Schedule Management: an Object Oriented Approach

by

G. Wolf

Computing Science Note 91/22
Eindhoven, September 1991
This is a series of notes of the Computing Science Section of the Department of Mathematics and Computing Science Eindhoven University of Technology. Since many of these notes are preliminary versions or may be published elsewhere, they have a limited distribution only and are not for review. Copies of these notes are available from the author.

Copies can be ordered from:
Mrs. F. van Neerven
Eindhoven University of Technology
Department of Mathematics and Computing Science
P.O. Box 513
5600 MB EINDHOVEN
The Netherlands
ISSN 0926-4515

All rights reserved
editors: prof.dr.M.Rem
        prof.dr.K.M.van Hee.
Abstract. In this paper we consider resource-constrained time-dependent scheduling systems, i.e. decision support systems with time as the important planning component. Instead of dealing with optimization aspects of the planning problem, we concentrate on schedule management, i.e. stepwise planning with respect to primitive functions, like handling single decisions or constraints.

The design of these systems is based on a mathematical model, giving a formal characterization of a class of scheduling problems and allowing generic descriptions of scheduling objects like processors, operations, decisions and constraints. The model is applied to various scheduling problems, like resource constrained project scheduling, car routing, the construction of time tables both for schools and for nursery in hospitals.

An object oriented implementation of the model, based on the natural hierarchy of scheduling problems, turns out to lead to a clear separation between the generic and the domain specific components of the schedule manager, minimizing redundant code and resulting in software with a high degree of maintainability.

Keywords. Decision Support System, Scheduling, Object Oriented Design.
1. Introduction

Software for decision support systems is very sensitive to change requests. A minor change in the underlying model or in the limiting conditions can make it necessary to redesign the whole system. There are hardly reusable components. The claim for a more flexible DSS-design is uttered by many authors [2,6,10,20].

The reason for this inflexibility can be found among others in the fact, that conventional DSS-systems are dominated by the algorithmic aspect. Although it is stated, that the software components of a DSS should relate to both information management and optimization [3,6,16], more attention is paid to the latter aspect, especially in the first design phase. Mostly mathematicians working in the field of operations research were involved in the software development. In spite of the mathematical relevance of algorithmic aspects, the software dealing with it plays an inferior role related to other software components like the management of decisions and the user interface.

In view of the fact, that software maintainability is our highest objective in this paper, the following requirements can be formulated:

- the software should contain reusable modules, that can be applied to a wider class of planning problems;
- the software architecture should allow a clear separation between the generic components and the domain specific components of the software;
- the software should be a platform with respect to the implementation for both generic algorithms (simulated annealing, genetic algorithms, etc.) [1] and powerful domain specific optimization procedures;
- the software should be flexible with respect to changes in the (domain specific) model and it should be easy to add or to change constraints.

Although several attempts have been made to increase flexibility of optimization components by using general search methods applicable in a wide range of planning situations [10], we concentrate in this paper on data manipulation aspects, particularly oriented towards scheduling problems. The basic system functionality is schedule management. Primitive functions should be provided to support stepwise planning [15]. The user should have among others the possibility to add (and delete) scheduling objects like "decisions" and "constraints"; the system should calculate the consequences of user actions and check feasibility. Optimization is not relevant in this context.

Scheduling is the allocation of resources over time to perform certain tasks [5]. Scheduling problems can be formulated as sequencing problems with time as the important planning component. Moreover, we require that all feasibility constraints and schedule evaluation criteria can be expressed as measures of time. The class of these scheduling problems is rather extensive and does include job-shop planning, car routing, school time tables and time tables for nursery in hospitals. We will give a formal definition of this class. Although several scheduling models exist [5,9,13], which are suitable to formulate optimization problems, they are not flexible enough for our purposes.

Because we need the flexibility of generic models as well as the power of problem specific models, modelling on more abstraction levels is necessary. This implies a software architecture consisting of generic and domain specific modules [20]. A specific
schedule manager is built up by linking the generic modules with the relevant group of
domain specific modules.

Because of our high requirements for software maintainability, system implementa-
tion is based on an object oriented design [17,19]. The use of object oriented techniques
for DSS-software has been up to now rather limited [8,12]. There are however many
reasons, that plead for their application:

- the concept of encapsulation, to specify objects both by their attributes and by their
  methods, is a powerful mechanism to represent scheduling objects like processors,
  operations, decisions and constraints;

- the concept of inheritance (and abstract datatyping), to formalize IS-A relationships
  between classes and subclasses, is necessary to represent the natural hierarchy of
  scheduling problems and scheduling objects to get a clear separation between the
  generic and the domain specific components;

- polymorphism, i.e. the possibility to use object methods in different subclasses under
  the same name, but with specific functionalities, will enlarge the flexibility of the
  software.

2. Notations

We will follow some notational conventions:

\[ A \approx B \] : set \( A \) and set \( B \) are equivalent (of equal power);
\[ [A \rightarrow B] \] : the set of functions from \( A \) to \( B \);
\( A^* \) : the set of strings of elements from set \( A \);
\( \mathcal{P}(A) \) : the powerset of \( A \) (set of all subsets of \( A \));

\[ \bigcup \{A(i) \mid P(i)\} \] : the union of all sets \( A(i) \), for which \( P(i) \) holds;
\[ \sum \{a(i) \mid P(i)\} \] : the sum of all \( a(i) \), for which \( P(i) \) holds.

Let \( \text{Time} \) be a finite interval on the real axis. With \( B(\text{Time}) \) we denote the set
algebra generated by subintervals of \( \text{Time} \).

Let \( A \) be an entitytype with attributes \( A_1, \ldots, A_n \), i.e. \( A \approx A_1 \times \ldots \times A_n \), and let
\( t \in A \) be an entity of that type. With \( A_i(t) \) we denote the projection of \( t \) to \( A_i \).

3. A basic general model

We restrict ourselves to resource-constrained time-dependent planning problems with
the property, that schedules can be represented as Gantt charts. This problem class is
very wide and includes the general jobshop problem, car routing and the construction
of time tables both for schools and for nursery in hospitals.
Existing models for scheduling problems are either too specific with regard to constraints [5,9,13] or too general to be powerful enough to characterize the time aspects of scheduling problems properly [10,14]. For this reason we present in this section a basic general model to characterize our problem class; extensions and problem specific interpretations follow later. Our object definitions (see section 7) will be based on this model.

3.1. Scheduling problems

A class of scheduling problems can be characterized by the following tuple

\((P, O, D, pr, op, S, f, g, t_{\text{max}})\) with

- \(P\), the set of all possible processors (or resources);
- \(O\), the set of all possible operations (or tasks) to be assigned to processors;
- \(D\), the set of all possible decisions; we have \(D = D_+ \cup D_0\), with \(D_+\) being the set of decisions, that assign an operation to a processor (real decisions) and \(D_0\) being the set of decisions, that put a processor in the waiting state (idle decisions);
- the function \(pr : D \rightarrow P\) assigning the processor, involved in a decision;
- the function \(op : D_+ \rightarrow O\) assigning the operation, involved in a real decision;
- \(S\), the set of all possible states, that a processor can have with the function \(s_0 : P \rightarrow S\) assigning the initial states;
- the function \(f : D \times S \rightarrow S\) describing a state transition, being the effect of a decision with respect to a processor;
- the function \(g : D \times S \rightarrow R_+\) to determine the duration, that a certain decision will be active;
- \(t_{\text{max}}\), a positive real number, defining the scheduling interval \(\text{Time} = (0, t_{\text{max}}]\).

3.2. Schedules

Scheduling problems are distinguished from other planning problems by the aspect of time. All time aspects should be handled in the generic part of the system.

Let \(p_i\) be a processor from a finite subset of \(P\) and let \(\{d_{i,j}\}_{j=1,n_i}\) be a finite sequence of decisions with respect to \(p_i\). The decision \(d_{i,j}\) is effectuated at the time \(t_{i,j}\) and is active till \(t_{i,j+1}\), the time, that the next decision with respect to \(p_i\) will be effectuated. We call \(t_{i,j}\) a decision-point. Let \(s_{i,j}\) be the state of processor \(p_i\) after decision \(d_{i,j}\) \((j > 0)\). Then we have the following equations

\[
\begin{align*}
\text{s}_{i,0} &= s_0(p_i) \\
\text{s}_{i,j} &= f(\text{s}_{i,j-1}, d_{i,j}) \quad (j > 0) \\
t_{i,1} &= 0 \\
t_{i,j} &= t_{i,j-1} + g(\text{s}_{i,j-2}, d_{i,j-1}) \quad (j > 1) .
\end{align*}
\]
We call \( \text{sn}(p_i) = s_{i,n_i} \) the final state of processor \( p_i \), which is the state after the last decision.

We can define a schedule as a tuple \((P_S, O_S, F_S)\) with

- \( P_S \) a finite subset of \( P \),
- \( O_S \) a finite subset of \( O \),
- a function \( F_S : P_S \rightarrow D^* \) assigning a finite sequence of decisions to every processor of \( P_S \),

\[
F_S : p_i \mapsto \{d_{i,j}\}_{j=1}^{n_i}, \quad (p_i \in P_S; \; d_{i,j} \in D) \quad \text{with} \quad \Pr(d_{i,j}) = p_i, \quad (\op(d_{i,j}) \in O_S \text{ or } d_{i,j} \in D_0) \quad \text{and} \quad \sum (g(d_{ij}, s_{ij}) | 0 < j \leq n_i) \leq t_{\text{max}}.
\]

As a consequence of this definition only one operation at a time can be allocated to a processor, whereas more processors can be assigned to the same operation simultaneously.

3.3. Constraints

Constraints are used to define the feasibility of schedules. It is important to have the flexibility to add and to change constraints dynamically. To realize this we should represent planning knowledge and especially constraints declaratively in a domain independent way. Several ideas and concepts going in this direction, can be found in the literature [7,11,12,20].

Although we assume a certain homogeneity with respect to scheduling objects like resources, operations and decisions, this is not the case with constraints. On the domain specific level various subclasses (types) of constraints may be defined (deadline constraints, precedence constraints, etc.). It is therefore important to distinguish which properties are adherent to constraint classes and which to constraint instances.

We distinguish hard and weak constraints. The feasibility of a schedule with respect to hard constraints is guaranteed by the schedule manager, whereas violations of weak constraints are only signalized to the user, who is ultimately responsible for feasibility. The hardness of a constraint is specified on instance level.

We require, that constraint violations can be expressed by time intervals, denoting the period, when the constraint is violated or when decisions should be effectuated to undo the constraint violation. This is just a generalization of the convention to specify constraints by Boolean functions. This is done for the following reasons:

- the time aspect of the constraint violation is more emphasized, which is rather important for scheduling problems;
- we have a measure of infeasibility with respect to constraints, which is important for weak constraints;
a wide class of schedule quality measures can be formulated as simple expressions of critical regions, belonging to the same class of constraints (make span, earliness and tardiness for the job-shop problem, etc.).

Now we come to a formal definition of constraints. Let \( \text{Sch} \) be the set of all schedules with respect to a certain problem class. We define a constraint-type \( c \) by its specific domain and its evaluator, i.e. a pair \( (Z_c, \text{eval}_c) \) with \( Z_c \) a set and \( \text{eval}_c \) a (penalty) function with
\[
\text{eval}_c : \text{Sch} \times Z_c \rightarrow B(\text{Time})
\]
A constraint instance of type \( c \) can be defined as a pair \( (z_c, \text{hd}) \) with \( z_c \in Z_c \) and \( \text{hd} \), being a Boolean to specify the hardness of a constraint instance.

Evaluation of the penalty function \( \text{eval}_c \) for a particular constraint instance and schedule will produce a subset of the scheduling interval \( \text{Time} \) to indicate the time period, when the constraint is violated or when decisions should be effectuated to undo the constraint violation. This time period, which we call the critical region, can be represented as a finite set of disjunct intervals. The critical region can be empty in the case of feasibility or contain just the whole time interval as extreme cases.

In the Section 4 this definition will be illustrated by a variety of examples.

3.4. Quality measures

Beside constraints, that define the feasibility of schedules, there are quality measures for comparison of schedules. These criteria are often defined as penalty functions, mapping a schedule to a real value [13]. It makes sense to couple quality measures to constraints or sets of constraints by measuring the extent of violation with respect to these constraints. For the most common quality measures (make span, total or weighted tardiness, etc.) corresponding constraints can be defined.

The relation between constraints and quality measures will be formalized by the following definition. A quality measure can be defined by a pair \((C_{\text{list}}, \text{agg})\) with \( C_{\text{list}} \) being a finite list of constraints and \( \text{agg} \) being a real (aggregate) function defined on the set of strings of real numbers
\[
\text{agg} : R^* \rightarrow R.
\]
Examples of such aggregate functions are \( \text{sum}(X) \), \( \text{max}(X) \), \( \text{variance}(X) \), etc. with \( X \) a sequence of reals.

The constraint list mentioned above will contain often all constraint instances of a certain type, but this is not necessary.

Given a particular schedule a quality value \( q \) with respect to a quality measure \((C_{\text{list}}, \text{agg})\) can be calculated by applying the specified aggregate function to the list of the lengths of the critical time regions, defined by the constraint evaluation functions corresponding to the specified list of constraints (see fig. 3.1); the length of an element of \( B(R) \) is defined by its Lebesque measure \( \mu \).
For example we can define the make span of a schedule for the job-shop problem by (virtual) deadline constraints (with deadlines equal to zero) and the aggregate function giving the maximum of a string of real numbers.

3.5. Extended schedules

Let \((P, O, D, pr, op, S, f, g, t_{max})\) be a class of scheduling problems and let \(C\) be the set of all possible constraints for this class. We define an extended schedule by a tuple \((P, O, F, C, S, f, g, t_{max})\) with \((P, O, F, C, S)\) being a schedule, \(C_S\) a subset of \(C\) and \(Q_s\) a set of quality measures with \(c \in C_S\) for every constraint instance \(c\) being involved in \(Q_s\) (i.e. \(c\) an element of some \(Clist\) with \((Clist, agg) \in Q_s\)).

3.6. Expressions

In this section we assume a fixed given schedule \((P, O, F, S)\). The set \(D_S\) of decisions involved is determined by
\[
D_S = \{d_{i,j} | d_{i,j} = (F_S(p_i))j, p_i \in P_S, 0 < j \leq n_i\}.
\]
We can now introduce some expressions, that are needed in the following paragraphs:

- the start time of decision \(d_{i,j}\):
  \[
  \text{start}(d_{i,j}) = \sum (g(s_{i,k-1}, d_{i,k}) | 0 < k < j) \quad (j > 1)
  \]
  \[
  \text{start}(d_{i,1}) = 0
  \]

- the end time of decision \(d_{i,j}\):
  \[
  \text{end}(d_{i,j}) = \begin{cases} 
  \text{start}(d_{i,j}) + g(s_{i,j-1}, d_{i,j}) \\
  \text{start}(d_{i,j+1}) \quad (j < n_i)
  \end{cases}
  \]

- the activity interval of decision \(d_{i,j}\):
  \[
  \text{interval}(d_{i,j}) = (\text{start}(d_{i,j}), \text{end}(d_{i,j}))
  \]

- the time period, when processor \(p\) executes operation \(o\):
  \[
  \text{pr-op-set}(p, o) = \bigcup (\text{interval}(d_{i,j}) | \text{pr}(d_{i,j}) = p, \text{op}(d_{i,j}) = o, d_{i,j} \in D_S)
  \]
- the executing period of operation \( o \):

\[
\text{op-exec-set}(o) = \bigcup (\text{pr-op-set}(p, o) \mid p \in P_S)
\]

- the start time of operation \( o \):

\[
\text{start-op}(o) = \begin{cases} \min(\text{start}(d_{i,j}) \mid \text{op}(d_{i,j}) = o, d_{i,j} \in D_S) & \text{if op-exec-set}(o) \neq \emptyset \\ t_{\text{max}} & \text{otherwise} \end{cases}
\]

- the end time of operation \( o \):

\[
\text{end-op}(o) = \begin{cases} \max(\text{end}(d_{i,j}) \mid \text{op}(d_{i,j}) = o, d_{i,j} \in D_S) & \text{if op-exec-set}(o) \neq \emptyset \\ t_{\text{max}} & \text{otherwise} \end{cases}
\]

- the finishing time of processor \( p \):

\[
\text{end-pr}(p) = \sum(g(d_{i,j}, s_{i,j-1}) \mid \text{pr}(d_{i,j}) = p, d_{i,j} \in D_S)
\]

- the allocation of processor \( p \) to operation \( o \) at time \( t \):

\[
\text{busy}(p, o, t) = \begin{cases} \top & \text{if } t \in \text{pr-op-set}(p, o) \\ \bot & \text{otherwise} \end{cases}
\]

- the cumulative activity time of decision \( d_{i,j} \) at time \( t \):

\[
\text{dur}(d_{i,j}, t) = \begin{cases} \text{end}(d_{i,j}) - \text{start}(d_{i,j}) & \text{if } t \geq \text{end}(d_{i,j}) \\ t - \text{start}(d_{i,j}) & \text{if } \text{start}(d_{i,j}) < t < \text{end}(d_{i,j}) \\ 0 & \text{otherwise} \end{cases}
\]

- the cumulative allocation time of processor \( p \) to operation \( o \) at time \( t \):

\[
\text{pr-op-dur}(p, o, t) = \sum(\text{dur}(d_{i,j}, t) \mid \text{pr}(d_{i,j}) = p, \text{op}(d_{i,j}) = o, d_{i,j} \in D_S)
\]
4. Domain specific interpretations of the generic model

In this section we give domain specific interpretations of the model. The basic model may be extended with domain specific parameters. We specify attributes of the relevant entities and give specifications of the relevant functions. Constraints are defined by their specific domain and their evaluators (see section 3.3).

4.1. Job-shop

Problem description:

We consider general job-shop scheduling with job splitting and preemption. Operations (tasks) may be allocated several times to processors (resources). An operation has a size, i.e. the total resource capacity necessary to complete the task. Processors can execute operations with different speeds. To execute operations certain abilities of processors are required. The sequence of task execution is restricted by precedence constraints. Deadline and release time constraints are related to the finishing and starting times of operations. Processors are only at certain time windows available.

Domain specific model parameters:

\( A \) – a set of abilities.

Processors:

\( id \) : identifier
\( abilset \) : set of abilities
\( speed \) : executing velocity

\[ P \approx N \times \mathcal{P}(A) \times R_+ \]

Operations:

\( id \) : identifier
\( req-abilset \) : set of required abilities
\( size \) : processing capacity

\[ O \approx N \times \mathcal{P}(A) \times R_+ \]

Real decisions:

\( pr \) : processor reference
\( op \) : operation reference
\( time \) : duration

\[ D_+ \approx P \times O \times R_+ \]
Idle decisions:

\( pr \): processor reference

\( wait \): waiting time

\( D_0 \approx P \times R_+ \)

States (of a processor):

\( \text{exectime} \): the total executing time of a processor

(till a decision point)

\( S \approx R_+ \)

Initial state (of a processor):

\( s_0(p) = 0 \ \forall p \in P \)

State transition:

\[
f(state, dec) = \begin{cases} 
exectime(state) + time(dec) & \text{if } dec \in D_+ \\
exectime(state) & \text{if } dec \in D_0 \end{cases}
\]

Duration of decision:

\[
g(state, dec) = \begin{cases} 
time(dec) & \text{if } dec \in D_+ \\
wait(dec) & \text{if } dec \in D_0 \end{cases}
\]

Constraints:

- release time constraints: \( Z_c = O \times R_+ \)

\[
eval_c(sch, o, rl) = \{ t \in Time \ | \ t < rl \} \cap \text{op-exec-set(o)}
\]

- deadline constraints: \( Z_o = O \times R_+ \)

\[
eval_c(sch, o, dl) = \{ t \in Time \ | \ t > dl \} \cap \text{op-exec-set(o)}
\]

- tasks should be completed: \( Z_c = O \)

let \( \text{compl(o)} \) be the completed part of an operation:

\( \text{compl(o)} = \sum (pr-op-dur(p_i, o, t_{max}) \ast speed(p_i) \ | \ p_i \in P_S) \)

and let \( \alpha = t_{max} \ast \text{compl(o)}/\text{size(o)} \)

\[
eval_c(sch, o) = \begin{cases} 
\emptyset & \text{if } \text{compl(o)} \geq \text{size(o)} \\
[\alpha, t_{max}] & \text{otherwise}
\end{cases}
\]
- avoiding superfluous task execution: \( Z_c = O \)

\[
eval_c(sch, o) = \text{op-exec-set}(o) \cap \\
\{t \in \text{Time} | \sum_{\text{pr-op-dur}(p_i, o, t) \ast \text{speed}(p_i) > \text{size}(o) | p_i \in P_S}\}
\]

- required processor abilities for task execution: \( Z_c = P \times O \)

\[
eval_c(sch, p, o) := \begin{cases} \\
\emptyset & \text{if } \text{abilset}(p) \cap \text{req-abilset}(o) \neq \emptyset \\
\text{pr-op-set}(p, o) & \text{otherwise}
\end{cases}
\]

- generalized precedence constraints: \( Z_c = O \times O \times R_+ \)

\[
eval_c(sch, o_1, o_2, \text{waiting}) := \{ t \in \text{Time} | t < \text{end-op}(o_1) + \text{waiting} \} \\
\cap \{ t \in \text{Time} | t > \text{start-op}(o_2) \} \text{ with } o_1 \neq o_2
\]

- processor unavailability: \( Z_c = P \times B(\text{Time}) \)

\[
eval_c(sch, p, \text{period}) := \text{period} \cap \left( \bigcup \{ \text{pr-op-set}(p, o_k) | o_k \in O_S \} \right)
\]

4.2. Car routing

Problem description:

Cargo (operations) has to be transported by trucks (processors) from one location to another. Transport time is dependent on the distance of the locations, the speed of a truck and the loading resp. unloading time. The amount of cargo that is transported simultaneously is restricted by a maximum volume, which is truck specific. Starting points may be different. Loading or unloading is only possible in specific time windows.

Domain specific model parameters:

\( L \) – a set of locations.  
\( \text{dist} : L \times L \rightarrow R_+ \) a distance function.

Processors:

\begin{align*}
\text{id} & : \text{car-identifier} \\
\text{start-loc} & : \text{start location} \\
\text{speed} & : \text{speed of the truck} \\
\text{maxvolume} & : \text{maximal total volume}
\end{align*}

\( P \approx N \times L \times R_+ \times R_+ \)
Operations:

\[ id : \text{cargo-identifier} \]
\[ from-loc : \text{place of depart} \]
\[ to-loc : \text{destination} \]
\[ load-time : \text{time to load or to unload} \]
\[ volume : \text{volume of the cargo} \]

\[ O \approx N \times L \times L \times R_+ \times R_+ \]

Real decisions:

\[ pr : \text{processor reference} \]
\[ op : \text{operation reference} \]
\[ loading : \text{loading or unloading} \]

\[ D_+ \approx P \times O \times \{ \top, \bot \} \]

Idle decisions:

\[ pr : \text{processor reference} \]
\[ wait : \text{waiting time} \]

\[ D_0 \approx P \times R_+ \]

States (of a processor):

\[ loc : \text{the location of the truck} \]
\[ vol : \text{the total volume of the cargo} \]

\[ S \approx L \times R_+ \]

Initial state (of a processor):

\[ s_0(p) = (\text{start-loc}(p), 0) \; \forall p \in P \]

State transition:

\[
\begin{align*}
    f(state, dec) = & \begin{cases} 
        (\text{from-loc}(op(dec)), vol(state) + volume(op(dec))), & \text{if } dec \in D_+, \text{ loading}(dec) = \top \\
        (\text{to-loc}(op(dec)), vol(state) - volume(op(dec))), & \text{if } dec \in D_+, \text{ loading}(dec) = \bot \\
        state & \text{if } dec \in D_0
    \end{cases}
\end{align*}
\]
Duration of decision:

\[
g(state, dec) = \begin{cases} 
\text{dist}(\text{loc}(\text{state}), \text{from-loc}(\text{op}(\text{dec}))) \ast \text{speed}(\text{pr}(\text{dec})) + \text{load-time}(\text{op}(\text{dec})) & \text{if } dec \in D_+, \text{ loading}(\text{dec}) = T \\
\text{dist}(\text{loc}(\text{state}), \text{to-loc}(\text{op}(\text{dec}))) \ast \text{speed}(\text{pr}(\text{dec})) + \text{load-time}(\text{op}(\text{dec})) & \text{if } dec \in D_+, \text{ loading}(\text{dec}) = \perp \\
\text{wait}(\text{dec}) & \text{if } dec \in D_0
\end{cases}
\]

Constraints:

- An operation is executed by one processor only: \( Z_c = 0 \)
  
  \[
eval_c(sch, o) = \begin{cases} 
opexecset(o) & \text{if } \# \{ i \mid \text{pr-op-set}(p_i, o) \neq \emptyset \} > 1 \\
0 & \text{otherwise}
\end{cases}
\]

- Operations should be completed: \( Z_c = 0 \)
  
  \[
eval_c(sch, o) = \begin{cases} 
\emptyset & \text{if } \nopexecset(o) \neq \emptyset \\
(0, t_{\text{max}}] & \text{otherwise}
\end{cases}
\]

- Release time constraints: see job-shop

- Deadline constraints: see job-shop

- Loading/unloading is not allowed in certain time periods: \( Z_c = L \times B(\text{Time}) \)
  
  \[
eval_c(sch, loc, period) = \text{period} \cap (\bigcup \{ \text{end}(d_{i,j}) - \text{loadtime}(\text{op}(d_{i,j})), \text{end}(d_{i,j}) \mid d_{i,j} \in D_+, \text{ from-loc}(\text{op}(d_{i,j})) = \text{loc} \} \vee \text{to-loc}(\text{op}(d_{i,j})) = \text{loc})
\]

- Trucks should not be overloaded: \( Z_c = P \)
  
  \[
eval_c(sch, p) = \bigcup \{ \text{interval}(d_{i,j}) \mid d_{i,j} \in D_+, \text{ pr}(d_{i,j}) = p, \text{ vol}(s_{i,j}) > \text{ supvol} \}
\]

- Trucks should return to their starting points: \( Z_c = P \)

Let \( \text{returntime}(p) \) be the time necessary to return from the final location back to the starting point:

\[
\text{returntime}(p) = \text{speed}(p) \ast \text{dist}(s_{f\text{in}}(p), \text{ start-loc}(p))
\]

and let \( \alpha = \text{maximum}(0, t_{\text{max}} - \text{returntime}(p)) \)

\[
eval_c(sch, p) = \begin{cases} 
\emptyset & \text{if } \text{loc}(s_{f\text{in}}(p)) = \text{ start-loc}(p) \\
(\alpha, t_{\text{max}}] & \text{otherwise}
\end{cases}
\]
two (or no) decisions are related to a processor and an operation: \( Z_c = P \times O \)

\[
eval_c(sch, p, o) = \begin{cases} 
pr-op-set(p, o) & \text{if } \# \{(i, j) \mid di,j \in D_+, pr(di,j) = p, op(di,j) = o\} \neq 2 \\
\emptyset & \text{otherwise}
\end{cases}
\]

- cargos should be loaded before being unloaded: \( Z_c = P \times O \)

\[
eval_c(sch, p, o) = \begin{cases} 
pr-op-set(p, o) & \text{if } \exists i, j_1, j_2 : j_1 < j_2, di,j_1, di,j_2 \in D_+, \\
pr(di,j_1) = pr(di,j_2) = p, op(di,j_1) = op(di,j_2) = o, \\
\text{loading}(di,j_1) = \bot \lor \text{loading}(di,j_2) = \top) & \text{otherwise}
\end{cases}
\]

4.3. Time tables for schools

Problem description:
Weekly time tables for schools are considered. Classrooms (operations) are to be allocated to courses (processors) at certain time units. Allocation time is constant. It is assumed, that the allocation of teachers and classes to courses is not effectuated by a decision, since this is no scheduling problem. One should care, that teachers, classes and classrooms cannot be allocated to different courses simultaneously. However more than one class can participate in one course simultaneously. A classroom has a capacity, that has to correspond to the size of the classes.

Domain specific model parameters:

\[ K \] – a set of classes.
\[ M \] – a set of teachers.
\[ n_{\text{max}} \] – maximal number of lessons in a week.

Processors:

\[ id \] : course-identifier
\[ teacher \] : teacher of the course
\[ classes \] : set of classes participating in the course
\[ subject \] : subject of the course
\[ students \] : number of participating students
\[ lessons \] : number of lessons weekly

\[ P \approx N \times M \times \mathcal{P}(K) \times N \times N \times N \]

Operations:

\[ id \] : classroom-identifier
\[ capacity \] : capacity of the classroom

\[ P \approx N \times R_+ \]
Real decisions:

- \( pr \) : processor reference
- \( op \) : operation reference

\[ D_+ \approx P \times O \]

Idle decisions:

- \( pr \) : processor reference

\[ D_0 \approx P \]

States (of a processor):

- \( \text{exectime} \) : the number of lessons executed

\[ S \approx R_+ \]

Initial state (of a processor):

\[ s_0(p) = 0 \ \forall p \in P \]

State transition:

\[ f(state, dec) = \begin{cases} 
\text{exectime}(state) + 1 & \text{if } dec \in D_+ \\
\text{exectime}(state) & \text{if } dec \in D_0 
\end{cases} \]

Duration of decision:

\[ g(state, dec) = 1 \]

that means:

\[ \text{interval}(d_{i,j}) = (j-1, j] \]

Scheduling interval:

\[ (0, t_{\text{max}}] = (0, n_{\text{max}}] \]

Constraints:

- teachers are indivisible: \( Z_t = M \)

\[ \text{eval}_t(sch, m) = \bigcup((j-1, j] | \# \{i \mid \text{teacher}(pr(d_{i,j})) = m \} > 1) \]

- classes are indivisible: \( Z_c = K \)

\[ \text{eval}_c(sch, k) = \bigcup((j-1, j] | \# \{i \mid k \in \text{classes}(pr(d_{i,j})) \} > 1) \]
- classrooms are indivisible: \( Z_c = 0 \)
  
  \[ \text{eval}_c(sch, o) = \bigcup((j - 1, j]) \mid \#\{i \mid \text{op}(d_{i,j}) = o\} > 1 \]

- restricted classroom capacity: \( Z_c = 0 \)
  
  \[ \text{eval}_c(sch, o) = \bigcup(\text{interval}(d_{i,j}) \mid d_{i,j} \in D_+, \text{op}(d_{i,j}) = o, \text{students}(pr(d_{i,j})) > \text{capacity}(o)) \]

- not too many lessons for a course: \( Z_c = P \)
  
  \[ \text{eval}_c(sch, p) = \bigcup(\text{interval}(d_{i,j}) \mid d_{i,j} \in D_+, \text{pr}(d_{i,j}) = p, \text{execitime}(s_{i,j}) > \text{lessons}(p)) \]

- enough lessons for a course: \( Z_c = P \)
  
  \[ \text{eval}_c(sch, p) = \bigcup(\text{interval}(d_{i,j}) \mid d_{i,j} \in D_0, \text{pr}(d_{i,j}) = p, (\text{lessons}(p) - \text{execitime}(s_{i,j})) > (n_{\text{max}} - j)) \]

- teachers unavailability: \( Z_c = M \times B(\text{Time}) \)
  
  \[ \text{eval}_c(sch, m, \text{period}) := \text{period} \cap \left( \bigcup(\text{op-set}(p_i, o_k) \mid p_i \in P_S, o_k \in O_S, \text{teacher}(p_i) = m) \right) \]

4.4. Time tables for nursery staff

Problem description:

Duty rosters for nursery staff in hospitals for a fixed time period is considered. Nurses (processors) can be on or off duty. To be on duty is the only operation. Nurses have different maximal working hours and different functions (sets of abilities). A certain minimal crew is required, which is time dependent and which is expressed in numbers of necessary functions (abilities).

Domain specific model parameters:

\( A \) - a set of abilities.

Processors:

\( id \) : identifier

\( abilset \) : set of abilities
suptime : maximal working time

\[ P \approx N \times \mathcal{P}(A) \times R_+ \]

Operations:

\( id \) : identifier (constant)

\( O \approx \{ o_1 \} \) (singular set)

Real decisions:

\( pr \) : processor reference
\( op \) : operation reference
\( time \) : duration

\[ D_+ \approx P \times O \times R_+ \approx P \times R_+ \]

Idle decisions:

\( pr \) : processor reference
\( wait \) : waiting time

\[ D_0 \approx P \times R_+ \]

States (of a processor):

exectime : the total executing time of a processor
(till a decision point)

\[ S \approx R_+ \]

Initial state (of a processor):

\( s_0(p) = 0 \ \forall p \in P \)

State transition:

\[ f(state, dec) = \begin{cases} 
\text{exectime}(state) + \text{time}(dec) & \text{if } dec \in D_+ \\
\text{exectime}(state) & \text{if } dec \in D_0 
\end{cases} \]

Duration of decision:

\[ g(state, dec) = \begin{cases} 
\text{time}(dec) & \text{if } dec \in D_+ \\
\text{wait}(dec) & \text{if } dec \in D_0 
\end{cases} \]
Constraints:

- maximal working times for nurses: $Z_c = P$

  $\text{eval}_c(p) = \{ t \in \text{Time} \mid \text{pr-op-dur}(p, o_1, t) > \text{suptime}(p) \} \cap \text{pr-op-set}(p, o_1)$

- minimal crew requirements: $Z_c = B(\text{Time}) \times P(A) \times N$

  $\text{eval}_c(sch, period, req-abilset, min-charge) := period \cap$
  
  $\{ t \in \text{Time} \mid \text{min-charge} > \# \{ i \mid \text{busy}(p_i, o_1, t), \text{abilset}(p_i) \cap \text{req-abilset} \neq \emptyset \} \}$

- processor unavailability: see job-shop
5. System functionality and user actions

A prototype is implemented based on our model with a primitive user interface. This system supports step-wise planning. The user can perform user actions by evoking interactively primitive functions like adding or deleting "scheduling objects" like processors, operations, decisions or constraints. The system will evaluate these user actions, compute their consequences and do some feasibility checking.

These primitive functions are grouped with respect to the relevant object classes. After the choice of the relevant class (schedule, processor, etc.) the user may evoke a function (find, insert, etc.). The system will execute that function and perform state transitions if necessary.

A transaction is a sequence of user actions followed by a commit. After a commit command all constraint instances will be evaluated resulting in a set of critical regions. Moreover quality values will be calculated for all defined quality measures. In the case that any hard constraint is violated all actions performed since the last commit are rolled back.

Grouping of user actions by means of the transaction concept is necessary, because a single user action may result in an infeasible schedule (some decisions may be effec­tuated at the same time involving different processors). To avoid the violation of hard constraints, eventually constraint checking should be done with respect to more than one user action.

With respect to the scheduling objects processor, operation, decision and constraint the following primitive functions are defined:

find, insert, update, delete

With the function find a specific object instance is made current within his class. Several options of this command are possible: find direct, find first, find next, find last and find previous. Direct search is always based on the object identifier. Sequential search for processors, operations and constraints is based on the order of the object identifiers; sequential search for decisions is based on the chronological order of their decision points. By means of the options find within processor, find within operation and find within constraint type the search process for decisions or constraints may be restricted.

The functions update and delete are related to scheduling objects, which are made current by previous find functions. When processors or operations are deleted, all related decisions and constraints are removed from the schedule as a side effect.

The function insert is rather obvious for processors and operations. To insert a constraint instance the relevant constraint type should be specified by the user. Before inserting a decision, first a processor, an operation and an decision should be made current by appropriate find functions. The new decision is inserted chronologically after the current decision and is related to the current processor and the current operation.

With respect to schedules we have the following primitive functions:

load, save, show, commit

With the function show the schedule will be represented on the screen. With the function load and save schedules can be loaded from or written to files, specified by the user.
By executing the function `commit` constraint instances will be evaluated and all previous user actions may be rolled back if any hard constraint is violated (see above).

It is important to observe, that the system can be used to manipulate decisions as well as processors, operations and constraints dynamically. The distinction between strategic, tactical and operational planning [4] is less strict in this context.

6. Architecture and implementation

6.1. Architecture

The architecture of a specific DSS is defined by an hierarchy of modules. On the top level we have generic modules. In these modules the superclasses of all scheduling objects (resource, decision, constraint, etc.) including their operators are defined. Only those attributes and methods that are common for all subclasses are defined here. Also time handling, schedule evaluation and feasibility checking belong to this level.

On a lower level we have domain-specific modules, containing the definitions of all relevant sub-classes of scheduling objects, with their specific attributes and operators. For example, in the job-shop case we can distinguish various constraints, like precedence-, deadline- and release time constraints. Also problem specific optimization routines could be defined in these modules eventually.

Constraint types are defined in separate modules. This makes it possible to change and add constraint-types easily. As a consequence constraint types can be used in several planning situations.

A problem-specific DSS is generated by linking the generic modules with the relevant domain-specific modules.

Finally object-instances are supplied by the end user by means of the primitive actions, described in the last section. It is important to observe, that the manipulation of object instances by the end-user is related to decisions as well as to resources, operations and constraints.

The classification of “stakeholders” [20] involved in the development and the use of DSS in three groups (toolsmith, designer and manager) corresponds with the different levels of abstraction in the DSS-architecture: generic modules (superclasses), domain specific modules (subclasses) and object instances. Each group is on exactly one abstraction level involved in the development or the use of the system.

6.2. Implementation

A set of modules is implemented in C++ resulting in decision support systems for several planning problems [18]. These systems run on UNIX and have the basic functionality as described in the last section.

To have a simple measure of the reusability of a certain DSS, we take the ratio of generic software within the whole system. As module size we use the number of lines of pure source code. Table 6.1 shows reusability factors of systems for different planning situations, depending on the number of implemented constraints (indicated by #c). The numbers are smoothed to eliminate the influence of the order of the constraints. Furthermore we consider all modules as specific, that can be used in several but not
all planning situations. Especially this applies to modules describing constraint-types. Though matters are rather simplified, the table gives a rough idea, in how far our approach can lead to software reusability and code reduction.

7. Object definitions

In this section we give some relevant object definitions based on our model by means of a pseudo-language. We will define both generic and domain specific object classes with respect to the job-shop problem. The definitions are simplified for the sake of clearness.

7.1. Object representation

To describe our objects and to write pseudocode we use in this paper a fictive pascal-like object-oriented pseudolanguage. To define classes we have in this language structures called objects. Objects are comparable with records, with the difference that object-components can be both attributes and object-methods (functions or procedures):

```pascal
object figure
    size: integer;
    colour: colourtype;
    method moveto(x, y: integer);
    method draw
end;
```
To define subclasses we can construct hierarchies of objects. An object is denoted as descendant of an other object in a hierarchy by specifying the name of the ancestor in parentheses (after the object name); attributes and methods of the higher class are inherited:

```
object triangle(figure)
```

If in the declaration of an object the name of a method is followed by the keyword virtual, this means that this method can be used in different subclasses under the same name, but with specific functionality (polymorphism). The semantics of such a method are not determined until run-time, when the method is invoked (dynamic binding).

We use the pseudo language structure list of to specify sequences of elements of a certain type

```
real_sequence = list of real
```

Moreover, we assume the following type as predefined

```
interval = record
      left, right: real;
   end;
```

### 7.2. Generic object classes

In this section we define the generic object classes:

```
schedule = object
   sch.processors : list of processor;
   sch.operations : list of operation;
   sch.decisions : list of decision;
   sch.constraints: list of constraint;
   method schedule.evaluate: boolean;
   method schedule.show;
end;
```

```
scheduling._object = object
   s.id : integer; (*object identifier*)
   method s.insert (s: schedule); virtual;
   method s.delete (s: schedule); virtual;
   method s.retrieve (s: schedule); virtual;
   method s.update (s: schedule); virtual;
end;
```

```
processor = object (scheduling._object)
   p.initstate: state;
end;
```

```
operation = object (scheduling._object)
end;
```
One of the most important methods is the function `schedule_evaluate` belonging to the object-class `schedule`, which checks the feasibility of a schedule by evaluating all its constraints. Constraint evaluation on its turn is performed by the method `c.evaluate` belonging to the object-class `constraint`. The semantics of the latter method are not determined until run-time, when the method is invoked and the constraint is initialized. This property is called *dynamic binding* and it is above all this property, which makes it possible to treat constraints so generically.
7.3. Domain specific object classes (job-shop)

In this subsection we define the domain specific object classes for the job-shop problem.

ability = (al,...,an);

rcps_processor = object (processor)
    p.abilset : set of ability;
    p.speed : real;
end;

rcps_operation = object (operation)
    o.abilset : set of ability;
    o.size : real;
end;

rcps_real_decision = object (real_decision)
    d.duration: real; (*duration of assignment*)
    method d.getduration: real; virtual;
end;

rcps_idle_decision = object (idle_decision)
    d.waiting: real; (*waiting time*)
    method d.getduration: real; virtual;
end;

rcps_state = object (state)
    busytime: real;
    method state.transition (dec: decision); virtual;
    method state.show; virtual;
end;

deadline_constraint = object (constraint)
    c.op : rcps_operation;
    c.deadline : real;
    method c.evaluate (s: schedule); virtual;
end;

precedence_constraint = object (constraint)
    c.op1, c.op2 : rcps_operation;
    c.waiting : real
    method c.evaluate (s: schedule); virtual;
end;
availability_constraint = object (constraint)
c_pr : rcps_processor;
c_intervals : list of interval;
method c_evaluate (s: schedule); virtual;
end;

8. Conclusions and extensions

Although scheduling problems vary a lot with respect to their optimalization aspect, they have much more in common, when we restrict ourselves to stepwise planning. When designing a decision support system in a more generic way, we should start therefore with components for pure schedule management. Algorithmic components may be implemented later in layers on top of the basic software. This may be the case both for flexible generic optimization procedures (like simulated annealing) as for powerful problem specific algorithms.

In view of the natural hierarchy of scheduling problems, an object oriented implementation seems to be suitable. It leads to a clear separation between the generic and the domain specific components of the schedule manager, minimizing redundant code and resulting in software with a high degree of maintainability and modularity.

In this paper we made a simplification by assuming only two levels of abstraction. The classification tree of scheduling problems has in fact more levels, which should reflect in the system architecture. Software modules should correspond always to a specific node in that tree. Another simplification is the assumed homogeneity of resources, operations and decisions, which may be not always realistic.

We have to emphasize again, that the optimality of schedules is not an objective in this paper. Important is the optimality of the software in the sense of maintainability, flexibility and reusability.

References


In this series appeared:

89/1 E.Zs.Lepoeter-Molnar
Reconstruction of a 3-D surface from its normal vectors.

89/2 R.H. Mak
P.Struik
A systolic design for dynamic programming.

89/3 H.M.M. Ten Eikelder
C. Hemerik
Some category theoretical properties related to a model for a polymorphic lambda-calculus.

89/4 J.Zwiers
W.P. de Roever
Compositionality and modularity in process specification and design: A trace-state based approach.

89/5 Wei Chen
T.Verhoeff
J.T.Udding
Networks of Communicating Processes and their (De-)Composition.

89/6 T.Verhoeff
Characterizations of Delay-Insensitive Communication Protocols.

89/7 P.Struik
A systematic design of a parallel program for Dirichlet convolution.

89/8 E.H.L.Aarts
A.E.Eiben
K.M. van Hee
A general theory of genetic algorithms.

89/9 K.M. van Hee
P.M.P. Rambags
Discrete event systems: Dynamic versus static topology.

89/10 S.Ramesh
A new efficient implementation of CSP with output guards.

89/11 S.Ramesh
Algebraic specification and implementation of infinite processes.

89/12 A.T.M.Aerts
K.M. van Hee
A concise formal framework for data modeling.

89/13 A.T.M.Aerts
K.M. van Hee
M.W.H. Hesen
A program generator for simulated annealing problems.

89/14 H.C.Haesen
ELDA, data manipulatie taal.

89/15 J.S.C.P. van der Woude
Optimal segmentations.

89/16 A.T.M.Aerts
K.M. van Hee
Towards a framework for comparing data models.

89/17 M.J. van Diepen
K.M. van Hee
A formal semantics for Z and the link between Z and the relational algebra.
<table>
<thead>
<tr>
<th>Page</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>90/2</td>
<td>K.M. van Hee P.M.P. Rambags</td>
<td>Dynamic process creation in high-level Petri nets, pp. 19.</td>
</tr>
<tr>
<td>90/4</td>
<td>A. Peeters</td>
<td>Decomposition of delay-insensitive circuits, p. 25.</td>
</tr>
<tr>
<td>90/5</td>
<td>J.A. Brzozowski J.C. Ebergen</td>
<td>On the delay-sensitivity of gate networks, p. 23.</td>
</tr>
<tr>
<td>90/8</td>
<td>M.B. Josephs</td>
<td>Receptive Process Theory, p. 16.</td>
</tr>
<tr>
<td>90/9</td>
<td>A.T.M. Aerts P.M.E. De Bra K.M. van Hee</td>
<td>Combining the functional and the relational model, p. 15.</td>
</tr>
<tr>
<td>90/10</td>
<td>M.J. van Diepen K.M. van Hee</td>
<td>A formal semantics for Z and the link between Z and the relational algebra, p. 30. (Revised version of CSNotes 89/17).</td>
</tr>
<tr>
<td>90/11</td>
<td>P. America F.S. de Boer</td>
<td>A proof system for process creation, p. 84.</td>
</tr>
<tr>
<td>90/12</td>
<td>P.America F.S. de Boer</td>
<td>A proof theory for a sequential version of POOL, p. 110.</td>
</tr>
<tr>
<td>90/14</td>
<td>F.S. de Boer</td>
<td>A proof system for the language POOL, p. 70.</td>
</tr>
<tr>
<td>90/15</td>
<td>F.S. de Boer</td>
<td>Compositionality in the temporal logic of concurrent systems, p. 17.</td>
</tr>
<tr>
<td>90/16</td>
<td>F.S. de Boer C. Palamidessi</td>
<td>A fully abstract model for concurrent logic languages, p. 23.</td>
</tr>
<tr>
<td>90/17</td>
<td>F.S. de Boer C. Palamidessi</td>
<td>On the asynchronous nature of communication in logic languages: a fully abstract model based on sequences, p. 29.</td>
</tr>
</tbody>
</table>
Design and implementation aspects of remote procedure calls, p. 15.

Two Case Studies in ExSpect, p. 24.

The Nature of Delay-Insensitive Computing, p. 18.

Data, Process and Behaviour Modelling in an integrated specification framework, p. 37.


Implication. A survey of the different logical analyses "if...then...", p. 26.

Parallel Programs for the Recognition of P-invariant Segments, p. 16.

Performance Analysis of VLSI Programs, p. 31.

An Implementation Model for GOOD, p. 18.

SPECIFICATIEMETHODEN, een overzicht, p. 20.

CPO-models for second order lambda calculus with recursive types and subtyping, p.

Terminology and Paradigms for Fault Tolerance, p. 25.

Interval Timed Petri Nets and their analysis, p. 53.

POLYNOMIAL RELATORS, p. 52.

Relational Catamorphism, p. 31.


A note on Extensionality, p. 21.

The PDB Hypermedia Package. Why and how it was built, p. 63.
<table>
<thead>
<tr>
<th>Paper ID</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>91/16</td>
<td>A.J.J.M. Marcelis</td>
<td>An example of proving attribute grammars correct: the representation of arithmetical expressions by DAGs, p. 25.</td>
</tr>
<tr>
<td>91/18</td>
<td>Rik van Geldrop</td>
<td>Transformational Query Solving, p. 35.</td>
</tr>
<tr>
<td>91/19</td>
<td>Erik Poll</td>
<td>Somé categorical properties for a model for second order lambda calculus with subtyping, p. 21.</td>
</tr>
<tr>
<td>91/23</td>
<td>K.M. van Hee, L.J. Somers, M. Voorhoeve</td>
<td>Z and high level Petri nets, p. 16.</td>
</tr>
<tr>
<td>91/25</td>
<td>P. Zhou, J. Hooman, R. Kuiper</td>
<td>A compositional proof system for real-time systems based on explicit clock temporal logic: soundness and completeness, p. 52.</td>
</tr>
<tr>
<td>91/27</td>
<td>F. de Boer, C. Palamidessi</td>
<td>Embedding as a tool for language comparison: On the CSP hierarchy, p. 17.</td>
</tr>
<tr>
<td>91/28</td>
<td>F. de Boer</td>
<td>A compositional proof system for dynamic process creation, p. 24.</td>
</tr>
</tbody>
</table>