = 3$, since it performs three jumps and is the shortest sequence that solves the problem.*

2.2. Combinatorial Reconfiguration Challenge

Similar to the International Planning Competition (IPC), CoRe 2022 featured different tracks. They can be separated into two main categories: graph tracks and solver tracks.

Graph Tracks In the *graph* tracks the objective was to construct an ISR instance such that the shortest reconfiguration sequence is as long as possible. For CoRe 2022 there were three graph tracks, one each for graphs with 10, 50 and 100 nodes, and the team that constructed the instance with the longest shortest reconfiguration sequence won the respective track.

Solver Tracks In total, there were three different solver tracks in CoRe 2022: the *existent*, the *shortest* and the *longest* track, each further subdivided into a *single solver* sub-track and a *portfolio solver* sub-track. In the *existent* track, each solver that provided a reconfiguration sequence for or detected unsolvability of an ISR instance received one point. In contrast, the *shortest* and *longest* tracks considered the quality of the solutions, and solvers that provided the shortest/longest (among the participants) reconfiguration sequence for an instance received one point.[2] The winning solver for each track was the one that received the most points across all benchmark ISR instances.

Note that the names *shortest* and *longest* are somewhat misleading. The aim in these tracks is to find a solution, aiming at as short/long loopless solutions as possible, but no guarantees on optimality are needed. To draw parallels between these tracks and International Planning Competition (IPC) tracks, the *shortest* track is actually more similar in that respect to the *satisficing* IPC track. Currently, there is no equivalent in planning competitions to the *longest* track. The *existent* track is somewhat similar to the *agile* IPC track.

2.3. Classical Planning

In this paper, we propose to model the ISR problem as a classical planning problem. For this, we consider the *Planning Domain Definition Language* (PDDL) [7] and the SAS$^+$ formalism [8] to describe classical planning problems. A *(classical) planning problem* is a concise representation of a transition system with a single initial state, a compact description of the set of goal states, and a set of actions with preconditions and effects that describe the transitions. The objective is to derive a course of action that transforms the initial state into one of the

[2] Reconfiguration sequences must be non-repeating. Therefore, participants must search for loopless solutions in the longest track.

```
(:action move
    :parameters (?l1 ?l2 - loc)
    :precondition (and
        ; Source has token
        (tokened ?l1)
        ; Destination is free
        (free ?l2)
        ; Destination's neighbors are free
        (forall (?l3 - loc) (imply
            (and (not (= ?l1 ?l3))
                 (edge ?l2 ?l3))
            (free ?l3))))
    :effect (and
        ; Source is free
        (not (tokened ?l1)) (free ?l1)
        ; Destination has token
        (tokened ?l2) (not (free ?l2))))
```

Listing 1: Single PDDL encoding using one move action.

goal states. While the full details of PDDL are beyond the scope of this paper and are not necessary to follow the content of the paper, the excerpts presented in this paper suffice to present our contributions. For a more detailed account, we refer the reader to Haslum et al. (2019).

An SAS$^+$ task formally is a tuple $\langle \mathscr{V}, \mathscr{A}, \mathscr{I}, \mathscr{G} \rangle$, where $\mathscr{V}$ is a finite set of *variables* V, each with a finite domain $dom(V)$, $\mathscr{A}$ is a finite set of *actions*, $\mathscr{I}$ is the *initial state*, and $\mathscr{G}$ is the *goal*. Partial variable assignments p map a subset of variables $vars(p) \subseteq \mathscr{V}$ to values in their domain. Variable assignments s with $vars(s) = \mathscr{V}$ are called *states*. A partial variable assignment p is satisfied in a state s if p and s agree on $vars(p)$. Each action $a \in \mathscr{A}$ consists of a precondition $pre(a)$ and an effect $eff(a)$, both partial variable assignments. An action is applicable in a state if its precondition is satisfied, and applying it updates the state with values defined in its effect. A planner finds a sequence of actions that is sequentially applicable and leads from the initial state $\mathscr{I}$ to some state satisfying the partial variable assignment $\mathscr{G}$.

3. Planning Encoding

The planning domain definition language (PDDL) is the de-facto standard language for modeling planning tasks [9], and most planning tools are built with PDDL as their input language. The ISR problem can be encoded in PDDL by introducing a single lifted action to *move* a token from one location to another. Listing 1 shows the PDDL code for this *move* action, with comments interleaved. We call this the *single* encoding. While the encoding itself is quite compact, grounding these tasks is slow. In an ISR instance with n nodes, n^2 *move* actions have to be created.

```
(:action pick
    :parameters (?l1 - loc)
    :precondition (and
        ; Not holding a token
        (handfree)
        ; Source has token
        (tokened ?l1))
    :effect (and
        ; Holding a token
        (not (handfree)) (holding)
        ; Source is free
        (free ?l1) (not (tokened ?l1))))

(:action place
    :parameters (?l1 - loc)
    :precondition (and
        ; Holding a token
        (holding)
        ; Destination is free
        (free ?l1)
        ; Destination's neighbors are free
        (forall (?l2 - loc) (imply
            (edge ?l1 ?l2) (free ?l2))))
    :effect (and
        ; Not holding a token
        (not (holding)) (handfree)
        ; Destination has token
        (not (free ?l1)) (tokened ?l1)))
```

Listing 2: Split PDDL encoding using two actions.

As we are dealing with graphs of up to 40000 nodes, this can be problematic. To overcome this issue, we tested two approaches. The first is manual pre-grounding, called *single-grounded*. This does not help with the quadratic number of actions but avoids overhead creating the SAS$^+$ representation. The second approach, called *split*, is to split the move action into two actions, *pick* and *place*. It is presented in Listing 2. In this encoding, we only need $2n$ actions but plans are twice as long and have to be post-processed. Even this encoding can be slow to ground and can be sped up significantly with pre-grounding, which we call *split-grounded*. Ultimately, we found the *split-grounded* encoding to be the most efficient, and so we used it for all tracks and solvers.

The planning systems we used are all built on the Fast Downward planning system [10], which first translates the input PDDL into SAS$^+$ [8] before searching for a plan. While we used the aforementioned PDDL encodings for the bulk of the development work for the contest, our final submission directly encodes the input tasks into the *split* SAS$^+$ format to save on the computational effort required by this translation.

We encode a given ISR problem $\langle G, I_s, I_t \rangle$ with a graph $G = \langle V, E \rangle$, as an SAS$^+$ task $\langle \mathscr{V}, \mathscr{A}, \mathscr{I}, \mathscr{G} \rangle$ in the follow-

ing way. The variables $\mathcal{V} = V \cup \{hand\}$ contain one binary variable for each node in the graph to represent if there is a token on this node, and a binary variable *hand* to represent if we are currently holding a token. The domain of all variables is $\{free, occupied\}$. The initial state is $\mathcal{I} = \{v \mapsto occupied \mid v \in I_s\} \cup \{v \mapsto free \mid v \in V \setminus I_s\} \cup \{hand \mapsto free\}$, and the goal is $\mathcal{G} = \{v \mapsto occupied \mid v \in I_t\} \cup \{v \mapsto free \mid v \in V \setminus I_t\} \cup \{hand \mapsto free\}$. Note that specifying the occupied nodes in the goal would also be sufficient but specifying a value for all variables can help the planners realize that there is exactly one goal state.

The actions are analogous to the ones shown in Listing 2. There is an action $pick(v) \in \mathcal{A}$ for every $v \in V$ and it has the precondition $pre(pick(v)) = \{v \mapsto occupied, hand \mapsto free\}$ and effect $eff(pick(v)) = \{v \mapsto free, hand \mapsto occupied\}$. I.e., picking up a token is possible from all nodes that have a token, as long as we are not already holding one. Additionally, there is an action $place(v) \in \mathcal{A}$ for every $v \in V$ and it has the precondition $pre(place(v)) = \{v \mapsto free, hand \mapsto occupied\} \cup \{v' \mapsto free \mid \{v, v'\} \in E\}$ and effect $eff(place(v)) = \{v \mapsto occupied, hand \mapsto free\}$. So placing a held token is only possible on positions that currently have no token and have only free neighbors. The latter ensures every reachable configuration is an independent set.

4. Finding Solutions

We use sequential algorithm portfolios for each of the three solver tracks. That is, we run a sequence of algorithms, each with an associated time limit. The next section describes the algorithms that we use in our sequential portfolios.

4.1. Planning Algorithms

After testing various planning heuristics from the literature in exploratory experiments, we found that *landmark*-based heuristics to work well on ISR tasks. Relaxation-based heuristics, such as FF [11] and Red-black [12] did not contribute to search performance. Interestingly, both for satisficing and optimal planning, it is best to combine the landmark costs admissibly.

A*+Landmarks We run an A* search [13] with a *landmark count* heuristic [14] that uses two different kinds of landmarks: h^1 landmarks [15] and RHW landmarks [16]. The landmark costs are combined with *uniform cost partitioning* [17], which ensures that the resulting heuristic is admissible. As a result, this algorithm is optimal, sound, and complete, i.e., if it reports a plan, this is a shortest plan, if it reports unsolvability, the task is indeed unsolvable, and given sufficient resources, it will terminate.

GBFS+Landmarks We run a greedy best-first search (GBFS) [18] with a *landmark count* heuristic [14] over h^1 landmarks [15]. Again, the landmark costs are combined with *uniform cost partitioning*. This algorithm is sound and complete, but not optimal.

Symbolic Search We run a forward symbolic blind search [19, 20] using Binary Decision Diagrams [21] as the underlying data structure. The symbolic planner we use is SymK [22], which uses CUDD [23] as its decision diagram library. This search is optimal, sound and complete.

Symbolic Top-k Search The problem of finding a plan that is as long as possible is not commonly considered in the planning community, but only in the context of approximating the longest possible solution in SAT-based planning [24]. Interestingly, the search for the longest path in a compactly represented graph is NEXP-TIME-hard [25] and is therefore considered more complex than ordinary satisficing or optimal planning, which are known to be PSPACE-hard [6]. Cohen et al. (2020) investigated heuristic search for finding the longest path for a given explicitly represented graph. While this is an interesting line of research to be applied in the context of planning, in the CoRe 2022 challenge we were interested in finding a long plan, but not necessarily the longest.

To find long plans, we run a forward symbolic blind search based on the algorithm SymK-LL [27], implemented in the symbolic search planner SymK [22], which iteratively finds and generates all loopless plans for a task. However, we have made the following adjustments to find long loopless plans. First, once we find a goal state reachable with cost c, we reconstruct only one loopless plan with cost c and ignore all other plans with the same cost. Second, since the split encoding introduces intermediate states in which a token is held, we ignore these artificial states when evaluating if a plan is loopless during the plan reconstruction of SymK-LL. This algorithm iteratively finds longer plans, starting with the shortest one, and eventually finds the longest loopless plan, given enough resources.

Counter Abstraction We abstract the problem to a planning problem that counts how many tokens are in certain positions and check for unsolvability in the abstraction. Since this algorithm is new, we describe it in more detail in Section 4.6. We now describe our sequential algorithm portfolios. Our portfolio for the *existent* track is identical to the one for the *shortest* track.

	existent		shortest				longest				CoRe 2022 limits		
	#c	#e	#c	#e	c score	e score	#c	#e	c score	e score	time (s)	mem (GB)	cores
ReconfAIGERation	257	245	152	212	201.36	212.00	54	30	83.02	30.00	10000	128	4
junkawahara	122	144	110	128	110.00	128.00	21	29	44.16	55.87	600	32	1
PARIS	334	314	275	268	282.74	268.00	143	230	183.24	246.45	62610	16	32*
telematik_tuhh	326	302	280	266	280.00	266.00	27	31	76.51	88.23	144000	60	2
toda5603	207	209	134	77	164.36	115.97	31	67	60.45	107.24	~ 10000	32	1
recongo	244	238	238	234	238.00	234.00	115	25	155.93	25.00	12600	96	1

Table 1
Coverage results from both the competition (c) and our experiments (e). # indicates the total number of problems solved or found to be the shortest/longest. "score" refers to the IPC-style calculation over all of the problems (see text for further details). The last column reports the limits used by the teams in the competition. If different limits were used in different tracks, we report the maximum. Our solver mostly runs on 1 core but the MIP solver used by numerical abstractions used 32.

4.2. Portfolio for *shortest* and *existent* Tracks

The competition enforced no resource constraints and left it up to the competitors for how long they want to run their solvers. We decided on the following time limits for our portfolio based on some initial test that showed diminishing returns for higher limits. If one step in the portfoilo finds a solution, the remaining steps are skipped.

1. Counter abstraction: 10 seconds
2. Symbolic search: 70 minutes
3. A*+Landmarks: 70 minutes
4. GBFS+Landmarks: 70 minutes
5. Counter abstraction: 14 hours

Note that we use counter abstractions twice: first with a small time limit at the start of the portfolio to handle all cases where we can quickly prove unsolvability. Then again with a large time limit after all other components to catch unsolvable instances that are hard to prove unsolvable.

4.3. Single Solver for *shortest* and *existent* Tracks

We ran GBFS+Landmarks for 70 minutes as our single-solver submission because it has the highest coverage among all solvers in the portfolio.

4.4. Portfolio for *longest* Track

Our portfolio for the *longest* track ran two components: (1) GBFS+Landmarks: 330 seconds; and (2) Symbolic top-k search: 70 minutes. When GBFS+Landmarks finds a solution, we use the cost of that solution as the lower solution bound for the subsequent symbolic top-k search, so that only solutions that are longer than the solution we already have are reconstructed. As a fallback, if neither of the two approaches produced a solution longer than the shortest/existent tracks, we used the solution to the shortest/existent tracks as a default.

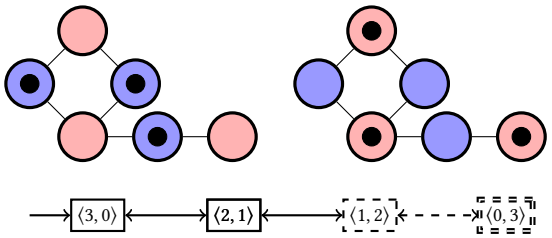

Figure 2: Example coloring for the counter abstraction approach. Top: initial state (left) and goal state (right). Nodes are colored blue if they have a token in the initial state but not the goal state and red if they have no token in the initial state but a token in the goal state. Bottom: the abstract state space. Dashed nodes are pruned.

4.5. Single Solver for *longest* Track

We ran symbolic top-k search for 70 minutes as our single-solver submission for the *longest* track. Note that we did not use the "fallback" option for this single-track submission.

4.6. Counter Abstraction

The *counter abstraction* component of our solver tries to detect if the task is unsolvable by abstracting it to a planning problem that counts the number of tokes in certain locations. This idea is inspired by counter abstractions in the area of model checking (e.g., Wahl and Donaldson 2010). Similar ideas where proposed in the area of planning as well [29]. In model checking, counter abstractions are usually used for symmetry reduction, whereas we do not require the abstracted parts to be symmetric to each other.

Given an ISR problem, we produce a *coloring* of the vertices in the graph, i.e., a function that maps each vertex of the graph to one color. Many different ways of coming up with a good coloring are conceivable but we opted for a simple strategy that uses up to four colors: one each for nodes that

- contain a token both in the initial and in the goal state;
- contain a token only in the initial but not in the goal state;
- contain a token only in the goal but not in the initial state;
- are empty in the initial and goal state.

Colors for situations that do not occur are not used. For example, the task in Figure 2 only requires two colors.

Given a coloring, each original state can be abstracted to a state with one counter variable per color that tracks how many tokens currently are on vertices with this color. For example, the initial state in Figure 2 has 3 tokens on blue nodes and 0 on red nodes, so it can be represented as the state $\langle 3, 0\rangle$. The goal has all three tokens on red nodes and none on blue, so it can be represented by $\langle 0, 3\rangle$. Moving a token from a node colored c_i to a node colored c_j changes the abstract state from $\langle c_1, \dots, c_i, \dots, c_j, \dots, c_n\rangle$ to $\langle c_1, \dots, c_i - 1, \dots, c_j + 1, \dots, c_n\rangle$. The main observation is that if any solution to the full problem exists, there has to be a solution in the abstraction as well. We thus construct the state space of the abstraction in the following way.

For a state $s = \langle c_1, \dots, c_n\rangle$, we construct one successor for each pair of unique colors c_i and c_j that differs from s by a single token that moved from c_i to c_j. In our running example, the initial state $\langle 3, 0\rangle$ has a single successor $\langle 2, 1\rangle$, and this state has two successors $\langle 3, 0\rangle$ (which we skip because we have already seen this state) and $\langle 1, 2\rangle$. The latter state now has the abstract goal $\langle 0, 3\rangle$ as a successor (Figure 2).

Whenever we generate a state, we check whether such a state is possible (independent of whether it is reachable). If it is not possible to place the tokens on the respective colors in the required way, we do not have to consider it or its successors. In our running example, the state $\langle 1, 2\rangle$ is not realizable: no matter where we place the blue token, it blocks two of the three red nodes. We use a mixed-integer program (MIP) solver to test if a state s is realizable by checking if the following system of constraints has a solution:

$$\begin{aligned} x_i + x_j &\leq 1 && \text{for all edges } \langle i, j\rangle \text{ in the graph} \\ \sum_{i \in N_c} x_i &\geq s[c] && \text{for all colors } c \\ x_i &\in \{0, 1\} && \text{for all nodes } i, \end{aligned}$$

where N_c is the set of all nodes with color c and $s[c]$ is the amount of tokens that should have color c in state s. The abstract state s is realizable iff the constraints have a solution.

If we generate a state that matches the goal state ($\langle 0, 3\rangle$ in our example), we know that there is an abstract plan. In this case, we still do not know if there is a real plan and return `unknown` (this component of the solver is incomplete). However, if there is no solution to the abstract problem, there cannot be a solution to the original problem. The abstract state space is usually small enough to explore completely. In our running example, it only has 4 states, and we only have to explore 3 of them, as we prune state $\langle 1, 2\rangle$.

While the MIP we use to check for realizability of abstract states is specific to ISR, the rest of the technique is domain-independent, and we will explore this further in the future.

4.7. Other Competitors

Across all solver tracks, seven teams competed at CoRe 2022, including our team (PARIS). Three of them were classified as portfolios: our portfolio, the submission by Turau and Weyer (telematik_tuhh), and the one by Froleyks, Yu, and Biere (ReconfAIGERation) in the existent track.

The solver telematik_tuhh tackles the problem by searching in the space of independent sets with two algorithms running concurrently: an iterative deepening A^* search using the number of misplaced tokens as heuristic value for finding optimal solutions, and a breadth-first search for detecting unsolvability. These algorithms are enhanced by domain-specific successor generation and memory optimization.

In the existent track ReconfAIGERation first transforms the problem to circuits represented as and-inverter graphs in the AIGER format [30], and then solves them with ABC [31], a model checker that runs several algorithms concurrently. In the other tracks it represents tasks as SAT formulas encoding increasingly longer reconfiguration sequences. The resulting bounded model checking problems are solved by the incremental SAT solver CaDiCaL [32].

Among non-portfolio entries, the one by Yamada, Kato, Kosuge, Takeuchi, and Banbara (recongo) achieved strong results. They translate instances into answer set programs and leverage clingo [33] as an off-the-shelf solver. Toda (toda5603) employs a modular strategy by initially running a greedy search and directly returning its suboptimal solution upon success. If it does not reach the goal, the problem is recast to a bounded model checking task where the state reached by the search is the initial state. This step is further informed by edge clique covers computed by ECC [34] and solved by the bounded model checker NuSMV [35].

Kawahara and Yamazaki (junkawahara) work with

families of independent sets, such as the initial independent set, or the family of all independent sets. They represent such families as zero-suppressed binary decision diagrams [36] and generate successors using set operations on ZDDs implemented using Graphillion [37].

Lastly, Blé, Cui, Wu, and Zhong (tigrisg) rely on a state-action-reward-state-action approach. We refer to Soh et al. [38] for the full solver descriptions.

5. Evaluation

In our experimental setup, we converted the docker images of each competing solver to singularity images (for improved performance) and ran all solvers in a unified setup with 2 hours timeout, 60 GB memory and 10 cores. All evaluations were run on Intel Xeon Silver 4114 processors running at 2.2 GHz. We could not include team tigris in the experiment since we could not run their docker container, and their team lost contact with the person who created it.

The shift in evaluation methodology is worth highlighting. In contrast with the contest parameters (where competitors were welcome to run their methods on their own hardware without any real resource constraints), we wanted to have a uniform analysis of the various approaches. This mitigates any bias that may stem from one team's computing infrastructure being more formidable than another. We can also see in the last columns of Table 1 that teams allocated very different amounts of resources to their solvers. By fixing one set of limits, we might bias the results towards a solver but we tried to select limits sufficiently high that all solvers can show their strengths and we will analyze the performance for lower time limits as well.

The first columns of Table 1 compare the coverage results we obtained with the ones from the competition. In most cases, coverage dropped compared to the competition since the competition gave no restrictions on resource usage and most submissions had a significantly higher timeout. Team junkawahara is an exception with only running their solver for 10 minutes for the competition, and gained 22 tasks when using a 2-hour time limit. The portfolio approaches ReconfAIGERation, PARIS and telematik_tuhh lost the most coverage in the *existent* track, since their resource allocation was built for an increased resource limit. However, the relative ordering of performance among the solvers remains.

For the *shortest* and *longest* tracks, solvers only gained coverage for a task if their solution was the best one amongst all competitors. This makes an analysis between the competition and our evaluation difficult since different best solutions might have been found. We note that for *shortest*, ReconfAIGERation shows improved performance, most likely because their submission used only 32GB of memory while we used 60GB. For *longest*, both ReconfAIGERation and recongo dropped significantly since they used a much higher time/memory limit in the competition; while PARIS performed significantly better. The latter is because in the competition, our submission for *longest* used the solutions from *shortest* as a seed to find longer plans, and we accidentally passed information that was processed incorrectly when we did not find a solution for *shortest*. For this experiment, we instead recomputed a (not necessarily shortest) plan in the beginning and handled the case of no found plan correctly.

	existent		shortest		longest	
	+	-	+	-	+	-
ReconfAIGERation	70	1	56	0	201	1
junkawahara	181	10	150	10	215	14
telematik_tuhh	22	10	10	8	200	1
toda5603	106	1	191	0	209	46
recongo	77	1	35	1	208	3

Table 2
Per-task comparison showing how often PARIS performed better (+) or worse (-).

We also included a scoring function that gives partial points for finding some solution; for *shortest* it is the ratio of the minimal reported solution and the found solution (analogous to the quality score used in IPC), for *longest* it is the ratio of the found solution and the maximal reported solution. The score suggests that in *shortest*, many solvers compute only minimal length solutions since their score is identical to their coverage. The picture is quite different for *longest*, showing that solvers that performed poorly often did find a decently long solution but not the longest overall.

Table 2 reports the number of tasks PARIS solved that others did not (+) and vice versa (-). It shows that overall junkawahara is the most complimentary to our approach, with telematik_tuhh also solving some problems we could not on existing and *shortest*, and toda5603 solving the most problems we did not on *longest*.

Figure 3 shows how many tasks each solver solves within a given time limit. In most cases, many problems are solved early on with more problems only trickling in slowly. The exceptions are the portfolio approaches, showing a sharp increase in coverage around the time the next component is started. A more sophisticated interleaving of the portfolio components could be used to smooth this process.

Finally, we reran each component of our *existent* portfolio separately to analyze their contribution. We report for each component how many tasks it could solve within 70 minutes and 16GB while none of the previous component solved it. Symbolic search alone solved

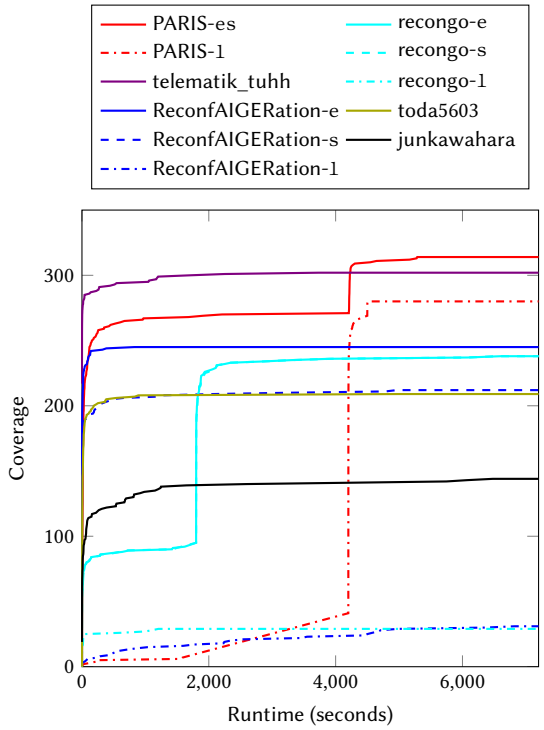

Figure 3: Number of tasks solved within a given time limit. If the same team had dedicated solvers for certain tracks, the track is indicated in the letter(s) behind the hyphen.

223 problems, while A*+Landmarks added another 45, GBFS+Landmarks 16 more and finally the counter abstraction detected 39 problems as unsolvable. In the competition we ran the counter abstraction for longer, leading to 7 more problems detected unsolvable, bringing the total coverage up to 330. The remaining difference from our competition result (334) is most likely due to hardware differences and noise. We also compared how many tasks could only be solved by a single approach: 15 for symbolic search, 0 for A*+Landmarks, 16 for GBFS+Landmarks and 39 for the counter abstraction. While A*+Landmarks is dominated by GBFS+Landmarks, it returns optimal solutions, making it an important contributor for the *shortest* track. The counter abstraction was invaluable for detecting unsolvable tasks, since no other component could decide any of the tasks that were solved by it.

6. Conclusions

In this paper, we introduced the independent set reconfiguration problem, one of the most-studied problems of combinatorial reconfiguration, as a testbed for planning algorithms. We modeled this problem as a planning task and adapted different planning techniques for solving it, including a new technique for detecting unsolvable ISR problems that we think can be generalized to unsolvability planning in general. The resulting solvers participated successfully in the 1st Combinatorial Reconfiguration Challenge (2022), winning the majority of awards in multiple tracks. We re-ran all solvers of the competition under equal computational conditions for a more thorough analysis and investigated the strengths and weaknesses of our planning-based solvers. Our findings show that the independent set reconfiguration problem is an interesting and challenging problem for planning, and our algorithms are currently among the best approaches for solving it. We believe these findings will lead more planning researchers to develop ideas for the ISR problem and create a cross-fertilization of the fields.

References

[1] T. Ito, E. D. Demaine, N. J. A. Harvey, C. H. Papadimitriou, M. Sideri, R. Uehara, Y. Uno, On the complexity of reconfiguration problems, in: Proc. ISAAC 2008, 2008, pp. 28–39.

[2] T. Ito, E. D. Demaine, N. J. A. Harvey, C. H. Papadimitriou, M. Sideri, R. Uehara, Y. Uno, On the complexity of reconfiguration problems, Theoretical Computer Science 412 (2011) 1054–1065.

[3] L. Cereceda, Mixing graph colourings, Ph.D. thesis, London School of Economics and Political Science, 2007.

[4] M. Kaminski, P. Medvedev, M. Milanic, Complexity of independent set reconfigurability problems, Theoretical Computer Science 439 (2012) 9–15.

[5] N. Nishimura, Introduction to reconfiguration, Algorithms 11 (2018) 52.

[6] T. Bylander, The computational complexity of propositional STRIPS planning, AIJ 69 (1994) 165–204.

[7] D. McDermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, D. Wilkins, PDDL – The Planning Domain Definition Language – Version 1.2, Technical Report CVC TR-98-003/DCS TR-1165, Yale Center for Computational Vision and Control, Yale University, 1998.

[8] C. Bäckström, B. Nebel, Complexity results for SAS$^+$ planning, Computational Intelligence 11 (1995) 625–655.

[9] P. Haslum, N. Lipovetzky, D. Magazzeni, C. Muise, An Introduction to the Planning Domain Definition Language, volume 13-2 of *Syn-*

thesis Lectures on Artificial Intelligence and Machine Learning, Morgan & Claypool, 2019.
[10] M. Helmert, The Fast Downward planning system, JAIR 26 (2006) 191–246.
[11] J. Hoffmann, B. Nebel, The FF planning system: Fast plan generation through heuristic search, JAIR 14 (2001) 253–302.
[12] C. Domshlak, J. Hoffmann, M. Katz, Red-black planning: A new systematic approach to partial delete relaxation, AIJ 221 (2015) 73–114.
[13] P. E. Hart, N. J. Nilsson, B. Raphael, A formal basis for the heuristic determination of minimum cost paths, IEEE Transactions on Systems Science and Cybernetics 4 (1968) 100–107.
[14] E. Karpas, C. Domshlak, Cost-optimal planning with landmarks, in: Proc. IJCAI 2009, 2009, pp. 1728–1733.
[15] E. Keyder, S. Richter, M. Helmert, Sound and complete landmarks for and/or graphs, in: Proc. ECAI 2010, 2010, pp. 335–340.
[16] S. Richter, M. Helmert, M. Westphal, Landmarks revisited, in: Proc. AAAI 2008, 2008, pp. 975–982.
[17] M. Katz, C. Domshlak, Structural patterns heuristics via fork decomposition, in: Proc. ICAPS 2008, 2008, pp. 182–189.
[18] J. E. Doran, D. Michie, Experiments with the graph traverser program, Proceedings of the Royal Society A 294 (1966) 235–259.
[19] Á. Torralba, V. Alcázar, P. Kissmann, S. Edelkamp, Efficient symbolic search for cost-optimal planning, AIJ 242 (2017) 52–79.
[20] D. Speck, F. Geißer, R. Mattmüller, When perfect is not good enough: On the search behaviour of symbolic heuristic search, in: [39], 2020, pp. 263–271.
[21] R. E. Bryant, Graph-based algorithms for Boolean function manipulation, IEEE Transactions on Computers 35 (1986) 677–691.
[22] D. Speck, R. Mattmüller, B. Nebel, Symbolic top-k planning, in: Proc. AAAI 2020, 2020, pp. 9967–9974.
[23] F. Somenzi, CUDD: CU decision diagram package - release 3.0.0, https://github.com/ivmai/cudd, 2015. Accessed: 2023-04-03.
[24] M. Abdulaziz, C. Gretton, M. Norrish, A state-space acyclicity property for exponentially tighter plan length bounds, in: Proc. ICAPS 2017, 2017, pp. 2–10.
[25] C. H. Papadimitriou, M. Yannakakis, A note on succinct representations of graphs, Information and control 71 (1986) 181–185.
[26] Y. Cohen, R. Stern, A. Felner, Solving the longest simple path problem with heuristic search, in: [39], 2020, pp. 75–79.
[27] J. von Tschammer, R. Mattmüller, D. Speck, Loopless top-k planning, in: Proc. ICAPS 2022, 2022, pp. 380–384.
[28] T. Wahl, A. Donaldson, Replication and abstraction: Symmetry in automated formal verification, Symmetry 2 (2010) 799–847.
[29] P. Riddle, J. Douglas, M. Barley, S. Franco, Improving performance by reformulating PDDL into a bagged representation, in: ICAPS 2016 Workshop on Heuristics and Search for Domain-independent Planning, 2016, pp. 28–36.
[30] A. Biere, K. Heljanko, S. Wieringa, AIGER 1.9 And Beyond, Technical Report 11/2, Institute for Formal Models and Verification, Johannes Kepler University, 2011.
[31] R. Brayton, A. Mishchenko, ABC: An academic industrial-strength verification tool, in: Proc. CAV 2010, 2010, pp. 24–40.
[32] A. Biere, K. Fazekas, M. Fleury, M. Heisinger, CaDiCaL, Kissat, Paracooba, Plingeling and Treengeling entering the SAT Competition 2020, in: Proc. SAT Competition 2020, 2020, pp. 51–53.
[33] M. Gebser, R. Kaminski, B. Kaufmann, T. Schaub, Multi-shot ASP solving with clingo, Theory and Practice of Logic Programming 19 (2019) 27–82.
[34] A. Conte, R. Grossi, A. Marino, Large-scale clique cover of real-world networks, Information and Computation 270 (2020) 104464. Special Issue on 26th London Stringology Days & London Algorithmic Workshop.
[35] A. Cimatti, E. Clarke, E. Giunchiglia, F. Giunchiglia, M. Pistore, M. Roveri, R. Sebastiani, A. Tacchella, NuSMV 2: An opensource tool for symbolic model checking, in: Proc. CAV 2002, 2002, pp. 359–364.
[36] S. Minato, Zero-suppressed BDDs for set manipulation in combinatorial problems, in: Proc. DAC 1993, 1993, p. 272–277.
[37] T. Inoue, H. Iwashita, J. Kawahara, S. Minato, Graphillion: software library for very large sets of labeled graphs, International Journal on Software Tools for Technology Transfer 18 (2016) 57–66. doi:10.1007/s10009-014-0352-z.
[38] T. Soh, Y. Okamoto, T. Ito, Core challenge 2022: Solver and graph descriptions, 2022.
[39] Proc. ICAPS 2020, 2020.