Agent-Based Patient Scheduling in Hospitals

Torsten O. Paulussen, Anja Zöller, Franziska Rothlauf, Armin Heinzl

University of Mannheim, Department of Business Administration and Information Systems, {paulussen | zoeller | rothlauf | heinzl}@bwl.uni-mannheim.de

Lars Braubach, Alexander Pokahr, Winfried Lamersdorf

University of Hamburg, Department of Computer Science, Distributed and Information Systems, {braubach | pokahr | lamersd}@informatik.uni-hamburg.de

Abstract. Patient scheduling in hospitals is a very complex task. This complexity stems from the distributed structure of hospitals and the dynamics of the treatment process. Hospitals consist of various autonomous, administratively distinct units which are visited by the patients according to their individual disease. However, the pathways (the needed medical actions) 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. Moreover, the durations of the treatments and examinations are stochastic. Additional problems for patient scheduling in hospitals arise from complications and emergencies. Thus, patient scheduling in hospitals requires a distributed and flexible approach. To this end, a flexible, agent-based approach to patient scheduling is developed in this chapter. After a description of the addressed patient scheduling problem, the proposed mechanism for patient-scheduling is presented and evaluated.

1 Introduction

In this chapter, an agent-based coordination mechanism for patient scheduling in hospitals is described. Patient scheduling is concerned with the optimal assignment of the patients to the scarce hospital resources [Schl1990], where the goal of the patients is to minimize their stay time and the goal of the resources is to minimize their idle time. However, patient scheduling in hospitals resolves as a very complex task. Hospitals consist of several autonomous, administratively distinct wards and ancillary units [DeLi2000] [KuOP1993], which are visited by the patients for their treatments and examinations in accordance with their illness [Schl1990]. However, the pathways (the needed medical actions) 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

[PJDH2003]. Further, the durations of the treatments and examinations are stochastic [BrWo1991] [RoCh2003]. Additional problems for patient scheduling in hospitals arise from complications and emergencies, which are in urgent need for treatment [PJDH2003]. Due to this, patient scheduling in hospitals requires an approach which is distributed, in order to leave the authority at the responsible units, and flexible, to be able to react to new information in a timely manner.

For this reason, a multi-agent based approach was chosen for this problem, because it allows the representation of every coordination object as a single autonomous agent with own goals [WeGo1996]. Further, the agents can react with the needed flexibility to changes (as new information about the health status of a patient becomes available) and disturbances (emergencies and complications) through proactiveness and responsiveness [Jenn2001]. In this context, the notion of flexibility refers to the term “technical flexibility”, that is, the ability to react adequately to external influences (see Part I, Chapter 3).

The remainder of this chapter is structured as follows. Section two elaborates the patient scheduling problem in hospitals. Based upon this, the conceptual framework of the proposed multi-agent system is developed in the third section. In the forth section, a prototypic implementation of the coordination mechanism is evaluated and benchmarked against the status quo of patient scheduling in hospitals. This chapter closes with conclusions and an outlook to future work in the fifth section.

2 The patient scheduling problem in hospitals

Hospitals are service providers with the primary aim to improve the health state of their patients, where the treatment of the patients is the main value adding process in hospitals [Fein1999] [GrTT997]. Hospitals consist of several autonomous, administratively distinct wards and ancillary units [DeLi2000] [KuOP1993] [PJDH2003]. During hospitalization, the patients reside at the wards and visit the ancillary units for treatments according to their individual disease, where the *treatment process* comprises the medical tasks which must be performed for the patients during hospitalization.

The service provision in a hospital can be viewed from a patient (or job) perspective and from a resource perspective. While the patients focus on the sequence of their medical tasks with the goal to minimize their stay time, the resources focus on the treatments and examinations within the resources with the goal to minimize their idle times [DeLi2000] [KuOP1993].

The *patient scheduling* is now concerned with the (optimal) temporal assignment of the medical tasks for the patients to the (scarce) hospital resources [Schl1990]. However, the patient scheduling problem in hospitals is confronted with a high degree of uncertainty. The patients arrive continuously at the hospital and the necessary medical treatments are often not completely determined at the beginning of the treatment process. Moreover, the new findings during diagnostic examinations might change the (medical) priority of the patients, invoke additional treatments or examinations, and make other medical actions obsolete [PJDH2003]. Further, the service times of treatments and examinations are stochastic [BrWo1991] [RoCh2003]. Finally, complications and arrivals of emergency patients – which are in urgent need for treatment – result in schedule disturbances.

To be able to handle the process dynamics in a distributed environment, hospitals commonly use a very flexible approach for patient scheduling which can be compared to a first-come first-served priority rule. Typically, a ward physician prescribes the necessary treatments and examinations for the patients. These prescribed medical tasks are send as treatment requests to the ancillary units. Based upon these requests the ancillary units order the patients from the wards when they deem appropriate [DeLi2000] [KuOP1993]. This allows the units to react very flexible 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 [DeLi2000]. Because the ancillary units only have a local view, that is, they do not – and cannot – take the whole pathway of the treated patients into their scheduling consideration, no inter-unit process optimization 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 Conceptual framework

In this section, the conceptual framework of an agent-based coordination mechanism for patient scheduling is developed. As described previously, the patient scheduling problem is concerned with the optimal assignment of the treatments and examinations of the patients to the scarce hospital resources; where the goal of the patients is to minimize their stay time and the objective of the resources is to minimize their idle time. In order to achieve this assignment in a goal-driven manner, the proposed coordination mechanism relies on the economic concept of mutual selection. In this context, the patients and resources can be identified as the coordination objects, which are modeled as autonomous agents; where the patient-agents try to acquire the needed medical services, that is, treatment or examination time slots, from the resource-agents. Because a resource is generally demanded by several patients, each resource-agent auctions off the medical services (time slots) of its hospital resource. In order to participate in the resource auctions, the patient-agents need utility functions, which enable them to determine the values of the required time slots and thus to generate the bids for the time slots.

3.1 Health-state dependent utility functions

In the proposed auction-based coordination mechanism, the patient-agents compete with each other over the scarce hospital resources in order to achieve the objectives of their corresponding patients as good as possible. This kind of coordination problem represents a worth-oriented environment [RoZl1994], in which the degree of goal achievement can be evaluated through a utility function cf. [PJDH2003]. The usage of continuous utility functions (instead of single values assigned to specific goals) allows the coordination objects to relax their goals, that is, to compromise in order to reach a better overall solution cf. [RoZl1994].

In contrast to the domain of electronic commerce or industrial production control, the bid-price for a resource time slot in hospitals can neither be based on the patient's willingness to pay for a time slot nor be derived from cost accounting, respectively. The preferences of the patient-agents rather have to be based upon medical priorities, that is, the health state of the patients [PJDH2003]. Because the patient-agents have to reason about the execution time of their treatments, time-dependent utility functions are developed which capture the health state development of the patients [Pau+2004]. In these utility functions the disease of a patient is viewed as disutility (decrease in quality of life) [HoRu1991] [PJDH2003]. Because the loss of utility adds up as long as the disease is not cured, this disutility over time can be viewed as opportunity costs for not curing the disease right away [PJDH2003]. Thus, the utility functions of the patient-agents are modeled as (opportunity) cost functions.

For the construction of these utility (or cost) functions, a cardinal measurement of the health state is required. Hospitals currently use numerous health state or priority measures, like the APACHE II score (*Acute Physi-*

ology and Chronic Health Evaluation) in intensive care units cf. [KDWZ1985], or a simple 1 to 6 priority scale to indicate the priority of cardiac patients (as observed in the performed field studies). In order to achieve inter-agent comparable priorities, the various used measures need to be expressed through one single health measure. For this reason, this work proposes a health state measure which was inspired by the (macro-economic) concept of *years of well being* [Torr1987], which already incorporates the notion of time. Here, a health state H of 1 denotes total health and 0 refers to death. However, the scale has no lower bound, as some health conditions might be valued worse than death (e.g. in some cases of coma patients). In order to determine the value of a health state H for a certain disease, it must be determined (by a decision maker) what time period xT of total health (H=1) equals one specific time period $1T$ with this disease, i.e.

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

Equation 1. Equivalence of time periods and health levels

Through this, the health state of a patient can be described in time units [PJDH2003]. For example, [Torr1987] determined a health state of 0.7 for a middle angina pectoris; in other words, that suffering one year from a middle angina pectoris equals 0.7 years of total health.

Because the loss of utility adds up as long as the disease is not cured, this disutility over time can be viewed as opportunity costs for not curing the disease right away [PJDH2003]. These opportunity costs $C(t)$ equal the difference between the achievable health state through treatment z and the patient's health state development over time without treatment $H(t)$. Because a treatment might not be able to restore total health, the achievable health state z might be lower than 1 (total health). In [Torr1987], for example, the health state after a kidney transplantation has a value of 0.84.

Further, the health state of a patient can either remain constant or can decrease over time. In case of a decreasing health state a linear reduction is assumed for practical reasons, i.e., $H(t)=s-bt$, where s denotes the initial health state and b the decrease rate [PJDH2003]. From this, the costs $C(t)$ are

$$C(t) = \int_0^t z - H(\tau)\,\mathrm{d}\tau = at + \frac{b}{2}t^2; a = z - s.$$

Equation 2. Opportunity cost

In other words, the initial health state *a* might also be viewed as the *severity* and the decrease rate *b* as the *criticality* of the patient's illness. Figure 1 shows an exemplary course of an illness with linear reduction of the health state, resulting in a quadratic, convex opportunity cost curve.

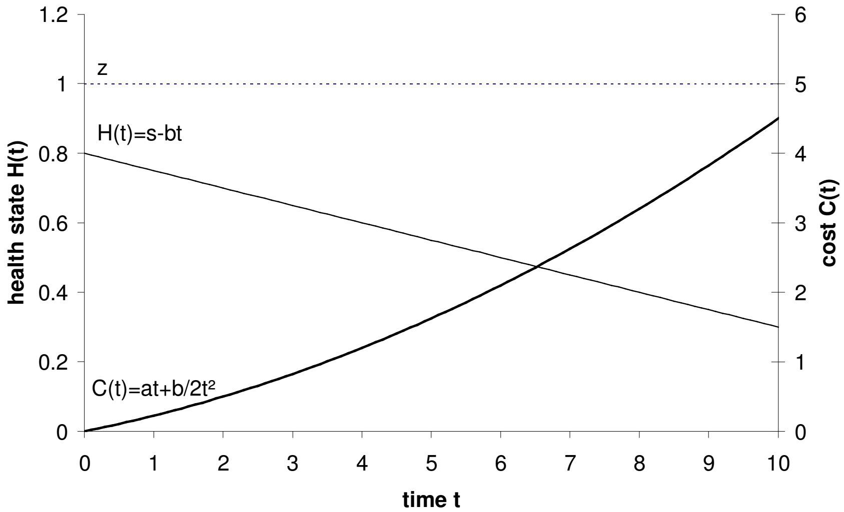

Figure 1. Linear reduction of the health state.

Finally, the achievable health state through treatment (z) might decrease. Because a decrease of the achievable health state due to late treatment would result in a lifelong decrease of the quality of life for the patient, these patients are treated immediately as emergency patients.

After the determination of the time dependent utility function, it is necessary to consider the case of *stochastic treatment durations*. Because the service times of the medical tasks might be stochastic, it is necessary for the patient-agents to consider this uncertainty in the bargaining process. For this reason, the cost function $C(t)$ has to be extended to a cost function $\tilde{C}(\mu,\sigma)$ based upon the expected mean μ and variance σ^2 of the starting time distribution $\varphi(t,\mu,\sigma)$. To calculate the value of $\tilde{C}(\mu,\sigma)$, the starting time distribution $\varphi(t,\mu,\sigma)$ has to be weighted with the cost function $C(t)$ of the patient agent, i.e.

$$\tilde{C}(\mu,\sigma) = \int_{-\infty}^{\infty} \varphi(t,\mu,\sigma)C(t)\,\mathrm{d}t.$$

Equation 3. Cost function depending on the expected mean and variance of the starting time (1)

Based upon decision theory the variance of the envisaged starting time for a task is viewed as risk (of delay), where a linear opportunity cost curve indicates risk neutrality, because the benefit from the chance to start earlier compensates (in case of a symmetric distribution function) the disutility through the chance of a delayed start. A convex opportunity cost function on the other hand indicates risk adversity, because the possible gains from an early start are outweighed by the possible losses due to a delayed start [Schn1991]. This should be illustrated by the following example equation, using a normal distribution and the described health state dependent cost function. The expected costs $\tilde{C}(\mu,\sigma)$ for a patient agent for a timeslot with a mean starting time μ and variance σ^2 can now be calculated by

$$\tilde{C}(\mu,\sigma) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \left(at + \frac{b}{2}t^2 \right) \mathrm{d}t = a\mu + \frac{b}{2}\left(\mu^2 + \sigma^2\right)$$

Equation 4. Cost function depending on the expected mean and variance of the starting time (2)

where the variance σ^2 is only influenced by b. With regards to decision theory, the health decrease rate b can be interpreted as the determinant of the agent's attitude to risk, that is, for b=0 the agent is risk neutral and for b>0 the agent is risk adverse [Schn1991]. This is illustrated in Figure 2, where curve A shows a risk adverse and curve B a risk neutral preference or cost function of the patients. However, if the starting time distribution is not symmetric, even risk neutral patients are sensitive to different variances.

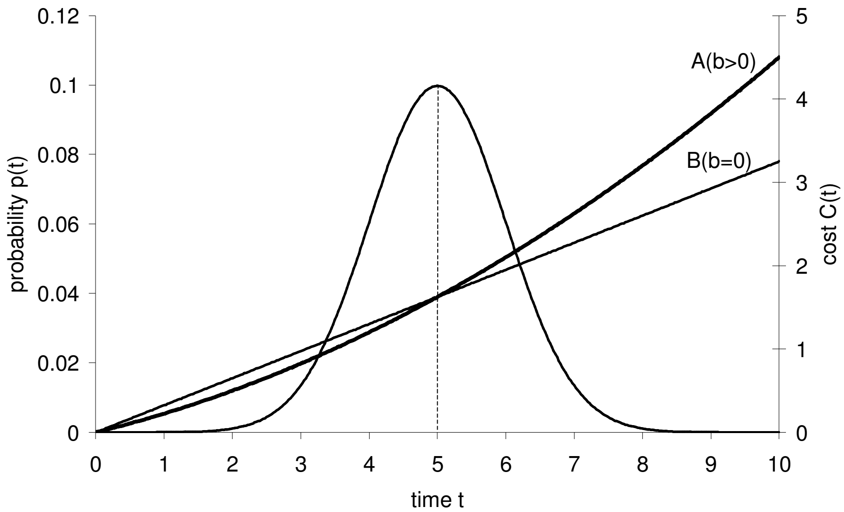

Figure 2. Stochastic treatment duration

Because the service times of the treatments and examinations in a hospital often do not correspond to a normal distribution, discrete distribution functions are used in this work. Therefore, the expected cost for a treatment results as the sum of the cost-values $at+b/2\times t^2$ of each time point (the n classes of the distribution) weighted by its probability p, i.e.

$$\tilde{C} = \sum_{i=1}^{n} p_i \left(at_i + \frac{b}{2} t_i^2 \right).$$

Equation 5. Expected cost for a treatment

3.2 Coordination mechanism

For the assignment of the treatments and examinations of the patients to the scarce hospital resources, a market inspired coordination mechanism (based upon the Contract Net Protocol [Smit1980] [DaSm1983]) is used. In this coordination mechanism the resources auction off their time-slots. Consequently, a resource time-slot is assigned to the patient-agent with the highest bid. The rational for this is, that the patient-agent who gains the highest utility from a specific time-slot is willing to pay the highest price for it (up to the expected utility).

In detail, the proposed coordination mechanism consists of four phases:

1. The *subscription phase*, where the patient-agents subscribe to the required resource-agents to inform them about their demanded medical tasks.
2. The *announcement phase*, in which the resource-agents initiate new auctions and announce them to the subscribed patient-agents.
3. The *bidding phase*, where the patient-agents generate and submit their bids for the needed time slot.
4. And the *awarding phase*, where the winner of the auction is determined.

Thus, this coordination mechanism turns the Contract Net Protocol on its head, as the potential contractors (resource-agents) announce their availability and the manager (patient-agents) bid for their pending tasks cf. [Durf2001].

3.2.1 Subscription phase

In order to participate at a resource auction the patient-agents must subscribe to the required resource-agents, that is, inform the resource-agents about the required treatments and examinations. However, to be able to subscribe to a resource-agent the patient-agents first must identify the resources capable of performing the needed treatments and examinations. Because the capabilities of the resources might overlap, i.e., different resources might be able to perform the same treatment, a *yellow page service* is used at which the resource-agents advertise their capabilities, and the patient agents inform themselves about the adequate resources. This allows the agents to flexible incorporate changes in the hospital environment (see also Part I, Chapter 3).

When a resource-agent receives a subscription from a patient-agent it informs the subscriber about the duration of the requested medical task. Because the service times of the medical tasks are stochastic, the resource-agent submits an array containing a discrete (empirical) distribution function, which is generated from historical task durations.

The subscription phase is somewhat distinct from the bidding and awarding phase. A subscription of a patient-agent is only needed if the treatment pathway of the corresponding patient is altered, that is, when a new treatment or examination is needed or an already registered medical task becomes obsolete. Thus, the main purpose of the subscription phase is to avoid unnecessary broadcast messages when a resource agent initiates a new auction, which is described next.

3.2.2 Announcement phase

The announcement phase initiates the actual auction mechanism. A resource-agent opens an auction for a new treatment or examination if its associated hospital resource has (almost) finished the current medical task. Similar to the current practice in hospitals, this allows the resource to react in an efficient manner to complications and emergency patients: if a treatment takes longer than expected or an emergency patient (who has not yet been entered into the information system) needs urgent treatment, the resource-agent just does not open a new auction until the exception is handled. Obviously, for this to work, the resource-agent needs some external input in order to update its beliefs about the state (busy or idle) of its physical resource.

In order to open a new auction, the resource-agent informs all subscribed patient-agents about the new auction and queries their envisaged starting time. In response, each contacted patient-agent replies the time at which the patient it represents is expected to be available. This can either be immediate if the patient is idle or later, otherwise. If the patient is not idle, the corresponding patient-agent transfers an array containing a distribution function of the finish time of the medical task the patient is currently engaged in. Based upon this information, the resource-agent computes the expected finish time distributions for all participating patient-agents. Then, the resource-agent submits these finish time distributions to the participants and calls for proposals (bids). This call for proposals initiates the bidding phase.

3.2.3 Bidding phase

In the bidding phase the patient-agents generate and submit bids for the prescribed medical procedures to the resource-agents. To be able to evaluate their current schedule and to calculate bid-prices for time slots, the patient agents rely on the utility functions described in Section 3.1. Based upon these utility functions the agents generate the bid-prices by calculating the expected loss of utility (cost of waiting) if they would loose a specific auction. In other words, the price a patient-agent is willing to bid for a specific time slot corresponds to the expected disutility the patient-agent would suffer if it does not win the auction. Therefore, the patient-agent has to determine the value of its own schedule with and without winning the desired time slot. To determine the schedule in case of loosing the auction, the aggregated finish time distributions of the other participating patient-agents are considered as block time of the corresponding resource. After a bid is generated, it is submitted (proposed) to the auctioneer.

3.2.4 Awarding phase

The last step of the coordination mechanism is the awarding phase. After the auctioneer (the resource-agent) has received the proposals containing the bids of the patient-agents, it awards its time slot to the patient-agent with the highest bid. In the case a patient-agent wins in more than one auction – as a patient-agent generally participates in multiple simultaneous auctions –, it must decide which award it should accept. Here, the patient-agent chooses the resource it gains the highest utility out of, i.e., to which it has submitted the highest bid. If an awarded resource time slot is rejected by the patient-agent, the corresponding resource-agent awards its auctioneered time-slot to the next best bidder until one patient-agent accepts the award or all participants rejected the time-slot.

4 Evaluation and benchmark

For the evaluation and benchmark of the proposed coordination mechanism, a prototypic multi-agent system was implemented. To test and evaluate the prototype under real-world conditions a simulation environment was built, that allows simulating different scenarios by varying several parameters, such as the hospital size, the divergence of treatment durations, or the probability of emergency cases. Additionally, in order to benchmark the proposed coordination mechanism against the status-quo patient scheduling in hospitals, a priority rule based strategy was also implemented.

4.1 Prototype realization

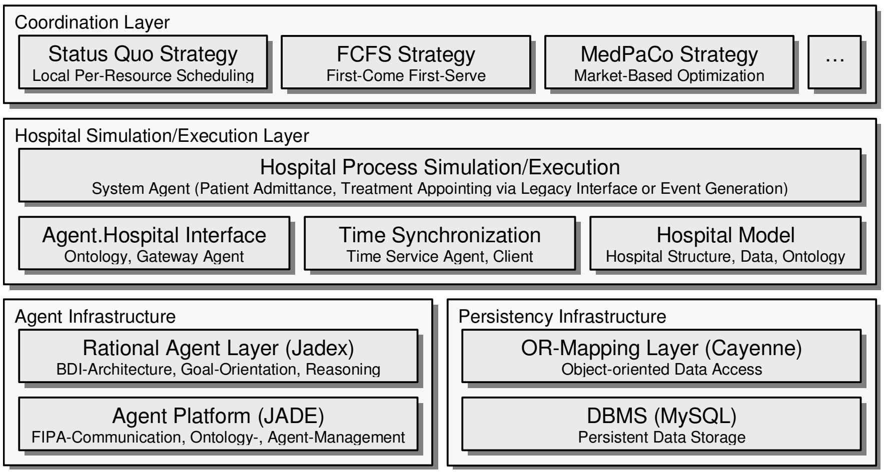

Figure 3. Prototype implementation.

The prototype implementation is organized in three separate layers: The coordination layer, the hospital layer, and the infrastructure layer (see Figure 3). The coordination layer is comprised of the different coordination mechanisms, each of which can be applied to perform the treatment scheduling. The coordination mechanism described in the previous sections and the alternative strategies used for benchmarking have been designed and implemented using agent-oriented tools and concepts. More details about the concrete realization of the coordination mechanisms can be found in [Pau+2004] [BrPL2004].

The hospital layer is designed to support the execution of the coordination by providing the facilities to perform simulation runs or to run the system as an application. When a simulation run is initiated, the information from the hospital model is used to create the hospital infrastructure consisting of initial patient and resource agents. During the run, the system agent uses different random distributions to approximate real arrival rates of patients and other occurrences like emergencies. It therefore decides at what time the next arrival or emergency will take place. The system agent is conceived to emulate all simulation external occurrences. Hence, for running the system as application instead of simulation it is merely required to adapt the system agent to react on some user interface and set-up the time service with real time.

The infrastructure layer provides system-level services for the implementation such as agent management and execution, as well as persistency. Basic agent services as the agent lifecycle management, agent communication and search facilities are provided by a FIPA-compliant agent middleware platform [PoCh2001]. These basic services are enhanced with a rational agent layer following the BDI-metaphor [RaGe1995], which enables the usage of goal-oriented concepts at the design and implementation level. Hence, it facilitates the development with the introduction of high-level agent-oriented programming concepts [PoBL2005]. The persistency infrastructure consists of a relational database management system, which is connected with an object-relational mapping layer. The mapping layer enables object-oriented access to the data by making the underlying relational database model transparent.

4.2 Scalability

The *scalability* denotes the additional computation effort (needed time to solve a problem) invoked by an increase of the problem size cf. [Durf2001] [LNND1998]. To be able to derive the scalability, the complexity of the test problems should only differ with respect to the problem size. For this, the open shop benchmark problems of [Tail1992] were used. In these problems the amount of jobs equal the amount of resources; thus the number of tasks is $n \times n$, where n denotes the number of resources or jobs, respectively.

The Taillard open shop benchmark consists of six different problem sizes (4×4, 5×5, 7×7, 10×10, 15×15, 20×20), each comprising ten problem instances (the used problems are available from [Tail1992]). Figure 4 shows the (logarithmic) mean run time of the proposed coordination mechanism ("Auctions") for each problem size ($n \times n$), and a curve representing a quadratic scaling ("$O(n^2)$") for comparison. Through comparison of the empirical run times of the proposed mechanism with the calculated $O(n^2)$ curve it can be stated, that the proposed mechanism approximately scales quadratic with the problem size.

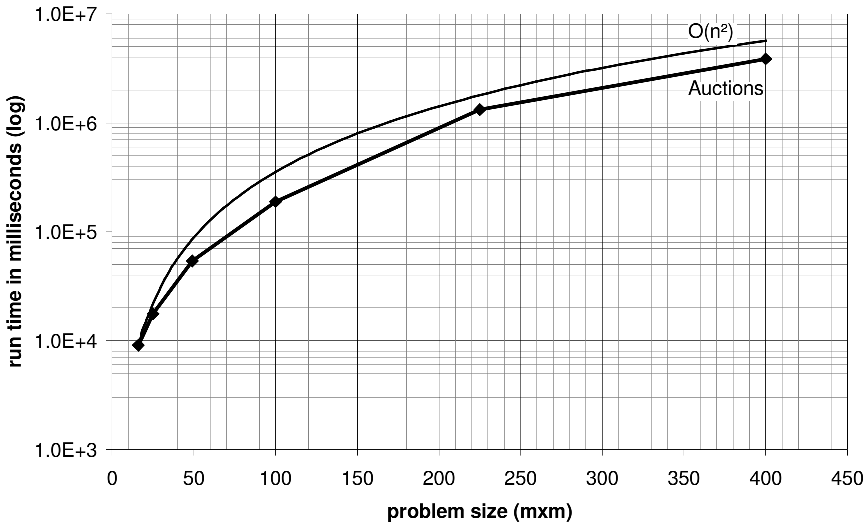

Figure 4. Scaling of Taillard n×n open shop problems

4.3 Continuous patient arrival

This subsection investigates the performance of the proposed coordination mechanism in a dynamic environment, where the patients arrive continuously at the hospital. To analyze the suitability, the mechanism is compared against a coordination mechanism using a *first-come first-serve* (FCFS) priority rule in three different scenarios:

1. short inter-arrival intervals with few medical tasks for each patient;
2. short inter-arrival intervals with many tasks; and
3. long inter-arrival intervals with many tasks.

A test of long inter-arrival intervals with few tasks was omitted, because pre-tests have shown that the problem was too easy (almost no resource conflicts occurred).

The setup of the tests is a follows. In all tests the last patient arrives not after the 400th minute. The *short* inter-arrival intervals where uniformly drawn out of the interval [1,10] (minutes), and the *long* inter-arrival intervals were randomly chosen between [1,60]. Because these tests are designed to analyze and compare the effect of short versus long inter-arrival intervals on the performance of the coordination mechanism separately, a uniform distribution of the inter-arrival intervals in both tests was chosen. The treatments for the patients were drawn out of a database containing

3393 actual, historically performed hospital treatments, involving the following six ancillary units: RAD (radiology), ECG (electrocardiography), ENDO (endoscopy), CT (computer-tomography), MR (magnetic-resonance-imaging), and NUC (nuclear-medicine).

For the test with the *few* medical tasks up to three treatments were assigned, and for the tests with *many* treatments one to seven treatments where drawn out of the treatment-database cf. [KuOP1993]. Thus, the most tasks have to be scheduled at the test with short inter-arrival intervals and many tasks. The best, average and worst achieved results (average patient idle time) of the test runs are given in Figure 5 (short-few), Figure 6 (short-many), and Figure 7 (long-many).

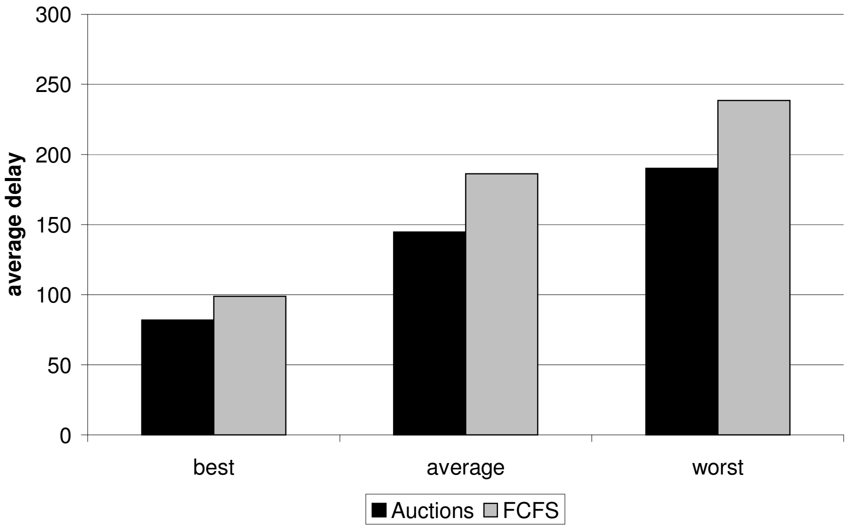

Figure 5. Results “short-few”

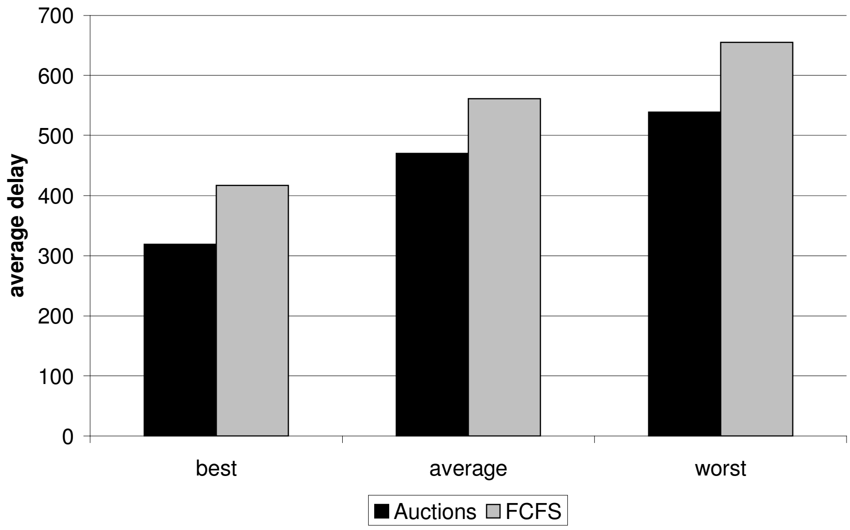

Figure 6. Results "short-many"

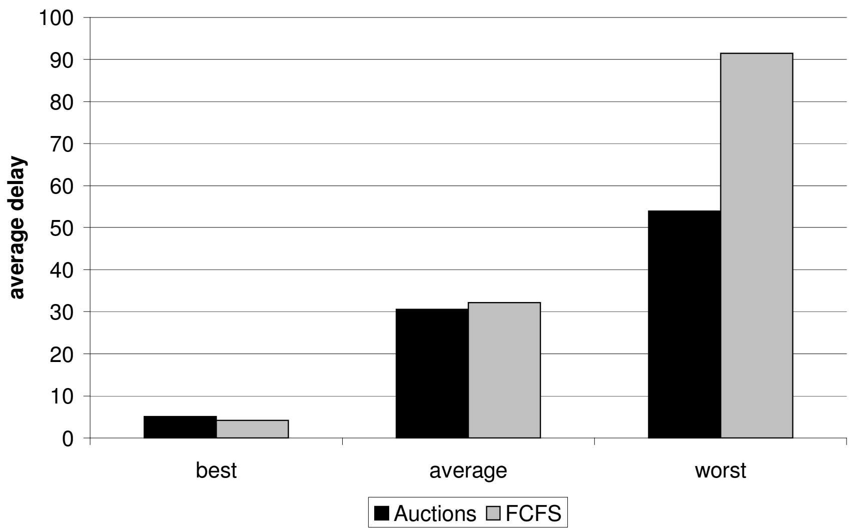

Figure 7. Results "long-many"

4.4 Resource capacity

In Figure 8, a solution achieved through the proposed coordination mechanism for one "small-many" test problem is given. To test the handling of multiple resources capable of performing the same treatment, the "small-many" test from Figure 8 was rerun with two additional radiological and endoscopic units, and one additional electrocardiography unit. The result of this modification is given in Figure 9, showing again a good load balancing behavior of the proposed mechanism.

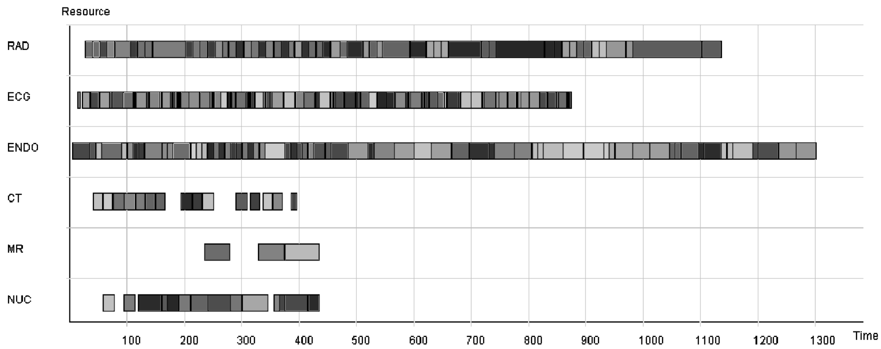

Figure 8. Single resource capacity

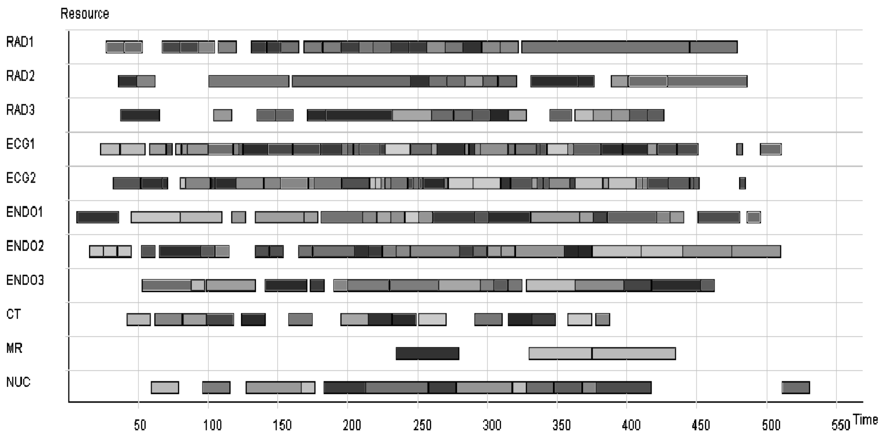

Figure 9. Increased resource capacity

4.5 Emergencies

To test the behavior of the proposed mechanism in case of emergencies, simulations were run with different emergency probabilities (the chance that an arriving patient is an emergency). Here, each emergency patient receives only one task, but this task must be performed immediately. Using the simulation setup of the previous subsection, tests with an emergency probability of 5, 10, 20, 25, and 50 percent were performed with the proposed coordination mechanism as well as with the first-come first-served mechanism. While normal patients do not arrive after the 300th minute, emergency patients can arrive until the 600th minute (the occurrence of emergencies must be bounded, because the system would run infinitely otherwise). Figure 10 shows a schedule with emergency patients (dark bars).

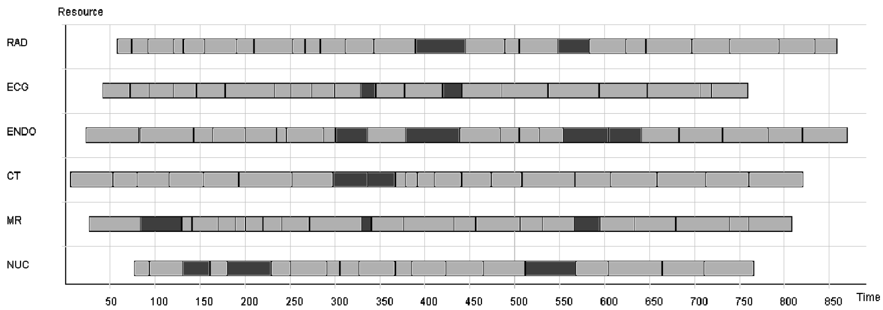

Figure 10. Resource allocation with emergency patients.

The results in percentage improvement of the idle time of the patients by the proposed coordination mechanism over the first-come first-served priority rule are given in Figure 11. Here, the achieved improvement over the hospital benchmark decreases with an increase in the emergency probability. However, this is plausible, because an increase of unpredictable tasks consequently reduces the scheduling potential of any scheduling approach.

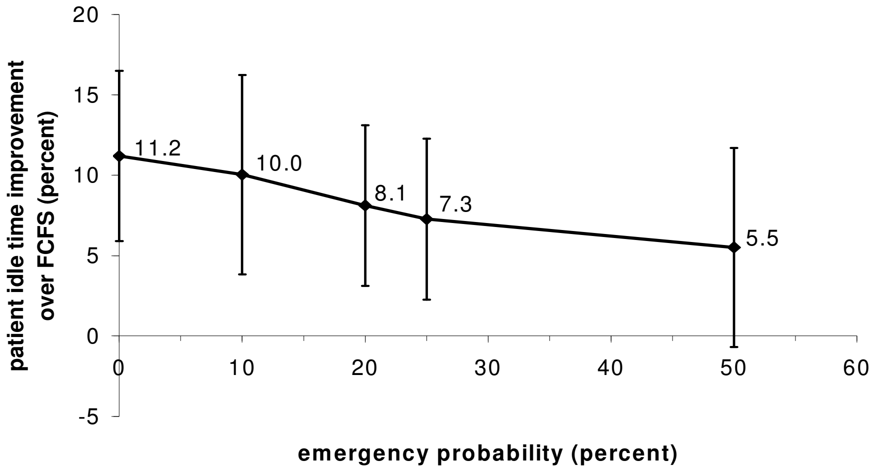

Figure 11. Results of simulation with emergency patients

5 Conclusions

Patient scheduling in hospitals requires a distributed and flexible approach in order to cope with the distributed structure of hospitals and to handle the inherent dynamics of the treatment processes. To this end, an agent-based coordination mechanism was presented in this chapter. Within this approach the patient-agents compete with each other over the scarce hospital resources. Through a decentralized auction mechanism the resource time slots are assigned to the patient-agents who gain the highest utility out of these time slots.

Because the utility of a patient in a hospital cannot – or at least should not – be based on the patient's willingness to pay for a specific resource time slot, it is important to develop utility functions which adequately represent the health state development over time. To this end, a novel health-state dependent utility function was introduced. Through these utility functions, the patient-agents can generate their bids for the time slot auctions at the resource-agents.

The proposed coordination mechanism significantly improves the current patient scheduling practice in hospitals (modeled as a first-come first-serve priority rule), while providing the required flexibility.

Currently the proposed coordination mechanism considers the health state of the patients as the only determinant of the patients priority. Thus, future research should address the question how the utility function of the patient agents can be adapted to handle multiple preferences. Here, the multi-attributive utility theory [Schn1991] might provide a good starting point. In this work, the proposed coordination mechanism was benchmarked against the status quo patient scheduling in hospitals. In future work, this benchmark will be extended to consider state of the art scheduling heuristics and meta-heuristics, e.g. genetic and evolutionary algorithms. Finally, the proposed mechanism will be evaluated in a real hospital. Here it is essential to integrate the existing legacy systems of the hospital. However, an agent-based system is assumed to be well suited for this, because the legacy systems can be encapsulated through agents, and thus easily be integrated into the framework cf. [Jen+2000].

References

[BrPL2004] Braubach, L.; Pokahr, A.; Lamersdorf, W.: MedPAge: Rationale Agenten zur Patientensteuerung. In: KI – Zeitschrift für Künstliche Intelligenz: Forschung, Entwicklung, Erfahrungen. Schwerpunktheft ‚Anwendungen von Softwareagenten' (2004)2, pp. 34-37.

[BrWo1991] Brahimi, M.; Worthington, D. J.: Queueing Models for Out-Patient Appointment Systems – A Case Study. In: Journal of the Operational Research Society 42(1991)9, pp. 733-746.

[DaSm1983] Davis, R.; Smith, R. G.: Negotiation as a Metaphor for Distributed Problem Solving. In: Artificial Intelligence, 20(1983)1, pp. 63-109.

[DeLi2000] Decker, K. S.; Li, J.: Coordinating Mutually Exclusive Resources using GPGP. In: Autonomous Agents and Multi-Agent Systems, 3(2000)2, pp. 133-157.

[Durf2001] Durfee, E. H.: Distributed problem solving and planning. In: Luck, M.; Marik, V.; Stepankova, O.; Trappl, R. (Eds.): Multi-agents systems and applications. Springer-Verlag, Berlin, 2001, pp. 118-149.

[Fein1999] Feinen, R.: Patientenbezogene Organisation von Behandlungsprozessen. In: Eichhorn, S.; Schmidt-Rettig, B. (Eds.): Profitcenter und Prozessorientierung: Optimierung von Budget, Arbeitsprozessen und Qualität. Kohlhammer, Stuttgart, 1999, pp. 188-199.

[GrTT1997] Greulich, A.; Thiele, G.; Thiex-Keye, M.: Prozeßmanagement im Krankenhaus. In: Schriftenreihe zum Managementhandbuch Krankenhaus, vol. 8. v. Decker, Heidelberg, 1997.

[HoRu1991] Horvitz, E.; Rutledge, G.: Time-Dependent Utility and Action Under Uncertainty. In: Seventh Conference on Uncertainty in

Artificial Intelligence. Morgan Kaufmann, San Mateo, 1991, pp. 151-158.

[Jenn2001] Jennings, N. R.: An agent-based approach for building complex software systems. In: Communications of the ACM 44(2001)4, pp. 35-41.

[Jen+2000] Jennings, N. R.; Faratin, P.; Norman, T. J.; O'Brien, P.; Odgers, B.: Autonomous Agents for Business Process Management. In: International Journal of Applied Artificial Intelligence, 14(2000)2, pp. 145-189.

[KDWZ1985] Knaus, W. A.; Draper, E. A.; Wagner, D. P.; Zimmermann, J. E.: APACHE II: A Severity of Disease Classification System. In: Critical Care Medicine 13(1985)10, pp. 818-829.

[KuOP1993] Kumar, A.; Ow, P. S.; Prietula, M. J.: Organizational Simulation and Information System Design: An Operations Level Example. In: Management Science 39(1993)2, pp. 218-240.

[LNND1998] Lee, L. C.; Nwana, H. S.; Ndumu, D. T.; De Wilde, P.: The Stability, Scalability and Performance of Multi-Agent Systems. In: BT Technology Journal 16(1998)3, pp. 94-103.

[Pau+2004] Paulussen, T. O.; Pokahr, A.; Braubach, L.; Zöller, A.; Lamersdorf, W.; Heinzl, A.: Dynamic Patient Scheduling in Hospitals. In: Bichler, M.; Holtmann, C.; Kirn, S.; Müller, J.; Weinhardt, C. (Eds.): Mulitkonferenz Wirtschaftsinformatik, Agent Technology in Business Applications. Essen, 2004.

[PJDH2003] Paulussen, T. O.; Jennings, N. R.; Decker, K. S.; Heinzl, A.: Distributed Patient Scheduling in Hospitals. In: Proceedings of the International Joint Conference on Artificial Intelligence. Acapulco, Mexico, 2003, pp. 1224-1232.

[PoBL2005] Pokahr, A.; Braubach, L.; Lamersdorf, W.: Jadex: A BDI Reasoning Engine. In: Bordini, R.; Dastani, M.; Dix, J.; Seghrouchni, A. (Eds.): Multi-Agent Programming, Kluwer (to appear in 2005).

[PoCh2001] Poslad, S.; Charlton, P.: Standardizing Agent Interoperability: The FIPA Approach. In: 9th ECCAI Advanced Course, ACAI 2001 and Agent Links 3rd European Agent Systems Summer School, EASSS 2001, Prague, Czech Republic. Springer, Heidelberg, 2001.

[RaGe1995] Rao, A.; Georgeff, M.: BDI Agents: from theory to practice, Systems (ICMAS'95). MIT Press, 1995.

[RoCh2003] Robinson, L. W.; Chen, R. R.: Scheduling doctors' appointments: optimal and empirically-based heuristic policies. In: IIE Transactions 35(2003), pp. 295-307.

[RoZl1994] Rosenschein, J. S.; Zlotkin, G.: Rules of Encounter. MIT Press, Cambridge, 1994.

[Schl1990] Schlüchtermann, J.: Patientensteuerung. Verlag Josef Eul, Bergisch Gladbach, 1990.

[Schn1991] Schneeweiß, C.: Planung. Springer-Verlag, Berlin, 1991.

[Smit1980] Smith, R. G.: The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. In: IEEE Transactions on Computers 29(1980)12, pp. 1104-1113.

[Tail1992] Taillard, E.: Benchmarks for basic scheduling problems. In: European Journal of Operations Research 64(1992), pp. 278-285.

[Torr1987] Torrance, G.-W.: Utility approach to measuring health-related quality of life. In: Journal of Chronic Diseases (1987)6, pp. 493-600.

[WeGo1996] Weinhardt, Ch.; Gomber, P.: Domänenunabhängige Koordinationsmechanismen für die dezentrale betriebliche Planung. In: Information Management (1996)1, pp. 6-16.