Distributed Patient Scheduling in Hospitals

T. O. Paulussen* — Inf. Sys. Dept., Univ. of Mannheim, Mannheim, Germany, paulussen@uni-mannheim.de

N. R. Jennings — Elect. & Comp. Sci. Dept., Univ. of Southampton, Southampton, UK, nrj@ecs.soton.ac.uk

K. S. Decker — Comp. & Inf. Sci. Dept., Univ. of Delaware, Newark, Delaware, USA, decker@cis.udel.edu

A. Heinzl* — Inf. Sys. Dept., University of Mannheim, Mannheim, Germany, heinzl@uni-mannheim.de

Abstract

Patient scheduling in hospitals is a highly complex task. Hospitals have a distributed organisational structure; being divided into several autonomous wards and ancillary units. Moreover, the treatment process is dynamic (information about the patients' diseases often varies during treatments, causing changes in the treatment process). Current approaches are insufficient because they either focus only on the single ancillary units, and therefore do not consider the entire treatment process of the patients, or they do not account for the distribution and dynamics of the patient scheduling problem. Therefore, we propose an agent based approach in which the patients and hospital resources are modelled as autonomous agents with their own goals, reflecting the decentralised structures in hospitals. In this multi-agent system, the patient agents compete over the scarce hospital resources. Moreover to improve the overall solution, the agents then negotiate with one another. To this end, a market mechanism is described, in which each self interested agent tries to improve its own situation. In particular we focus on how the agents can calculate demand and supply prices based upon their current schedule. Further, an evaluation of first results of the proposed method is given.

1 Introduction

Patient scheduling in hospitals is concerned with the optimal assignment of patients to hospital resources. Hospitals are divided into several autonomous wards and ancillary units, which are visited by the patients for treatments and examinations during hospitalisation in accordance with their illness [Schliichtermann, 1990]. However, the pathways (the needed treatments and examinations) and the medical priorities (the health condition of the patients) are likely to change due to new findings about the diseases of the patients during examination. Further, complications and arrivals of emergency patients, which are in urgent need for treatment, result in schedule disturbances[1] [Schliichtermann, 1990]. Due to this, patient scheduling in hospitals requires an approach which is distributed, in order to leave the authority at the responsible hospital units, and flexible, to be able to react to plan changes in an efficient manner.

We have chosen to adopt a multi-agent based approach to this problem, because such systems allow the representation of every single coordination object, i.e. the responsible entities, as single autonomous agents with their own goals [Weinhardt and Gomber, 1996]. This, in turn, reflects the existing decentralised structures in hospitals [Decker and Li, 2000]. Through social abilities, the agents can interact with each other to reach their own goals. Moreover they can react with the needed flexibility to changes (as new information about the health status of a patient becomes available) and disturbances (e.g. emergencies and complications) through proactiveness and responsiveness [Jennings, 2001].

This paper advances the state of the art in two main ways. First is the design of a novel procedure, i.e. a patient centred multi-agent based coordination mechanism, where the patients are represented as autonomous agents, trying to improve their current scheduling state by negotiation with other agents. Second is the derivation of health state dependent opportunity cost functions, based upon which the agents evaluate their current schedule and compute the gains and losses through plan modifications.

The remainder of the paper is structured as follows. Section 2 describes the domain of patient-scheduling in hospitals. Section 3 details the conceptual framework of our multi-agent system, the relevant coordination objects and the coordination mechanism. The results of the proposed mechanism are also evaluated. The paper ends with conclusions and an outlook to further work in section 4.

2 The Patient Scheduling Problem

In addition to the complexity arising from the distributed structure of hospitals, patient scheduling has to be performed in the face of a high degree of uncertainty about the treatment pathways of patients within the hospital. Thus patients arrive

* We would like to thank the *DFG, Deutsche Forschungsgemeinschaft* for funding this work as part of SPP 1083.

[1] This problem can be ameliorated by the provision of additional (exclusive) emergency resources (e.g. separate x-ray facilities). However, for economic reasons, hospitals try to minimise these extra resources.

continuously at the hospital and the necessary medical treatments are often not able to be completely determined at the beginning of the treatment process. Moreover the results of a diagnostic examination might change the (medical) priority of the patients, invoke additional activities and/or make other medical actions obsolete.

Due to this complexity, the application of traditional (operations research) methods from industrial scheduling to the patient scheduling problem is problematic [SchlCichtermann, 1990]. To be able to handle the process dynamics in a distributed environment, hospitals commonly use a very flexible approach for patient scheduling. Typically, the wards send treatment and examination requests to the ancillary units. Based upon these requests, the ancillary units order the patients from the wards. This allows the units to react very flexibly to changes with very low communication needs. If, for example, an emergency patient needs to be inserted, the next patient will simply be called from the ward later, leaving this patient available to other ancillary units.

However, because there is no inter-unit coordination, this procedure cannot resolve resource conflicts, which occur if the same patient is requested by more than one ancillary unit at the same time [Decker and Li, 2000]. Because the ancillary units only have a local view, i.e. they do not - and cannot - take the whole pathway of the treated patients into their scheduling consideration, no inter-unit process optimisation can be undertaken (i.e. the medical tasks for the patients cannot be scheduled and coordinated in an efficient manner). This causes undesired idle times as well as overtime hours for the hospital resources and extended patient stay times.

3 Agent-Based Patient Scheduling

In this section the conceptual framework of our multi-agent system is described. Then, the relevant coordination objects, i.e. the patients and hospital resources, are modelled as agents. After this, a coordination mechanism is introduced. Finally, examples are given and the first results of the implementation are presented.

3.1 Conceptual Framework

Using a multi-agent system, [Decker and Li, 2000] addressed the problem of resource conflicts in the patient scheduling domain. Specifically, by adding a coordination mechanism to handle mutually exclusive resources (i.e. the patients) to the Generalized Partial Global Planning (GPGP) approach[2], they achieved significant scheduling improvements compared to current patient scheduling without coordination. In their work, they took a function-oriented approach, where only the units are represented as autonomous agents. The treatment pathway of the patients is captured by the nursing units. However, the dynamics of the patient scheduling problem and the medical priorities of the patients were not considered.

For this reason, we use a patient-centred approach, in which the patients are modelled as agents, too. This allows the representation - and therefore consideration - of the hospital processes as a whole [Adam and Gorschlüter, 1999]. In contrast to the resource agents, who only see the patients as entities to be treated, the patient agents only see the medical actions as tasks that need to be performed. Due to these opposing forces, the patient agents ensure that the resource agents also consider the treatments of the patients outside their unit (without any explicit knowledge of them) and vice versa.

To this end, the relevant coordination objects in the hospital (i.e. the entities to be coordinated), which will be represented as agents, can be identified as patients and resources (i.e. rooms, machines and personnel). To reduce complexity, all resources needed for a specific medical action are represented by the hospital unit responsible for this action, which will be implemented as the resource agents.

If more than one patient agent now wants to use the identical resource at the same time interval, a resource conflict occurs. To solve these conflicts, a coordination mechanism is needed. For this inter-agent coordination, we decided to use a market mechanism. The rationale for this is that a market mechanism is a distributed coordination mechanism which facilitates efficient solutions with low communication needs. In particular, only prices for specific goods are communicated, keeping all other information private to the market participants [Wellman *et al.,* 2001]. Additionally, a market facilitates a dynamic environment, where the market participants take actions according to their current (dynamically changing) situation based upon private information and preferences. In this market, the coordination objects are modelled as autonomous, self-interested market participants, trying to improve their local schedule. In markets, the price mechanism leads to an efficient resource allocation because the resources are assigned to the agents that are willing to pay the highest price (assuming that the agents bid rationally, these are the agents who gain the highest utility from this resource). A pareto optimal solution will be achieved, because no agent will accept a deal, which worsens its current state. Therefore, the agents will trade resources until no agent can improve its schedule without harming another agent [Weigelt, 1994].

3.2 Coordination Objects

To implement a coordination mechanism, the above mentioned coordination objects have to be modelled. In this work the main goal for the patients is to minimise their stay time and for the resources to minimise idle times. This kind of scheduling problems represents a worth-oriented environment [Rosenschein and Zlotkin, 1994], where the degree of goal achievement can be evaluated by a utility function (rather than a state or task oriented environment, where a single goal cannot be achieved partially). The usage of continuous worth functions instead of single worth values assigned to specific goals enables the agents to relax their goals, i.e. to compromise in order to achieve a better solution [Rosenschein and Zlotkin, 1994]. For example, if two agents want the same time slots at the same resource, the agents can agree upon a solution which does not fully satisfy their own goals, but reaches a better overall solution. This situation is illustrated in the figures 2(a) and 2(b).

Because we adopt a patient-centred approach, the main focus will be on the patient agents, which will be modelled in

[2] GPGP is a domain independent, task environment centred, scheduling approach to coordination [Decker and Li, 2000].

the first subsection. Within this part, we introduce the health state dependent cost-functions, which are needed for price articulation in the market mechanism. In the second subsection, the resource agents are described.

Patient Agents

Each patient is represented by an agent. The patient agents can only see their own schedule, containing their pathway through the hospital. The pathway of a patient agent comprises the needed treatments and examinations as well as the order constraints between those tasks. The patient agents are also equipped with a private worth function to be able to negotiate in a goal driven manner. This worth function is realised as a cost function, which they try to minimise. These cost functions increase over time, setting incentives for the patient agents to schedule their treatments and examinations as early as possible in order to reduce costs.

In contrast to commercial domains, the utility or cost functions cannot be based upon monetary values in hospitals. While in e-commerce scenarios the human principals reveal their preferences through their willingness to pay (e.g. by specifying maximum buy and minimum sell prices), patients do not - and should not - reveal thei; preferences through their willingness to pay for a specific treatment time slot. The preferences rather have to be based upon medical priorities, i.e. their health state. This absence of normalised monetary units additionally causes the problem of inter-agent utility comparison. To solve this problem, we operationalised the progress of the patients' health state. In current hospital practice, numerous health states or patient priority measures and indices are in use (e.g. APACHE IT (Acute Physiology And Chronic Health Evaluation) [Knaus *et al.,* 1985] in intensive care units). The choice of the measurement to use is up to the hospital. However, we propose, that it should support a cardinal measurement of the health state progress over time.

The *cardinal measurement* is necessary for inter-agent utility comparison and transfer (as we have to calculate with these units). Because these health state (utility) units cannot be transferred between the actual patients in the hospital, the multi-agent system has to be separated from the actual patient world, where only the agents transfer utility units in order to reach a better overall solution (minimisation of health state adjusted patient stay time).

The *health state progress* is important, because the priority of the patients should not be based upon their current state, but on their health state development. For example, a patient with a currently reasonable but rapidly deteriorating health condition should have a higher priority than a patient with a (slightly) lower current, but continuously constant, health state.

For the necessary cardinal measurement of health, we rely on the concept *of years of well being* [Torrance, 1987] because it handles the health state progress over time (a good overview can be found in [Pedroni and Zweifel, 1990]). In this method, the question is what time period xT of total health (1) equals one specific time period IT of a certain health state H, i.e.

$$1T * H = xT * 1 \Leftrightarrow H = x.$$

Through this, the health state of a patient can be described in time units.

The primary goal of patients in hospitals is to increase their current health state through treatment, where a disease could be viewed as disutility (decrease in quality of life) for the patient. This loss of utility adds up as long as this disease is not cured. Based on this assumption, the (opportunity) cost C^{opp} for not curing the patient right away equals the difference between the achievable health state (through treatment) z and the patient's health state over the time $H(t)$. Formally, this can be expressed by

$$C^{opp}(t) = \int_0^t z - H(t)dt.$$

In addition, it has to be considered that the health state and the achievable health state can change over time. Therefore, the patient's opportunity costs are influenced by his current health state a, the development of his health state over time $H(t)$, and the maximal reachable health state through treatment z. If the health state does not change over time, i.e. $H(t)$ = a, the opportunity costs are

$$C^{opp}(t) = \mu t; \mu = z - a.$$

If the health state of a patient worsens over time, assumptions about the course of the health state have to be made by a physician. If we assume - for clarity - a linear reduction by b of the health state, i.e. $H(t) = a - bt$, we get

$$C^{opp}(t) = zt - at + \frac{b}{2}t^2 = \mu t + \frac{b}{2}t^2.$$

However, this approach works with any health decrease rate, as the health state of a patient normally does not decrease linearly. Nevertheless, a linear approximation of the decrease in health can be justified by practical reasons. Instead of trying to estimate the exact shape of the curve, a physician could rather specify two or more specific points in time and in between a linear reduction could be estimated. Figure 1 shows an exemplary course of an illness with linear reduction of the health state, resulting in a quadratic opportunity cost curve.

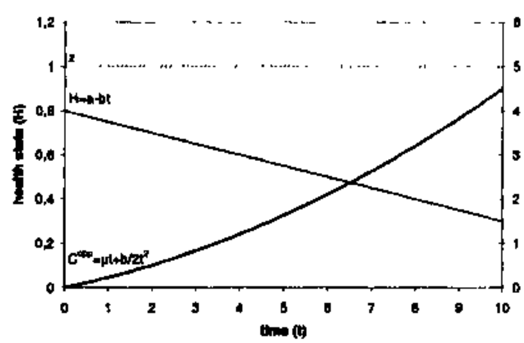


Figure 1: Linear reduction of health state.

Finally, if the maximal reachable health state decreases, the patient must be treated immediately to prevent lifelong damage to his health.

With these cost functions in place, the agents can now evaluate their current state, i.e. the degree of goal achievement. Further, based on these functions, the agents are able to compute supply and demand prices for time slots which correspond to the losses or benefits caused by plan changes, which is essential for a market mechanism. How these prices are computed will be explained later in the context of the description of the coordination mechanism (section 3.3).

Resource Agents

In contrast to the patients, the hospital resources are directly comparable with the resources in industrial scheduling domains. Their main goal can be described as maximising their utilisation or, equivalently, minimising idle time.

To reach this goal, cost-functions can also be articulated for the resource agents. These cost functions represent the reserve prices (i.e. the price that will be charged for an empty slot) for the possible appointments [Wellman *et al,* 2001]. Through this, priorities between different resources can be established that allow penalisation of undesired appointment times (e.g. evening shift or overtime hours). Here, the basis for the resource agents' cost functions comes from cost accounting.

However, for inter-agent utility comparison (i.e. between patient and resource agents) these measurements have to be equalised, which can be a very difficult task, because the health state of a patient has to be compared with monetary values from cost accounting. A good way out of this problem are *trade-off* considerations, e.g. what amount of idle time of a specific resource equals one hour waiting time for a patient with a specific health state.

3.3 The Coordination Mechanism

The main goal of the coordination mechanism is to minimise the health state adjusted stay time of the patients, which is equivalent to an overall minimisation of suffering for the patients. The basic idea of our coordination mechanism is that the patient agents try to buy into resource time slots for the needed treatments and examinations. However, the usage of (central) auction mechanisms is obstructed by the dynamics of the patient scheduling problem described in section 2.

To ensure feasible (i.e. conflict free) initial task appointments for the patients, all new treatments and examinations are scheduled on a *first-come first-served* (fefs) basis. To do this, the patient agents who want to add a task to their pathway contact the responsible unit for the execution of this task in order to obtain the earliest time slot which is available at the unit as well as in their own schedule. To illustrate this, figure 2(a) shows a possible initial resource allocation. These initial appointments determine the budget (or better the initial opportunity costs) of the agents. This is important because a hypothetical price system is used (as per section 3.2).

Based upon this initial schedule, the agents try to improve their schedule in order to reduce their opportunity costs. The price p they are willing to pay for a specific time slot (expected gain) or they charge for a time slot (expected loss) is the difference between the cost-value of the current allocation t^{old} and the cost-value for the wanted appointment t^{new}, according to their individual cost function described in the previous section, i.e.

$$p(t^{old}, t^{new}) = C^{opp}(t^{old}) - C^{opp}(t^{new}).$$

Because in this approach the (opportunity) costs of an appointment increase over time for the patient agents, they must try to schedule their treatments and examinations as early as possible. If a demanded time slot is already occupied by another patient agent, the initial demander must try to buy the time slot from the current owner. With respect to the properties of a market mechanism, the agents act in a rational, self-interested manner. Therefore, the owners of the time slots will only release them, if the price offered equals the losses invoked through rescheduling. Since they only charge for the costs invoked through rescheduling they can be viewed as acting in a partially cooperative manner.

The detailed negotiation process goes as follows:

1. A patient agent initiates a negotiation for rescheduling, if the pathway (additional or obsolete medical actions) or the health state of its patient has changed.
2. The initiating agent selects the task with the highest possible improvement (difference between the costs of the current owned and the best reachable time slot) and contacts the resource agent that is responsible for the execution of this task.
3. The resource agent reserves that time slot, and contacts all affected patient agents, i.e. the agents that currently own this interval, and informs them about the proposal of the initiator.
4. The affected patient agents (sellers) try to reschedule to the first nonreserved time slots (see step 3) and notify the initiator about their costs due to rescheduling. To prevent cycles, reserved intervals cannot be demanded by other agents.
5. If the alternative time slots for the sellers are already occupied, they again become demanders for those time slots and accumulate the invoked costs. Here, order constraints can invoke additional rescheduling in other resources.
6. After all prices are computed and submitted to the initiator, the initiator compares its expected gains from rescheduling to the total price asked for this interval. If the gains exceed the costs it accepts or rejects otherwise, and the negotiation for this time slot terminates.

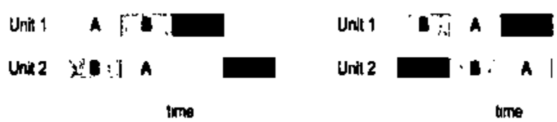
The former initiator continues its rescheduling activities by opening new negotiations for the next task with the (now) highest possible improvement until it cannot improve any task any further. Previously rejected time slots will not be considered unless these time slots are released by their owners. For concurrency issues only one (randomly chosen) agent can initiate a negotiation at a time.

In this coordination mechanism only the patient agents are active components. However, as described earlier, cost functions can be implemented into the resource agents, allowing prioritisation between the resources. The resource price for a time slot is charged when a patient agent buys a time slot, and reimbursed when the patient agent releases this time slot.

As described earlier, the pathway of the patients through the hospital is likely to change during the treatment process, that is, additional treatments and/or examinations may become necessary while other tasks may become obsolete. In bur coordination mechanism, additional tasks (as well as additional patients) can be added at any time using the fcfs rule explained above. If a treatment becomes obsolete, the responsible patient agent notifies the affected resource agent that, again, informs the other patient agents in its schedule.

3.4 Example Scenarios

Figure 2(a) and 2(b) illustrate the functioning of our coordination mechanism. Figure 2(a) shows an initial schedule with three patients (A,B,C). For illustration purposes, we start with identical cost-functions with an assumed initial health state *a* of 0.7, an achievable health state *z* of 1.0, a decrease rate *b* of 0.001 and equal possible starting times for all patients. The used cost function resolves as *0.3t* + 0.0005* . Further, in this example, all task durations are set to lOt. We will use the $-sign to indicate utility units.



(a) Schedule before negotiation. (b) Schedule after negotiation.

Figure 2: Example Schedule before and after negotiation.

Based upon this initial schedule, agent C negotiates with agent B in unit 2 in order to improve its current situation. Agent C's current opportunity costs in unit 2 are $12.8 (0.3 x 40 + 0.0005 x 4(r). Its opportunity costs for the first time slot would be $3.05 (i.e. a difference of $9.75). For the calculation of the offer price, agent B has to determine its additional rescheduling costs. Therefore agent B has to negotiate with agent A in unit 2 for the next best time slot and in unit I, because the patients are exclusive resources which can only perform one task at a time. Agent A looses $3.25 in unit 2 and another $3.15 in unit 1. Agent B looses $3.15 in unit 2 but gains $3.15 in unit 1. Agent B is charged with the losses of Agent A ($6.40) and adds this to its supply price for agent C (here $3.15 − $3.15 = $0.00). Because the total supply price ($6.40) is less than agent C 's gains ($9.75) agent C accepts the deal (because it can compensate the losses of the other agents) and the plan is changed. No more deals are possible and figure 2(b) shows the result of this negotiation.

In this example, patient agent A accepted a schedule of lower quality in favour of patient agent C. However, only patient agent A - not the real patient - has received a compensation from patient agent C. This illustrates the necessary disjunction between the real patients and the multi-agent system in the hospital domain as described earlier (section 3.2).

In the next step, we relax the assumption, that all agents have the same cost functions. If agent C would have had a lower health state (e.g. μ = 0.35) it could have even improved its appointment in unit 1 because it would have gained $3.75 from moving to the second slot (agent A in figure 2(b)) while agent A would only ask $3.25 for its slot (agent A moves to slot 1 and gains $3.15, and agent B moves to slot 3 and looses $6.40).

We have applied the described coordination mechanism to the first Taillard 5 × 5 open shop problem [Taillard, 1993] to analyse the behaviour for tasks with different durations. Figure 3 shows the resulting Gantt chart with a solution of 349. The optimal solution is 300. To analyse the convergence of the mechanism, we logged the changes of the latest task for each agent. Figure 4 shows the corresponding convergence behaviour, which indicate a fast convergence in the first third of the graph.

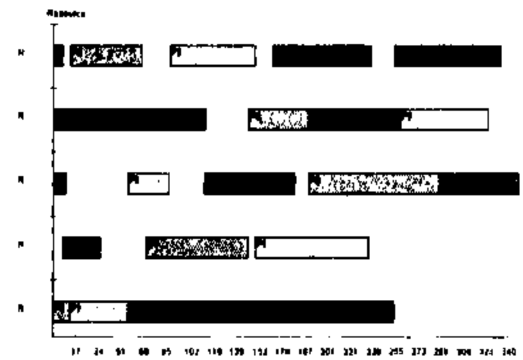


Figure 3: Gantt of the results of the Taillard 5x5 open shop problem.

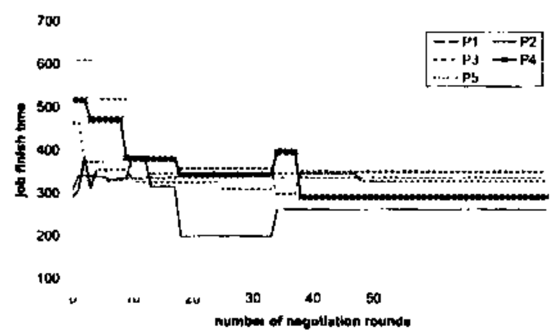


Figure 4: Converge of the Taillard 5x5 open shop problem.

The *number of negotiation rounds* counts each improvement attempt (successful or not) of the agents. The same procedure was executed for the $n \times n$ open shop problems with *equal task durations.* For these problems the optimal solution will be always achieved. The resulting Gantt chart is shown in figure 5. Figure 6 shows the corresponding convergence, illustrating the simplicity of this problem type in contrast to the used Taillard benchmark.

However, a better agent-based solution to this problem can be achieved through the usage of (central) appointment auctions. Implementing an auction based approach using the same cost functions resulted in a solution at 308. However, these auctions cannot handle the dynamics of the patient scheduling problem, because the auction process would have

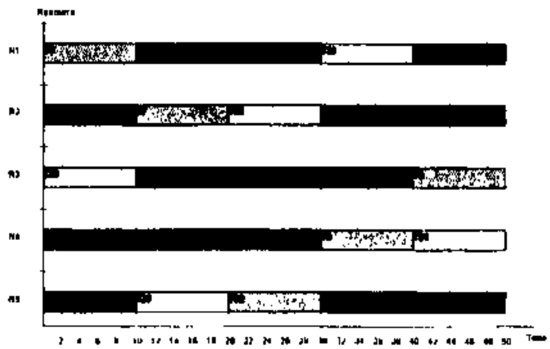


Figure 5: Gantt of the results of the easy 5x5 open shop problem.

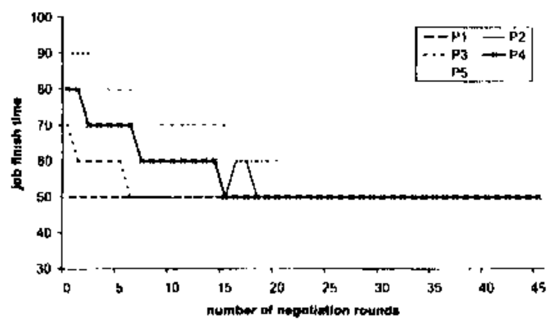


Figure 6: Converge of the easy 5x5 open shop problem.

to restart again after severe changes in the pathway occurred.

4 Conclusions and Future Work

In this paper, we presented a patient-centred coordination mechanism for inter-unit patient scheduling in hospitals. Therefore, health state dependent cost functions were developed. We described the implemented coordination mechanism and evaluated first results, showing the different convergence for problems with equal and different task durations.

In the next step, we will run our approach on real hospital data, already retrieved out of a field study in five hospitals. Based upon these results, our coordination mechanism will be enhanced in future work, where the main focus is on the development of a flexible auction mechanism in order to handle the dynamics of the patient scheduling problem. In this context, we will analyse how the usage of texture measurements (here: criticality and goodness measures) [Sycara *et al,* 1990] can facilitate our coordination mechanism in order to reduce backtracking. Additionally, our agent based approach will be benchmarked with a genetic and evolutionary algorithm, which is under development in the context of this work. Further, a decision theoretic approach facilitating stochastic task duration will be presented in the near future.

References

[Adam and Gorschluter, 1999] D. Adam and P. Gorschlutcr. Qualitatsmanagement im Krankenhaus. *Zeitschriftfur Betriebswirtschaftslehre (ZfB), Erganzungsheft 5: Krankenhausmanagement,* 1999.

[Decker and Li, 2000] K. Decker and J. Li. Coordinating mutually exclusive resources using GPGP. *Autonomous Agents and Multi-Agent Systems,* 3(2): 133-157,2000.

[Jennings, 2001] N. R. Jennings. An agent-based approach for building complex software systems. *Communications of the ACM,* 44(4):35-41, April 2001.

[Knaus *et ai,* 1985] W. A. Knaus, E. A. Draper, D. P. Wagner, and J. E. Zimmermann. Apache ii: a severity of disease classification system. *Critical Care Medicine,* 13(10):818~829,1985.

[Pedroni and Zweifel, 1990] G. Pedroni and P. Zweifel. *Wie misst man Gesundheit?* Pharma-Information, Basel, 1990.

LRosenschein and Zlotkin, 1994] J. S. Rosenschein and G. Zlotkin. *Rules of Encounter.* MIT Press, Cambridge, MA, 1994.

[Schluchtermann, 1990] J. Schluchtcrmann. *Patientensteuerung.* Verlag Josef Eul, Bergisch Gladbach, Germany, 1990.

[Sycara *et al,* 1990] K. Sycara, S. Roth, N. Sadeh, and M. S. Fox. An investigation into distributed constraint-directed factory scheduling. In *Proceedings of the Sixth IEEE Conference on AI Applications (CAIA-90),* Santa Barbara, CA, 1990.

[Taillard, 1993] E. Taillard. Benchmarks for basic scheduling problems. *European Journal of Operations Research,* 64:278-285, 1993.

[Torrance, 1987] G. W. Torrance. Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases,* 1987.

[Weigelt, 1994] M. Weigelt. *Dezentrale Produktionssteuerung mit Agenten-Systemen: Entwicklung neuer Verfahren und Vergleich mit zentraler Lenkung.* Dt. Univ.-Verl., Wiesbaden, 1994.

[Weinhardt and Gomber, 1996] Ch. Weinhardt and P. Gomber. Domanenunabhangigc Koordinationsmechanismen fur die dezentrale betriebliche Planung. *Information Management,* pages 6 16, 1996.

[Wellmane/a/.,2001] M. Wellman, W. Walsh, P. Wurman, and J. MacKie-Mason. Auction protocols for decentralized scheduling. *Games and Economic Behavior,* 35:271-303,2001.