Applying Automatic Planning Systems to Airport Ground-Traffic Control — A Feasibility Study

Sebastian Trüg, Jörg Hoffmann, and Bernhard Nebel

Institut für Informatik
Universität Freiburg
79110 Freiburg, Germany
⟨*last name*⟩@informatik.uni-freiburg.de

Abstract. Planning techniques have matured as demonstrated by the performance of automatic planning systems at recent international planning system competitions. Nowadays it seems feasible to apply planning systems to real-world problems. In order to get an idea of what the performance difference between special-purpose techniques and automatic planning techniques is, we applied these techniques to the airport traffic control problem and compared it with a special purpose tool. In addition to a performance assessment, this exercise also resulted in a domain model of the airport traffic control domain, which was used as a benchmark in the 4th International Planning Competition.

1 Introduction

Planning techniques have matured as demonstrated by the performance of automatic planning systems at recent international planning system competitions [1, 12]. Current planning systems are able to generate plans with hundred steps and more compared with ten steps or less ten years ago. Given the inherent complexity of the planning problem, this is a dramatic improvement. The reason for this performance boost is the use of new algorithms and the development of powerful heuristics [2, 3].

In order to assess how feasible it is to apply automatic planning systems to real-world problems, we used these techniques to solve the operational airport ground-traffic control problem [8]. This is an NP-complete problem that can be characterized as a job-shop scheduling problem with blocking. Using special-purpose techniques, it can be solved approximately for a realistic number of airplanes (roughly 50) on realistic airports (such as Frankfurt airport).

So how far do we get with planning techniques for this problem? As our experiments indicate, the *expressiveness* of current planning systems approaches the point where one can formulate the traffic problem as a planning problem. From a *performance* point of view, the results are less encouraging. The planners we evaluate in our experiments are only good enough for small airports with few airplanes. One has to put this result into perspective, however. Special-purpose scheduling techniques are, of course, highly optimized for the particular scheduling problem. Furthermore, the performance penalty to be expected for general-purpose solutions is always quite high and often prohibitive. On the positive

side, the exercise of formalizing the traffic problem has led to a new challenging, real-world benchmark problem for automatic planning systems.[1] Indeed, the domain was used as a benchmark in IPC-4, the 4th International Planning Competition. The results obtained by the competitors in IPC-4 are somewhat more encouraging than our own results. The IPC-4 results were available just a few days before the deadline for the conference version of this paper. We provide a short summary of the results.

The rest of the paper is structured as follows. In Section 2, we give a description of the airport ground-traffic control problem. Section 3 describes then the formalization using PDDL [11], the *de facto* standard for planning systems. In order to allow existing planning systems to be used, we also describe how to compile the PDDL specification into basic STRIPS. Based on that, in Section 4, the results of our experiments, and a summary of the IPC-4 results, are given. Section 5 summarizes and concludes.

2 The Airport Ground-Traffic Control Problem

In a nutshell, the airport ground-traffic control problem consists of coordinating the movements of airplanes on the airport so that they reach their planned destinations (runway or parking position) as fast as possible – whereby collisions shall be, of course, avoided.

The airplanes move on the airport *infrastructure*, which consists of *runways, taxi ways*, and *parking positions*. Airplanes are generally divided into the three 'Wake Vortex Categories': *light, medium*, and *heavy*, which classify them according to their engine exhaust. A moving airplane can either be *in-bound* or *out-bound*. In-bound airplanes are recently landed and are on their way from the runway to a parking position, usually a gate. Out-bound airplanes are ready for departure, meaning they are on their way to the departure runway. Since airplanes are not able to move backwards, they need to be pushed back from the gate on the taxiway where they start up their engines. Some airports provide different park positions that allow an airplane to start its engines directly but to simplify the situation we assume that an airplane always needs to be pushed back.

The ground controller has to communicate to the airplanes which ways they shall take and when to stop. While such guidance can be given purely reactively, it pays off to base decisions on anticipating the future. Otherwise it may happen that airplanes block each other and need more time than necessary to reach their destinations on the airport. The objective is to minimize the overall summed up traveling times of all airplanes.

From a formal point of view, one considers the problem with a time horizon of say one hour and schedules all movements, minimizing the movement times of the planes. Of course, because the situation changes continually (new planes arrive and schedules cannot be executed as planned), continuous rescheduling is

[1] The full encoding of the problem can be found in the technical report version of this paper [14].

necessary. We will consider, however, only the static optimization problem with a given situation on the airport and a time horizon of a fixed time span.

Our domain representation and implementation is based on software by Wolfgang Hatzack, namely on a system called *Astras*: Airport Surface ground TRAffic Simulator. This is a software package that was originally designed to be a training platform for airport controllers. Astras provides a two-dimensional view of the airport, allowing the user to control the airplanes by means of point and click. Astras also includes features for simulating the traffic flow on an airport over the course of a specified time window, as well as an automated controller (named *Acore*) driven by a greedy re-scheduling approach [8]. Our PDDL domain encoding is based on Astras's internal representation of airports. We generated our test instances by software that is integrated with Astras. During an airport simulation, if desired by the user our software exports the current traffic situation in various PDDL encodings.

3 The PDDL Encoding of the Airport Domain

The central object in the PDDL encoding of the airport domain is the airplane that moves over the airport infrastructure. The airport infrastructure is built out of segments. An airplane always occupies one segment and may block several others depending on its type. Our assumption here is that medium and heavy airplanes block the segment behind them whereas light airplanes only block the segment they occupy. Blocked segments cannot be occupied by another airplane. To handle terms like *behind* we need to introduce direction in segments. Since our segments are two-dimensional objects we need exactly two directions which we quite inappropriately call north and south. Every segment has its north end and its south end so it becomes possible to talk about direction in a segment.

To model the airplane movement we need at least two actions. The *move* action describes the normal forward movement of an airplane from one segment to another. The *pushback* action describes backward movement when an airplane is being pushed back from its park position.

We also introduce an airplane state. An airplane can either be moving, be pushed, be parked, or be airborne. We want to make sure that an airplane only moves backwards while being pushed from its park position and only moves forward if not. The parked state is necessary since a parked airplane's engines are off and thus the airplane does not block any segments except the one it occupies unlike when moving. If a plane is airborne, i.e. the plane took off already, then that means that the plane is not relevant to the ground traffic anymore.

The actions *park* and *startup* describe the transitions between the different states. As one may expect the park action makes sure the airplane only blocks the occupied segment while the startup action does the exact opposite. It initially blocks segments depending on the airplane type.

A last action is needed to completely remove the airplane from the airport after takeoff. This action is called *takeoff* and makes sure the airplane does not block or occupy any segment anymore.

In the following we describe our different encodings of the airport domain: a durative and non-durative ADL [13] encoding, a STRIPS [6] encoding where the ADL constructs were compiled out, finally a means to model runway blocking for landing airplanes.

3.1 ADL Encoding

Our domain has four types of objects: *airplane*, *segment*, *direction*, and *airplanetype*.

The airplane type (its Wake Vortex Category) is described with the *has-type* predicate:

```
(has-type ?a - airplane ?t - airplanetype)
```

The airplane state is described with four predicates:

```
(airborne ?a - airplane ?s - segment)
(is-moving ?a - airplane)
(is-pushed ?a - airplane)
(is-parked ?a - airplane ?s - segment)
```

The *is-parked* predicate has a second parameter stating the park position the airplane is parked at. The second parameter of the *airborne* predicate just states from which segment the airplane took off. This is useful in case an airport provides several departure runways and we want to force an airplane to use a specific one. Apart from the airplane state we also need to describe the current position of an airplane and its heading:

```
(at-segment ?a - airplane ?s - segment)
(facing ?a - airplane ?d - direction)
```

We need several predicates to describe the airport structure. All the following predicates are static; they will be set once in the initial state and never be changed.

```
(can-move ?s1 ?s2 - segment ?d - direction)
(can-pushback ?s1 ?s2 - segment ?d - direction)
(move-dir ?s1 ?s2 - segment ?d - direction)
(move-back-dir ?s1 ?s2 - segment ?d - direction)
(is-blocked ?s1 - segment ?t - airplanetype ?s2 - segment ?d - direction)
(is-start-runway ?s - segment ?d - direction)
```

The *can-move* predicate states that an airplane may move from segment *s1* to segment *s2* if its facing direction *d*. The *can-pushback* predicate describes the possible backward movement which is similar to the *can-move* predicate. In our encodings, the *can-move* predicate holds only for pairs of segments that belong to the *standard routes* on the airport – this is common practice in reality (as reroutes are likely to cause trouble or at least confusion), and yields much better planner performance.

The *move-dir* and *move-back-dir* predicates state the airplane's heading after moving from segment *s1* to segment *s2*. That means that in a correct airport domain fact file every *can-move* (*can-pushback*) predicate has its *move-dir* (*move-back-dir*) counterpart.

```
(:action move
 :parameters (?a  - airplane
              ?t  - airplanetype
              ?d1 - direction
              ?s1 - segment
              ?s2 - segment
              ?d2 - direction)
 :precondition (and
                  (has-type ?a ?t)
                  (is-moving ?a)
                  (at-segment ?a ?s1)
                  (facing ?a ?d1)
                  (can-move ?s1 ?s2 ?d1)
                  (move-dir ?s1 ?s2 ?d2)
                  (not (exists (?a1 - airplane) (and (not (= ?a1 ?a))
                                                     (blocked ?s2 ?a1))))
                  (forall (?s - segment) (imply (and (is-blocked ?s ?t ?s2 ?d2)
                                                     (not (= ?s ?s1)))
                                                (not (occupied ?s))
                  ))
               )
 :effect       (and
                  (at-segment ?a ?s2)
                  (not (at-segment ?a ?s1))
                  (occupied ?s2)
                  (not (occupied ?s1))
                  (when (not (= ?d1 ?d2))
                        (and (facing ?a ?d2)
                             (not (facing ?a ?d1))))
                  (blocked ?s2 ?a)
                  (when (not (is-blocked ?s1 ?t ?s2 ?d2))
                        (not (blocked ?s1 ?a)))
                  (forall (?s - segment) (when (is-blocked ?s ?t ?s2 ?d2)
                                               (blocked ?s ?a)
                  ))
                  (forall (?s - segment) (when (and (is-blocked ?s ?t ?s1 ?d1)
                                                    (not (= ?s ?s2))
                                                    (not (is-blocked ?s ?t ?s2 ?d2))
                                               )
                                         (not (blocked ?s ?a))
                  ))
               )
)
```

Fig. 1. ADL *move* action

The *is-blocked* predicate is used to handle all blocking. It says that segment *s1* will be blocked by an airplane of type t at segment *s2* facing into direction d. We will see the use of this predicate in the move action's effect in detail.

The last predicate, *is-start-runway*, states that an airplane at segment s facing direction d is allowed to takeoff. This is an essential predicate as we cannot allow an airplane to takeoff wherever it (or better: the planner) wants.

Finally, we need two predicates to describe the current situation on the airport regarding blocked and occupied segments:

```
(occupied ?s - segment)
(blocked ?s - segment ?a - airplane)
```

3.1.1 Non-durative actions Now we will look at the actions used in the non-durative ADL domain in detail.

The most important one is clearly the *move* action (see Figure 1). The parameters have the following meaning: an airplane a of type t facing direction *d1* on segment *s1* moves to segment *s2* and then faces direction *d2*. The first four predicates in the precondition make sure the airplane is in a proper state, meaning it is located in the right segment facing the right direction. Predicates five and six represent the airport structure as described above. The last two formulas are the really interesting ones as they deal with segment blocking. The first one makes sure that no other airplane blocks the segment our airplane is moving to. Here we use the full power of ADL to check for every airplane if it is blocking segment *s2*. Our moving airplane cannot block itself so we exclude it from the check. The second formula is used to make sure that none of the segments our airplane will block after the movement is occupied by another airplane. To achieve that we use the *is-blocked* predicate. Its instances give us the segments that will be blocked once our airplane has moved to segment *s2*. If we find such a segment we make sure its not occupied since a segment may not be occupied and blocked by different airplanes at the same time. We skip segment *s1* since it is always occupied by our airplane.

If all of these conditions apply, the *move* action may be executed. Most of the effects should be self-explanatory: updating of the occupied segment, changing the heading if necessary, and blocking the occupied segment. Again, the blocking is the part where full ADL is needed. We iterate over all segments to find those that are blocked from our airplane after the movement and those that were blocked before the movement. The latter ones need to be unblocked with exception of those blocked after the movement. Again, the *is-blocked* predicate is the central tool.

The *pushback* action differs only little from the *move* action. That is why we will skip a detailed description here and take a look at one of the more interesting actions used for performing the transitions of the airplane state.

The *startup* action (see Figure 2) represents the process of starting the airplane's engines after it has been pushed back from its park position and thus is clearly only needed for out-bound airplanes. That is why the preconditions contain the *is-pushing* predicate. The *startup* action represents the process of the airplane starting its engines. That means it begins blocking segments. So we need the exact same check for occupied segments as in the *move* action and also the blocking formula in the action's effect. Apart from that we only update the airplane state to *is-moving*.

The *park* action does the opposite of the *startup* action by unblocking all segments except the occupied one.

The takeoff action makes sure the airplane is completely removed from the airport meaning it does not block or occupy any segments anymore.

Definition 1. *A correct state of an airplane a is defined by the following facts:*

1. *a is at exactly one segment or is airborne.*

```
(:action startup
  :parameters (?a - airplane ?t - airplanetype ?s - segment ?d - direction)
  :precondition (and
                  (is-pushing ?a)
                  (has-type ?a ?t)
                  (at-segment ?a ?s)
                  (facing ?a ?d)
                  (forall (?s1 - segment)
                          (imply (and (is-blocked ?s1 ?t ?s ?d)
                                      (not (= ?s ?s1)))
                                 (not (occupied ?s1))
                  ))
                )
  :effect       (and
                  (not (is-pushing ?a))
                  (is-moving ?a)
                  (forall (?s1 - segment)
                          (when (is-blocked ?s1 ?t ?s ?d)
                                (blocked ?s1 ?a)
                  ))
                )
)
```

Fig. 2. ADL *startup* action

2. *a occupies exactly the segment it is at or is airborne.*
3. *If a is airborne it neither occupies nor blocks any segments.*
4. *a is facing in exactly one direction or is airborne.*
5. *If a is moving or being pushed it only blocks the segments determined by the is-blocked predicate and the one it occupies.*
6. *If a is parked it only blocks the segment it occupies.*
7. *a never blocks a segment occupied by another airplane.*
8. *a never occupies a segment blocked by another airplane.*

Proposition 1. *If all airplanes had a correct state before executing an action of the airport domain they also have correct states afterwards.*

The (straightforward) proof to Proposition 1 can be found in our TR [14].

3.1.2 Durative actions To obtain a domain with action durations, we enriched the above action encodings with the obvious "at start", "at end", and "over all" flags for the preconditions and effects. The duration of *move* and *pushback* actions is calculated as a function of segment length (airplane velocity is assumed as a fixed number). The duration of *startup* actions is calculated as a function of the number of engines. The *park* and the *takeoff* action's durations are both set to fixed values.

3.2 STRIPS

Most current PDDL planning systems can not handle full ADL, especially durative ADL. To work around this problem and make the airport domain accessible

```
(:action pushback_seg_pp_0_60_seg_ppdoor_0_40_south_south_medium
 :parameters    (?a - airplane)
 :precondition  (and
                 (has-type ?a medium)
                 (is-pushing ?a)
                 (facing ?a south)
                 (at-segment ?a seg_pp_0_60)
                 (not_occupied seg_ppdoor_0_40)
                 (not_blocked seg_ppdoor_0_40 airplane_CFBEG)
                 (not_blocked seg_ppdoor_0_40 airplane_DAEWH)
                )
 :effect        (and
                 (not (occupied seg_pp_0_60))
                 (not_occupied seg_pp_0_60)
                 (not (blocked seg_pp_0_60 ?a))
                 (not_blocked seg_pp_0_60 ?a)
                 (not (at-segment ?a seg_pp_0_60))
                 (occupied seg_ppdoor_0_40)
                 (not (not_occupied seg_ppdoor_0_40))
                 (blocked seg_ppdoor_0_40 ?a)
                 (not (not_blocked seg_ppdoor_0_40 ?a))
                 (at-segment ?a seg_ppdoor_0_40)
                )
)
```

Fig. 3. A STRIPS pushback action: pushing a medium airplane from seg_pp_0_60 to seg_ppdoor_0_40 heading south

to most planning systems we need to remove the ADL constructs to create a STRIPS only version.

Due to the lack of quantified variables there is no way to determine which segments need to be unblocked and unoccupied when moving. In our context this difficulty can be tackled by pre-instantiating. We create one *move* (and *pushback*) action for every pair of segments, for every airplane type, and for every possible pair of directions associated to moving between the segments. We can then do the required state updates without the de-tour to conditions over the *is-blocked* predicate. (More formally, the conditional effects disappear because their conditions are all static.) We end up with a single non-instantiated parameter for the *move* and *pushback* actions: the particular airplane that is moved.

The *blocked* predicate depends on an airplane. In the ADL version a quantified precondition iterates over all airplanes to check if a segment is blocked. In STRIPS, instead we use a separate check for every airplane on the airport.

STRIPS does not allow negated preconditions. We need negated preconditions to test for unblocked and unoccupied segments, so we emulate their behavior using a standard translation approach. We introduce *not_blocked* and *not_occupied* predicates, and make sure that these always exhibit the intended behavior (i.e. they are assigned the inverse values initially, and every action effect is updated to affect them inversely).

Figure 3 shows an example of a non-durative STRIPS *pushback* action.

Now that we moved almost everything from the fact into the domain declaration it is obvious that we need a separate domain definition for each airport

situation. The fact file merely defines the airplanes' starting state and the goal. All airport layout is implicitly defined by the move actions. Thus we can drop the *is-blocked*, *can-move*, *move-dir*, *can-pushback*, and *move-back-dir* predicates. The STRIPS domain version can be enriched with durations exactly like the ADL version.

3.3 Time Windows

The controllers of most large airports in the world only control their fore-field but not the airspace above the airport. In particular, they can not change the landing times of airplanes, and independently of what they decide to do the respective runway will be blocked when a plane lands.

We have to model something like *segment s is blocked from time x to time y.* Clearly this can only be achieved with durative PDDL. The language for IPC-4, PDDL 2.2 [5], introduced *Timed Initial Literals*, which provide a very simple solution to the time-window problem. They allow the specification of a literal together with a time stamp at which the literal will become true. For example:

```
(at 119 (blocked seg_27 dummy_landing_airplane))
```

The above statement in the ":init" section of the fact file will make sure that segment seg_27 gets blocked at time 119. To unblock the segment after the landing we use a similar statement.

```
(at 119 (not (blocked seg_27 dummy_landing_airplane)))
```

Since the blocking predicate's second parameter is an airplane and we only simulate the landings but not the airplanes themselves, we use a dummy airplane for all landing actions. The dummy airplane is used nowhere else. (It does not have a correct state in the sense of Definition 1, but obviously this does not affect the correctness of our overall encoding.)

3.4 Optimization Criterion

We were not able to model the real optimization criterion of airport ground traffic control. The standard criterion in PDDL is to minimize the execution time, the *makespan*, of the plan. In our encoding of the domain this comes down to minimizing the arrival time of the last airplane. But the real objective is, as said above, to minimize the overall *summed up* travel time of all airplanes. There appears to be no good way of modeling this criterion in current PDDL. The difficulty lies in accessing the *waiting* times of the planes, i.e. the times at which they stay on a segment waiting for some other plane to pass. If one introduces an explicit waiting action then one must discretize time, in order to be able to tell the planner how long the plane is supposed to wait. For PDDL2.2, the introduction of a special fluent "current-time" was considered, returning the time point of its evaluation in the plan execution. Using such a "look on the

clock", one could make each plane record its arrival time, and thus formulate the true optimization criterion in Airport. The IPC-4 organizing committee decided against the introduction of a "current-time" variable as it seems to be problematic from an algorithmic point of view (it implies a commitment to precise time points at planning time), and not relevant anywhere else but in the airport domain.

4 Results

To evaluate the performance of state-of-the-art planning systems in the Airport domain, we created five scaling example airports that we named "Minimal", "Mintoy", "Toy", "Half-MUC", and "MUC". The smallest of these airports is the smallest possible airport Astras can handle. The two largest airports correspond to one half of Munich Airport, MUC, respectively to the full MUC airport. Figure 4 shows sketches of the "Minimal" airport, and of the "MUC" airport.

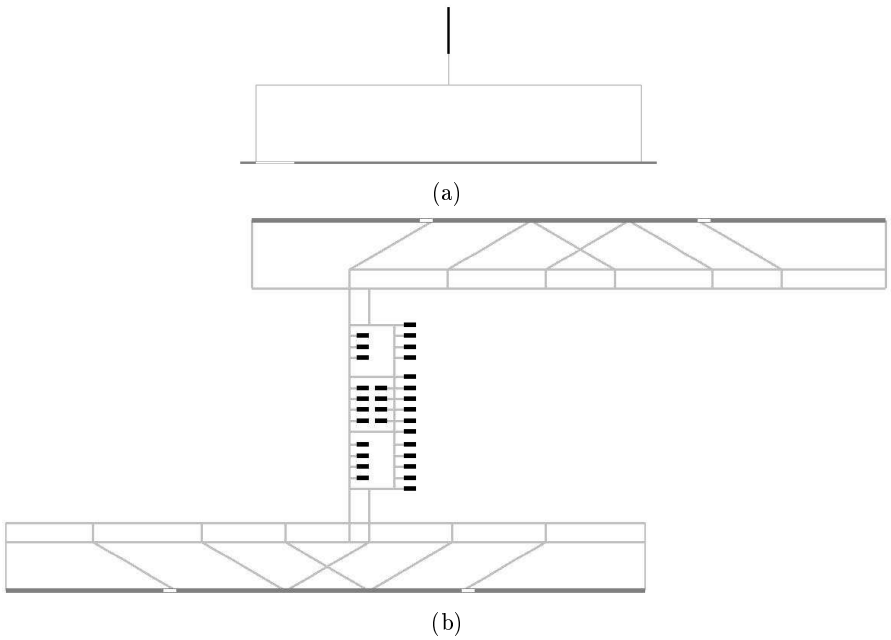

Fig. 4. Sketches of the "Minimal" (a) and "MUC" (b) airports. Park position segments are marked in black, while the segments airplanes can takeoff from are marked in white.

For each airport, we created a number of test examples that scaled by the number of airplanes to be moved, and by the number of time windows (if any). As indicated earlier in the paper, the test examples were generated by running an Astras simulation of the respective airport, then selecting various traffic situations during the simulation and putting them out in our different versions/encodings of the domain.

In our experiments, we ran the planners FF [9], IPP [10], MIPS [4], and LPG [7]. These planners are available online (programmed by one of the authors, in

the cases of FF and IPP), and are suitable to represent the state-of-the-art in sub-optimal and (step-)optimal fully automated planning during the last 5 or 6 years, judging from the results of the international planning competitions. FF was the winner of the 2000 competition, and is able to handle non-durational ADL representations. MIPS was awarded a 2nd place at both the 2000 and 2002 competitions, and handles durational STRIPS as well as timed initial literals. LPG was the winner of the 2002 competition, and handles durational STRIPS representations. IPP is the only optimal planner in our collection; it is based on the Graphplan [2] approach, handles non-durational ADL representations, and finds plans with a provably smallest number of parallel time steps.[2] IPP won the ADL track of the 1998 competition.

The planners were run under Linux on a 6 GB RAM computer with 4 Xeon 3.06 GHz CPUs. We report total runtimes as put out by the systems. Whenever a planning system ran out of memory on an example, the respective table entry shows a "—". We used a runtime cutoff of 10 minutes, and examples not solved within that time are marked by ">>". We first give our results for the non-durational Airport domain, then for the durational domain without time windows, finally for the full durational domain with time windows. We then provide a brief summary of the results obtained for the airport domain in the 2004 competition.

4.1 Non-durational results

In the non-durational domain, two of our tested systems – namely FF and IPP – could handle the original non-compiled ADL domain. We ran FF and IPP on both the ADL and STRIPS encodings of the domain, and we ran LPG and MIPS on the STRIPS encoding only. Table 1 shows the results. We only report runtimes, not plan quality. From a real-life perspective, talking about plan quality – summed up travel time – does not make much sense in a non-durational setting where actions don't take any time.

Instances with a number of airplanes not solved by any of our planners are not shown in the table. We observe the following. First, in the ADL encoding no planner is able to scale to the MUC airports. The failure is due to a memory explosion in the pre-processing routines of both FF and IPP, which ground out all actions.[3] Strangely, in those ADL instances that are feasible FF is a lot more efficient than in the corresponding STRIPS instances. We investigated this phenomenon but could not find an explanation for it. It appears to be due to the internal ordering of the action instances. Our second observation is that all planners scale reasonably well in the smaller artificial airports, but run into

[2] Of course, finding step-optimal movement sequences is a long way away from the real optimization criterion in controlling airport ground traffic. Nevertheless, we find it interesting to see how far an optimal planner scales in the domain.

[3] While grounding out the actions comes very close to what we did in the STRIPS compilation, FF and IPP do it in a general way and fail during the creation of the necessary thousands of 1st order formulas (i.e., pointer tree structures).

		ADL		STRIPS			
Airport	Nr. planes	FF	IPP	FF	IPP	LPG	MIPS
Minimal	1	0.03	0.02	0.01	0.01	0.02	0.09
Minimal	2	0.05	0.04	0.01	0.02	0.09	0.11
Mintoy	1	0.39	0.26	0.02	0.05	0.08	0.25
Mintoy	2	0.59	0.42	0.01	0.21	0.09	0.30
Mintoy	3	0.85	0.86	0.02	0.57	0.16	0.67
Mintoy	4	1.09	5.56	0.76	9.54	17.45	15.12
Toy	1	0.55	0.35	0.02	0.08	0.18	0.30
Toy	2	0.79	0.55	0.03	0.24	0.48	0.36
Toy	3	1.38	1.05	0.06	0.38	0.62	0.68
Toy	4	1.50	3.52	3.04	4.63	0.82	37.12
Toy	5	1.80	239.19	19.30	190.76	9.84	38.64
Toy	6	2.52	>>	5.57	>>	22.17	164.31
Toy	7	6.60	>>	>>	>>	—	>>
Half-MUC	2	—	—	0.25	8.96	>>	6.25
Half-MUC	3	—	—	0.37	18.73	—	13.51
Half-MUC	4	—	—	0.88	49.92	—	—
Half-MUC	5	—	—	1.34	61.14	—	—
Half-MUC	6	—	—	1.82	87.76	—	—
Half-MUC	7	—	—	3.59	131.13	—	—
Half-MUC	8	—	—	4.52	173.68	—	—
Half-MUC	9	—	—	5.60	240.01	—	—
Half-MUC	10	—	—	8.52	—	—	—
MUC	2	—	—	0.40	32.81	—	—
MUC	3	—	—	0.62	63.93	—	—
MUC	4	—	—	0.83	109.00	—	—
MUC	5	—	—	335.38	289.18	—	—

Table 1. Runtime results (in seconds) for the non-durational encodings of the Airport domain.

trouble on the real-life sized MUC airports. On the negative side, LPG and MIPS solve hardly any of the MUC instances – it is unclear if this is due to the size of the search spaces, or to implementational difficulties with the pre-compiled STRIPS encoding. On the positive side, FF and IPP can solve Half-MUC instances even with many airplanes, and scale up to 5 planes on the real-life MUC airport. It is interesting to note that the optimal planner IPP is competitive with the sub-optimal planner FF, and even outperforms LPG and MIPS – a very unusual phenomenon in today's planning landscape. Apparently, the heuristic function encoded in the planning graph is of high quality (comparable to the quality of heuristics based on ignoring the delete lists) in the airport domain. Altogether, it seems that planners are not too far away from real-life performance in the non-durational setting of this domain.

4.2 Durational results

From the real-life perspective, of course the durational version of the Airport domain is the more interesting one, particularly the version including time windows for the landing airplanes. Of our tested systems, only LPG and MIPS can handle durations, and only MIPS can handle the timed initial literals necessary to encode the time windows. Both LPG and MIPS handle only STRIPS representations so we could not run the ADL encoding. See the results for the STRIPS

encodings in Table 2. We only report runtime. The found plans are all optimal with respect to the summed up overall travel time, see the discussion below.

		No Time Windows		Time Windows
Airport	Nr. planes	LPG	MIPS	MIPS
Minimal	1	0.03	0.15	0.18
Minimal	2	0.04	0.26	0.30
Mintoy	1	0.44	0.42	0.55
Mintoy	2	1.63	2.06	2.71
Mintoy	3	0.17	13.56	13.79
Mintoy	4	>>	74.88	19.45
Toy	1	0.10	0.50	1.16
Toy	2	0.37	2.41	13.77
Toy	3	0.24	13.65	83.59
Toy	4	2.76	101.51	102.68
Toy	5	3.46	>>	>>
Toy	6	5.48	>>	>>
Toy	7	>>	>>	>>
Half-MUC	2	>>	94.52	157.76
Half-MUC	3	—	405.19	497.93

Table 2. Runtime results (in seconds) for the durational STRIPS encoding of the Airport domain, with and without time windows.

As above, instances solved by none of the planners are not shown in the table. The obvious observation regarding runtime is that the durational domain is much harder for our planners than the non-durational domain – though as above it is unclear if LPG's and MIPS's inefficiency in the MUC airports is due to search complexity, or to the pre-compiled STRIPS encoding. LPG is generally more efficient than MIPS in the smaller airports, but fails completely in the larger Half-MUC airport.

Regarding plan quality, as said above already LPG and MIPS find the optimal plans in all cases they can handle, i.e. they return the plans with the smallest possible summed up overall travel time. We checked that by hand. While this is a good result, one should keep two things in mind. First, LPG and MIPS do not know about the real optimization criterion so it is largely a matter of chance if or if not the plan they find is optimal with respect to that criterion. Second, the instances shown here – those cases that LPG and MIPS can handle – are very simple. With just a few airplanes, there is not much potential for (possibly) harmful interactions between the intended travel routes of these planes. In the above examples, often it is the case that there is just one non-redundant solution (a solution that does not leave planes standing around waiting without reason), and that this solution is the optimal one. Specifically this is the case in the two Half-MUC instances solved by MIPS.

We also wanted to run Acore, the (sub-optimal) scheduler integrated with Astras, on the above instances (i.e. in the respective traffic situations during the simulation with Astras), and compare the results with those of our planners. This turned out to not be feasible. In the small airports, there are a lot of parking conflicts, i.e. cases where an in-bound airplane is headed for a parking position

that is occupied by an out-bound airplane. Such situations do rarely occur in reality (in fact, the flight schedules try to avoid these situations), and Acore can't handle them. In the larger MUC airports, on the other hand, our planners could not solve many instances. In the two Half-MUC instances solved by MIPS, Acore finds the trivially optimal solutions just like MIPS does. Generally, concerning runtime Acore is vastly superior to our planners. Acore can solve instances with 50 planes and more on Frankfurt airport, which is far beyond the scalability of the tested planning systems.

4.3 IPC-4 results

As said in the introduction, the IPC-4 results became available just a few days before the deadline for the version of this paper to be included in the conference proceedings. We were therefore, and for space reasons, unable to include a detailed discussion of these results, but we consider it interesting to provide at least a brief account of what happened. The IPC-4 results were obtained on the same machine that we used in the experiments above.

Some progress was made in the non-durational performance. Three planners ("Fast Downward", "SGPlan", and the new version of LPG) were able to solve Half-MUC with up to 12 planes within 100 seconds. Fast Downward even solved a MUC example with 15 planes within 200 seconds. The progress made on the durational performance is yet more impressive: even in the presence of time windows, the performance of LPG and "SGPlan" was very similar to that in the non-durational domain, easily (within 100 secs) solving Half-MUC examples with up to 11 planes, and solving MUC examples with up to 5 planes within 30 minutes. For optimal planners, not so much progress could be observed. The most efficient optimal planner in the non-durational domain, "SATPLAN04", was roughly as efficient as IPP in our own experiments. There were only three optimal planners that could handle durations, and only a singe Half-MUC instance (with two planes) got solved by them within 30 minutes.

5 Conclusion

The results show that today's PDDL planning systems are not quite yet powerful enough to handle the airport domain when it comes to real-life problems – for that, the planners would have to be able to, like Acore, generate good solutions to large airports (like Frankfurt) with many airplanes (roughly 50) in a few seconds. Nonetheless, the results, especially those obtained for the durational domain by the sub-optimal planners in IPC-4, are very encouraging. They definetely show that today's state-of-the-art planners are *a lot* closer to real-life applicability than they were some years ago. They even suggest that real-life applicability, at least in this particular domain, has come within close reach.

The core problem in controlling the ground traffic on an airport is to resolve the *conflicts* that arise when two planes need to cross the same airport segment [8]. In our PDDL encoding, this core problem is hidden deep in the domain

semantics, and it seems likely that the automated planners spend most of their runtime unawares of the core difficulties. One can try to overcome this by not encoding in PDDL the physical airport, but only the conflicts and their possible solutions. Seeing if and how this is possible, ideally in connection with the real optimization criterion, is an important topic for future work.

References

1. Fahiem Bacchus. The AIPS'00 planning competition. *The AI Magazine*, 22(3):47–56, 2001.
2. Avrim L. Blum and Merrick L. Furst. Fast planning through planning graph analysis. *Artificial Intelligence*, 90(1-2):279–298, 1997.
3. Blai Bonet and Héctor Geffner. Planning as heuristic search. *Artificial Intelligence*, 129(1-2):5–33, 2001.
4. Stefan Edelkamp. Taming numbers and durations in the model checking integrated planning system. *Journal of Artificial Intelligence Research*, 20(195-238), 2003.
5. Stefan Edelkamp and Jörg Hoffmann. PDDL2.2: The language for the classical part of the 4th international planning competition. Technical Report 195, Albert-Ludwigs-Universität, Institut für Informatik, Freiburg, Germany, 2004.
6. Richard E. Fikes and Nils Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208, 1971.
7. A. Gerevini, A. Saetti, and I. Serina. Planning through stochastic local search and temporal action graphs in LPG. *Journal of Artificial Intelligence Research*, 20:239–290, 2003.
8. Wolfgang Hatzack and Bernhard Nebel. Solving the operational traffic control problem. In A. Cesta, editor, *Recent Advances in AI Planning. 5th European Conference on Planning (ECP'01)*, Toledo, Spain, September 2001. Springer-Verlag.
9. Jörg Hoffmann and Bernhard Nebel. The FF planning system: Fast plan generation through heuristic search. *Journal of Artificial Intelligence Research*, 14:253–302, 2001.
10. Jana Koehler, Bernhard Nebel, Jörg Hoffmann, and Yannis Dimopoulos. Extending planning graphs to an ADL subset. In S. Steel and R. Alami, editors, *Recent Advances in AI Planning. 4th European Conference on Planning (ECP'97)*, volume 1348 of *Lecture Notes in Artificial Intelligence*, pages 273–285, Toulouse, France, September 1997. Springer-Verlag.
11. Drew McDermott et al. *The PDDL Planning Domain Definition Language*. The AIPS-98 Planning Competition Comitee, 1998.
12. M.Fox and D.Long. The AIPS-2002 international planning competition. In *Proceedings of the 6th International Conference on Artificial Intelligence Planning Systems (AIPS-02)*. AAAI Press, Menlo Park, 2002.
13. Edwin P. D. Pednault. ADL: Exploring the middle ground between STRIPS and the situation calculus. In R. Brachman, H. J. Levesque, and R. Reiter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 1st International Conference (KR-89)*, pages 324–331, Toronto, ON, May 1989. Morgan Kaufmann.
14. Sebastian Trüg, Jörg Hoffmann, and Bernhard Nebel. Applying automatic planners to airport ground-traffic control. Technical Report 199, Albert-Ludwigs-Universität, Institut für Informatik, Freiburg, Germany, 2004.