

Towards Mixed Planning and Scheduling

Roman Barták¹

Abstract. The conventional wisdom and practice is that planning and scheduling tasks are solved separately using different methods and approaches. However, recent development in industrial planning and scheduling demands for mixing both tasks to allow modelling of wider class of problems.

The purpose of this paper is to present a framework for mixing planning and scheduling tasks within single system. We analyse traditional views of planning and scheduling and we highlight the drawbacks of separating both tasks when applied to modelling complex process environments. We give some real-life examples where the mixed approach helps to model the problems and we propose a generic framework for such mixture. We also argue for using constraint programming as underlying solving technology and, finally, we describe some constraint models based on the proposed framework.

Although, we concentrate on planning and scheduling in complex process environments we believe that the results contribute to both planning and scheduling communities in general.

1 INTRODUCTION

It is a common practice that planning and scheduling tasks are solved separately using different methods. The *planning task* deals with finding plans to achieve some goal, i.e., finding a sequence of activities that will transfer the initial world into one in which the goal description is true. Moreover, the possible sequences of actions must respect the limitations of the world. Planning has been studied in Artificial Intelligence (AI) for years and the methods developed there, like the STRIPS representation [6] and the Graphplan planning algorithm [3], are the core of many planning systems. Opposite to planning, the *scheduling task* deals with the exact allocation of activities to available resources over time respecting precedence, duration, capacity, and incompatibility constraints [4,5]. Traditional scheduling methods developed in Operations Research (OR) are now being absorbed by Constraint Programming (CP) technology that provides better declarative modelling capabilities [13].

Recent development shows that such strict distribution between planning and scheduling is not desirable in some problems and that using planning techniques in scheduling and vice versa may contribute both to richer modelling capabilities and to higher efficiency of the systems. Convergence may be observed on both sides of the border. The planning community tackles problems typical for scheduling like planning under resource constraints [8] and new algorithms were developed here using the constraint programming technology [7] that is more tributary to scheduling. Nevertheless, to be fair we should also mention that there exists reverse movement [11] asking for removing resources scheduling from planning algorithms. Scheduling community is a bit self-contained too but there exist feelings about the necessity to include some planning capabilities to schedulers [2]. Also, the developers of constraint satisfaction algorithms listen to the requirements of dynamic planning tasks more [9].

Moreover, in industrial applications the border between planning and scheduling is fuzzier than in academics and the discrimination criterion is shifted to a different level. Both industrial planning and scheduling deal with generation and allocation of activities; the difference between planning and scheduling is shifted here to the resolution level of the resulting plan or schedule. While industrial planning deals with the task of finding “rough” plans for longer period of time and uses department as the basic resource, the scheduling prepares detail schedules with higher time resolution for individual machines, workers etc.

The similar character of industrial planning and scheduling brings the idea of using single approach that can be applied to both areas. In the paper we propose such a framework for close co-operation between planning and scheduling modules. The basic idea is not very complicated: the planner introduces new activities to the system and these activities are immediately allocated to resources by the scheduler. The new contribution here is the instant allocation of activities after their introduction. Note that we do not mean fixing the activity completely but restricting the possible “positions” of the activity by means of constraint propagation. This prunes the search space of the planner and allows us to detect the clashes soon. Also, the scheduler may ask the planner to introduce new activities if the situation (the current partial schedule) requires them. Such behaviour of the scheduling engine was motivated by existence of real-life problems going beyond the horizon of conventional static scheduling and requiring the dynamics of planning.

The rest of the paper describes the proposed framework in detail. In Section 2 we specify the problem area and we outline the main difficulties. In Section 3 we overview the conventional definition of planning and scheduling tasks and we give examples where the strict separation of both tasks is too restrictive. Section 4 is dedicated to description of the basic structure for mixed planning and scheduling system. In Section 5, we look at the engine behind our framework and we describe how to express our ideas in terms of constraint programming technology. We conclude with summary of the paper and with description of early experience with the implementation.

2 PROBLEM AREA

The problem area that we deal with can be characterised as a complex process environment where a lot of complicated real-life constraints bind the problem variables. Typical examples of such environments can be found in plastic, petrochemical, chemical, pharmaceutical or food industries. The task is to schedule most profitable production for fixed period of time.

The problem domain is described as a heterogeneous environment with *several resources* interfering with each other. Currently we are working with producers, movers and stores, later other resources like workers and tools will be added. The task is to generate (to plan) activities necessary to satisfy custom orders (and other marketing requirements) and to allocate (to schedule) the activities to resources over time.

¹ Charles University, Faculty of Mathematics and Physics, Department of Theoretical Computer Science, Malostranské nám. 2/25, 118 00, Praha, Czech Republic, email: bartak@kti.mff.cuni.cz
Supported by the Grant Agency of the Czech Republic under the contract no. 201/99/D057.

There exist alternative resources for processing the activity and some resources can handle several activities at a time (this is called batch processing). In case of batch processing, compatibility and capacity constraints restricting which products and in what quantities can be processed, i.e., produced, moved or stored together, must be considered. Also the order of activities processed by the resource is not arbitrary but the currently processed activity influences what activities can follow. Consequently, we must follow the *transition patterns* and assume the *set-up times* between the activities as well. The processing time is usually variable and there is defined a *working time* when the activities can be processed in resources.

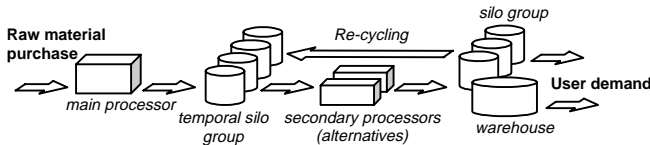


Figure 1. A complex-process environment

Alternative processing routes, *alternative production formulas* and *alternative raw materials* are other typical features of above mentioned industry areas. In addition to the core products it is possible to produce the *by-products*, typically during set-ups. The by-products should be used as a raw material in further production and there is a push to use them this way because they will fill-up the available storing capacity otherwise. Consequently we must schedule processing of by-products. During production of the core product some *co-products* may appear. The co-products can be used to satisfy other orders, they can be sold as an alternative to the ordered item or they can be processed further as a raw material. Again, processing of co-products must be scheduled as well because of limited capacity of warehouses where all the products are stored. Last but not least there is a possibility of *cycling*, i.e., processing the item for several times for example to change features of the item or just to clean up the store, and *re-cycling*, i.e., using of by-products and co-products as a raw material.

Typically, the production in complex process environments is not driven by the custom orders only but it is necessary to schedule the production for store according to the factory patterns and the forecast. It means that the scheduler should be able to introduce new activities during scheduling to “fill the gaps”.

3 TRADITIONAL VIEW OF PLANNING AND SCHEDULING

Let’s look at the traditional definition of planning and scheduling tasks first.

Planning. The traditional AI planning tackles the problem of finding plans to achieve some goal, i.e., finding a sequence of activities that will transfer the initial world into one in which the goal description is true. It means that a description of the initial world, the (partial) specification of the desired world and the list of available activities make the input of the planner. A solution is a sequence of activities that leads from the initial world description to the goal world description and it is called a plan.

Conventional AI planning techniques use highly specific representation and algorithms but there is a pressure to use more general search frameworks like CP [9]. The advantage of such general framework is wider applicability and availability of ready-to-use methods. The specific features of a particular problem are

then reflected at the modelling level only and not in the underlying search algorithms.

Scheduling. The traditional scheduling task deals with the exact allocation of activities to resources (or resources to activities) over time respecting precedence, duration, capacity, and incompatibility constraints [4,5]. The set of activities, the list of resources and the specification of the constraints make the input to the scheduler. The output of the scheduler consists of the exact allocation of the activities to the resources over time.

Scheduling tasks are usually solved using techniques from OR and CP. Both frameworks expect the task to be specified fully in advance, i.e. all the problem variables and constraints must be known beforehand. Recently, new problem areas like complex-process environments use a partial specification of the problem that requires adding new variables and constraints during scheduling.

As mentioned above, the border between planning and scheduling is fuzzier in industrial life so it is not surprising that there exist systems providing both planning and scheduling functionally. Nevertheless, in such Advanced Planning and Scheduling (APS) systems, the modules implementing the tasks are still separated following the above described traditional view. This decomposition seems natural because both tasks deal with a bit different problems (generation vs. allocation of activities) and different methods from different areas are used to solve the tasks. Also, the interface between modules implementing planning and scheduling is well defined: the planner generates a list of activities first and, then, push these activities to the scheduler that allocates them to available resources. Figure 2 describes the flowchart of such system with strictly separated planning and scheduling modules.

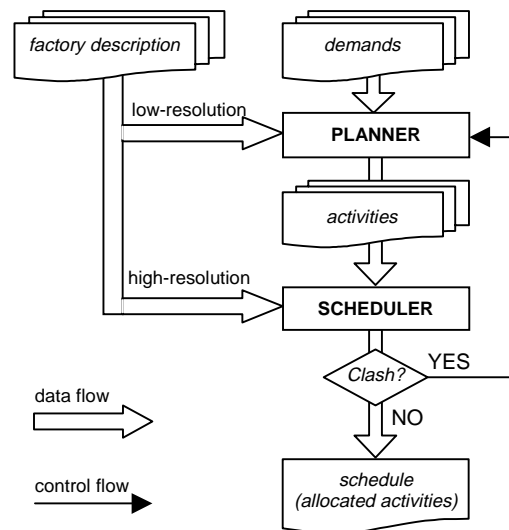


Figure 2. Separate planning and scheduling

3.1 Difficulties of separate planning and scheduling

The structure of APS system with separate planning and scheduling modules is satisfactory when the control and data flow through the system is linear, i.e. when we do not backtrack from the scheduler to the planner. In general such backtrack is not forbidden but, of course, it is not desirable because it decreases efficiency and complicates the interface between the planner and the scheduler.

We see two different reasons for backtrack from the scheduler to the planner. First, there is a possibility of clash in the plan that may be caused by choosing a bad alternative during planning which prevents the scheduler to allocate the activities to available resources. The second reason for backtrack is a low-profitable schedule that may occur when the plan does not utilise the resources fully. Because the scheduler is not able to repair the plan alone (there is no activity “re-generator” in the scheduler) it must ask the planner for help. Moreover, the scheduler should inform the planner about the reason of backtrack and the planner must be able to repair the plan accordingly.

To reduce the number of backtracks or to eliminate them at all we may use a more informed planner that generates conflict-free enough-profitable plans. However, this requires usage of higher-resolution data and taking into account the typical scheduling constraints [8]. Actually, such informed planner absorbs the scheduler or, at least, it solves part of the scheduling task.

Another possibility to avoid backtracking is to postpone planning decisions until enough information is available. Decision postponement is a popular approach in the planning community; its active version using constraint programming was implemented in the Descartes planning system [7]. However, in current systems such decision postponement keeps within the planning context. What we propose is to extend the decision postponement as long as to the scheduling stage.

Before we describe the proposed mixed planning and scheduling framework let's look at other problems from complex-process environments that motivate mixing planning and scheduling. All these real-life problems have one common feature: the appearance of some activity depends on allocation of other activities to resources. Naturally, such problems cannot be solved by more informed planner and they require closer co-operation between planning (introduction of activities) and scheduling (allocation of activities).

Set-ups and transitions. Scheduling set-ups and transitions is a crucial problem in complex-process environments because of their considerable cost and duration. It is impossible to predict appearance of set-ups/transitions before the activities are allocated to the resource so we need to introduce special set-up/transition activities during scheduling [10]. This is necessary especially if by-products result from such activities.

Consumption of by-products. Many current scheduling systems do not care about by-products but in complex-process environments this is not desirable. First, we may use by-products in further production as raw material; second, by-products may fill up the stores for final products. Unfortunately, until we know that the by-product appears (and this is during set-ups or transitions typically) we cannot introduce activities for processing the by-product.

Production for store. In some plant configurations, it is cheaper to continue in production rather than stopping the machine if all the ordered products were finished. Decision about such non-ordered production could be done at the planning level; i.e. the planner generates activities for non-ordered production. However, if the resulting plan is too ambitious then it may cause clashes in the schedule which results in backtrack to the planner. Therefore, it seems more appropriate to delegate the decision to the scheduler which may “fill the gaps” in the schedule by activities producing for store.

4 MIXED PLANNING AND SCHEDULING

In previous paragraphs and sections we argued for mixing planning and scheduling components to solve some problems that conventional separate planning and scheduling systems cannot tackle. We also sketched the basic ideas behind such mixed framework, in particular introduction of activities during scheduling and postponing planning decisions till scheduling. Actually, both these methods describe the same thing from two different views: postponement of planning decisions is realised (partially) by introduction of activities during scheduling.

We shall describe the structure of the mixed planning and scheduling system using the notions of production scheduling that is our problem area. However, we believe that the same structure is applicable to other planning and scheduling areas like transportation problems as well.

The system consists of two components, activity generator (former planner) and activity allocator (former scheduler). You may see the relation between them in Figure 3.

The *activity generator* is responsible for introduction of new activities to the system. It may use the information about allocation of already posted activities when deciding about new activity. Also, if the activity generator “is not sure” about the next activity it may introduce a slot for activity and the real activity will be filled in this slot later by the activity allocator. Finally, the activity generator is responsible for posting constraints among activities, e.g., it ensures that production activity is connected to supplying and consuming activities etc. The mechanism of activity slots and the possibility to introduce constraints describing relations among activities realises the decision postponement.

The *activity allocator* deals with the allocation of activities to resources over time. In general, activity allocator decides about values of activity attributes like start and completion time, the resource processing the activity etc. In fact, it may even decide about the activity itself if activity slot is being allocated. The activity allocator may also ask the activity generator to introduce new activity if it is missing and it may inform the generator about conflicts. Finally, it should be said that we might use the decision postponement strategy during activity allocation as well. Now, we do not choose the unique value for activity parameter immediately but we successively remove values from the variable domain that are inconsistent with values of other variables. By introduction of new activities and constraints we may restrict the domains further until they become singletons or conflict is detected. This technique is called constraint propagation and it is explained in detail in the following section.

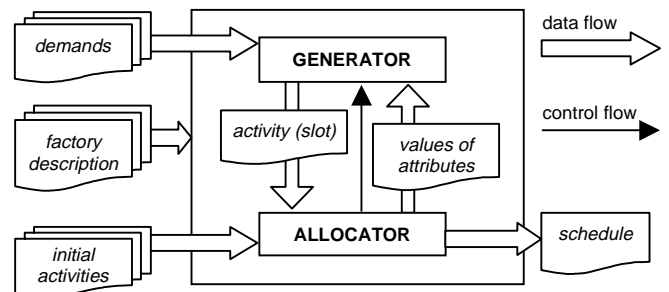


Figure 3. The structure of mixed planning and scheduling system

The scheduling (or is it planning?) is initiated by the list of demands, in production scheduling these demands corresponds to custom orders, and by some initial activities that may describe the

initial situation or requested future activities like maintenance. This generality of input data structure allows us to use the system in conjunction with traditional planners that generate activities in advance. The activities are posted to the activity allocator immediately while the demands are placed in the activity generator. Then, the activity generator uses the information about demands and about allocation of other activities (about attributes' values of activities being allocated) when deciding about introduction of new activity.

4.1 Benefits of proposed architecture

We proposed the mixed planning and scheduling framework to solve the problems in complex-process environments primarily. By allowing the dynamic introduction of activities during scheduling we are able to model problems like set-ups and transitions, processing of by-products and re-cycling, scheduling non-ordered production or choosing alternative processing routes using information about current (partial) schedule. Naturally, the capabilities of the system depends on implementation of individual modules and on choosing the model of given problem [1]. We may still identify some general advantages of the proposed framework.

Generality. The proposed architecture is very general and depending on implementation of activity generator and allocator we get various planning and scheduling systems. We may use it to design generic planners where the role of activity allocator is suppressed to solving planning constraints only. Or we may design the traditional scheduler where (almost) all the activities are inputted and only few activities are introduced during scheduling (for example to model set-ups).

Modularity. Note that the generator and the allocator are still separate modules so we may combine various planners and schedulers. We might even use the same architecture to implement system with separate planner and scheduler (if the planner does not care about attributes' values returned by the scheduler).

Formal interface. The important thing about the architecture is that it formalises the interface between the planner and the scheduler. The planner post activities and (some) constraints among activity attributes to the scheduler while the scheduler returns (partial) valuation of activity attributes.

Early detection of clashes. In general we cannot avoid backtracking during scheduling due to the search nature of most solvers. Because we allow successive introduction of activities, we may detect conflicts earlier so we backtrack sooner and it is not necessary to completely re-plan after the clash.

Active decision postponement. The architecture allows active decision postponement both in planning (via using activity slots and constraints) and scheduling (via partial labelling of attributes).

Single factory description. From the modelling point of view, it is nice that we use single factory description for both planner and scheduler. We may still use parts of the description dedicated to planner or to the scheduler but now the information is available to the other module as well via closer co-operation between the modules.

5 APPLYING CONSTRAINTS FOR PLANNING AND SCHEDULING

When proposing the mixed planning and scheduling framework we assumed constraint programming to serve as the underlying engine

behind the solvers. *Constraint programming* [12] is based on idea of describing the problem declaratively by means of constraints, logical relations among several unknowns (or variables), and, consequently, finding a solution satisfying all the constraints, i.e., assigning a value to each unknown from respective domain. It is possible to state constraints over various domains, however, currently probably more than 95% of all constraint applications deal with finite domains.

At the present time, scheduling is probably the most successful application area of CP [13] while application of CP to planning is not so spread [9]. The reason for this disproportion can be found in the conventional static formulation of constraint satisfaction problem that expects all the elements, i.e., all the variables and all the constraints, to be specified in advance. This is not an obstacle in scheduling tasks where all the activities are known beforehand, however, the plans are highly variable and it is impossible to predict which activities will be used in which combinations.

In the proposed framework, we are using constraint satisfaction technology in the scheduling module to evaluate attributes of activities primarily. The planning module uses the constraint engine via posting activities to the scheduler and via accessing the attributes' values. Naturally, the dynamic nature of the system, where activities (and thus the variables) and constraints are introduced during scheduling, must be considered when choosing the constraint solving package.

There exist two main approaches to solving CSPs (Constraint Satisfaction Problems); one based on constraint propagation, i.e., removing inconsistent values from variables' domains, second based on local search, i.e., altering complete but inconsistent labelling of variables towards (more) consistent labelling. Because of dynamic nature of mixed planning and scheduling, it is more complicated to use local search (we do not have a complete set of variables and constraints beforehand) but as already mentioned the constraint propagation is valuable technique especially for implementation of active decision postponement. We propose using CLP (Constraint Logic Programming) as underlying programming environment as it allows adding new constraints during search as well as removing constraints upon backtracking. Almost all CLP systems use constraint propagation techniques to solve the constraints.

5.1 Constraint models

General architecture of the mixed planning and scheduling provides enough flexibility to solve various problems but the quality (efficiency) of the resulting system is highly dependent on how the problem is modelled. By using the activity generator and the activity allocator we presume using activity based models, i.e., the problem area is described by means of activities. Nevertheless, in scheduling community another model is popular too, namely timetabling model.

Timetabling or time-line model uses discrete time divided into time slices (usually equally distributed) and, then, it describes the situation at each time slice. This model is less appropriate for pure planning as it enforces resource allocation but it can be used in mixed planning and scheduling environment without troubles. In such a case, the architecture of the system is even simpler because we do not need separate activity allocator. We may imagine behaviour of this system as generating all the activity slots first and filing these slots by real activities then. As showed in [1] this model is less appropriate for modelling large-scale problems because of huge number of variables and constraints introduced to model the problem. On the other side, because all the variables and

constraints are posted beforehand we may use popular and very efficient local search algorithms here.

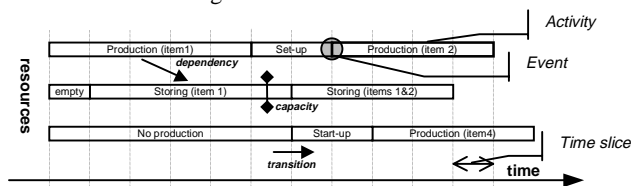


Figure 4. Constraint models and constraint classification

Activity-centric models use event based time and they seem to be better suited for modelling large-scale problems because of smaller memory consumption. In activity based models we may classify the constraints used into three groups:

- *resource constraints*, describing limitations of the resource like capacity,
- *transition constraints*, specifying available transitions between activities in single resource (like set-ups), and
- *dependencies*, capturing relations between activities of different resources (like supplier-consumer relations).

The distribution of the constraints to classes depends on particular problem and according to this distribution we may use either task-centric model where activities are grouped per task or resource-centric model where activities are grouped per resource. Currently, the task-centric models are used mostly due to absence of transition constraints and presence of simple resource constraints. However, as soon as more complicated transition and resource constraints appear, like in complex-process environments, the resource-centric model seems better. The comparison and arguments for choosing particular model can be found in [1,4].

Note that appearance of resource and transition constraints depends on allocation of activities to resources over time. Thus in the proposed mixed planning and scheduling framework, the activity allocator is responsible for posting these constraints. If the task-centric model is used then the planner generates activities per task so it knows the activities belonging to single task. Thus, it may post the dependency constraints to the scheduler. In case of resource-centric model, the situation is more flexible. Now, the generator knows the activities belonging to given resource, in particular it is aware of their order. Consequently, the generator may post the transition constraints. Opposite, a connection between activities using dependency constraint relies on activity allocation and, thus, the scheduler should introduce the dependency constraints. Nevertheless, the scheduler may ask the planner to find dependent activities and pass the responsibility on posting dependency constraints back to the planner.

6 CONCLUSIONS

In the paper we describe a framework for mixing planning and scheduling. We concentrate more on description of motivation for such framework, on highlighting general features and benefits and on explaining the technology behind rather than on specification and experiments with particular implementation.

Nevertheless, the proposed framework is currently being implemented as part of the generic scheduling engine within the VisOpt project. First experience confirms our expectation about general capabilities of the framework; there is no difficulty to model all the typical problems in complex-process environments. This is a big advantage over other scheduling systems that are not able to tackle such complex industries. The experiments with the

implementation also show some interesting results concerning efficiency of the system. It is necessary to find balance between active decision postponement and memory consumption especially when applied to large-scale problems. Also, the implementation indicates that the fully dynamic model, where all the variables and constraints are introduced dynamically, does not exploit the power of constraint propagation and thus the active decision postponement becomes more passive. Consequently, it is better to use semi-dynamic representation where as much as possible activities (activity slots) are posted beforehand (the role of activity generator is suppressed). This observation illustrates why most current scheduling systems deal with static problems only.

ACKNOWLEDGEMENTS

Author's work is supported by the Grant Agency of the Czech Republic under the contract number 201/99/D057 and by InSol Ltd. I would like to thank Yossi Rissin and the team of InSol for introducing me to the problem and for interesting and encouraging discussions concerning real-life problems of industrial planning and scheduling. I am also grateful to Helmut Simonis from Cosytec for discussions concerning the scheduling applications and for drawing my attention to some existing applications in the area.

REFERENCES

- [1] Barták, R.: Conceptual Models for Combined Planning and Scheduling. *Proceedings of the CP99 Post-conference Workshop on Large Scale Combinatorial Optimisation and Constraints*, Alexandria, USA (1999) 2-14
- [2] Barták, R.: On the Boundary of Planning and Scheduling: a Study. *Proceedings of the Eighteenth Workshop of the UK Planning and Scheduling Special Interest Group*, Manchester, UK (1999) 28-39
- [3] Blum, A. L. and Furst, M. L.: Fast Planning Through Planning Graph Analysis. *Artificial Intelligence* 90 (1997) 281-300
- [4] Brusoni, V., Console, L., Lamma, E., Mello, P., Milano, M., Terenziani, P.: Resource-based vs. Task-based Approaches for Scheduling Problems. *Proceedings of the 9th ISMIS96, LNCS Series*, Springer Verlag (1996)
- [5] Caseau, Y., Laburthe, F.: A Constraint based approach to the RCPSP. *Proceedings of the CP97 Workshop on Industrial Constraint-Directed Scheduling*, Schloss Hagenberg, Austria (1997)
- [6] Fikes, R. E. and Nilsson, N. J.: STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. *Artificial Intelligence* 3-4 (1971) 189-208
- [7] Joslin, D. and Pollack M.E.: Passive and Active Decision Postponement in Plan Generation. *Proceedings of the Third European Conference on Planning* (1995)
- [8] Koehler, J.: Planning under Resource Constraints. *Proceedings of 13th European Conference on Artificial Intelligence*, Brighton, UK (1998) 489-493
- [9] Nareyek, A.: AI Planning in a Constraint Programming Framework. *Proceedings of the Third International Workshop on Communication-Based Systems* (2000), to appear
- [10] Pegman, M.: Short Term Liquid Metal Scheduling. *Proceedings of PAPPACT98 Conference*, London (1998) 91-99
- [11] Srivastava, B. and Kambhampati, S.: Scaling up Planning by teasing out Resource Scheduling. Technical Report ASU CSE TR 99-005, Arizona State University (1999)
- [12] Tsang, E.: *Foundations of Constraint Satisfaction*. Academic Press, London (1995)
- [13] Wallace, M.: Applying Constraints for Scheduling. *Constraint Programming*, Mayoh B. and Penjaak J. (Eds.), NATO ASI Series, Springer Verlag (1994)