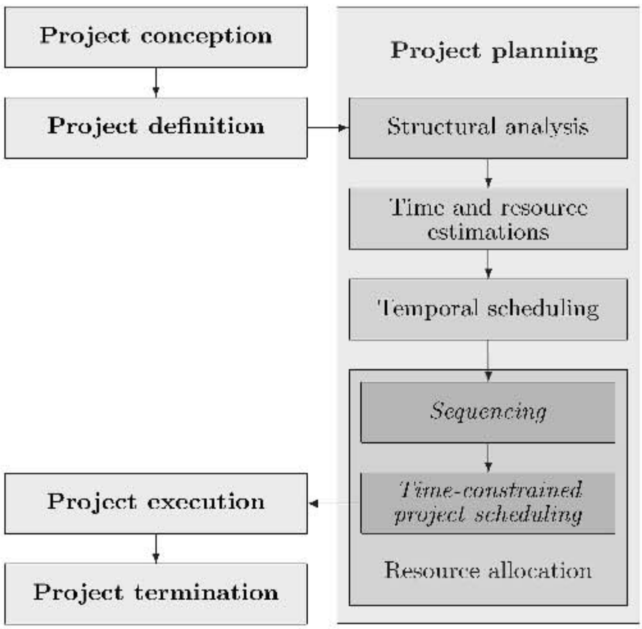
Introduction

Project management and resource allocation. A project is a major one-time undertaking dedicated to some well-defined objective and involving considerable money, personnel, and equipment. It is usually initiated either by some need of the parent organization or by a customer request. The life cycle of a project can be structured into five consecutive phases involving specific managerial tasks (cf. Klein 2000, Section 1.2). Starting with some proposal, several preliminary studies such as a feasibility study, an economic analysis, or a risk analysis are conducted in the *project conception phase* in order to decide whether or not a corresponding project will be performed. In the *project definition phase*, the objectives of the project are formulated, the type of project organization is selected, resources are assigned to the project, and different tasks with associated milestones are identified. Subsequently, the *project planning phase* at first decomposes each task into precedence-related activities by means of a structural analysis of the project. The time and resource estimations then provide the duration and resource requirements for each activity as well as temporal constraints between activities that are connected by precedence relationships. The result of the structural analysis and the time and resource estimations is the representation of the project as a network modelling the activities and the prescribed precedence relationships among them. Next, the temporal scheduling of the project provides the earliest and latest start times as well as the slack times of the activities, limitations with respect to resource availability yet being disregarded. The last and most complex issue of project planning consists in allocating the scarce resources over time to the execution of the activities. During the *project execution phase*, the implementation of the project is controlled by monitoring the project progress against the schedule which has been established in the project planning phase. In case of significant deviations from schedule, the resource allocation has to be performed again. The final *project termination phase* evaluates and documents the project after its completion to facilitate the management of future projects.

Each phase in the project life cycle requires specific project management techniques. Several recent textbooks on project management are devoted to the managerial and behavioral aspects of project conception, project definition, project planning, project execution, and project termination (see, e.g., Lewis 1998, Pinto 1998, Turner 1999, Keeling 2000, Meredith and Mantel 2002, or Kerzner 2003). This book is concerned with quantitative methods for the project planning phase and, more specifically, with the problem of optimally allocating resources over time.



Project planning within the life cycle of a project

The complexity of resource allocation arises from the interaction between the activities of a project by explicit and implicit dependencies, which may be subject to some degree of uncertainty. Explicit dependencies are given by the precedence relationships between activities emanating from technological or organizational requirements. In the course of time estimation, those dependencies are transformed into temporal constraints between activities. The scarcity of the resources used establishes an implicit dependency between activities, which can be formulated as resource constraints referring to sets of activities competing for the same resources or in terms of an objective function penalizing excessive resource requirements. The *resource allocation* problem consists in assigning time intervals to the execution of the activities while taking into account the prescribed temporal constraints and resource scarcity. If resource constraints are given, we also speak of a resource-constrained project scheduling problem. We distinguish between two subproblems: *sequencing* and *time-constrained project scheduling*. The limited availability of resources necessitates the definition of additional precedence relationships between activities when performing the resource allocation task. Again, those precedence relationships can be expressed in the form of temporal constraints. In contrast to the structural analysis, however, the precedence relationships to be introduced are subject to decision. This sequencing problem forms the core problem of project planning. Time-constrained project scheduling is concerned with computing the project schedule such that all temporal constraints – predetermined by the structural analysis or arising from sequencing – are observed and some objective function reflecting the managerial goal of the project is optimized. In the resource allocation methods developed in this book, sequencing and time-constrained project scheduling will be performed jointly in an iterative manner.

If activities can be performed in alternative execution modes that differ in durations and resource requirements, the selection of an appropriate execution mode for each activity may be included into the resource allocation problem. In that case, the time and resource estimations provide the sets of alternative execution modes, and solving the mode assignment problem constitutes the first step of resource allocation. Depending on whether the sets of execution modes are countable or uncountable, we speak of a discrete or a continuous mode assignment problem. A resource allocation problem that comprises a mode assignment problem is termed a multi-mode resource allocation problem.

Historical perspective and state of the art. Algorithms for resource allocation in project management date back to the 1960s, see Davis (1966), Lane (1968), Herroelen (1972), and Davis (1973) for overviews. The early work was concerned with three types of resource allocation problems: the time-cost tradeoff problem, the project duration problem, and the resource levelling problem. For all three problem types it is assumed that a strict order in the set of activities specifies completion-to-start precedence constraints among activities. The *time-cost tradeoff problem* is a multi-mode resource allocation problem which arises when certain activity durations can be reduced at the expense of higher direct cost. The project budget is then regarded as the resource to be allocated. If for each activity the cost incurred is a convex function in the activity duration, the continuous mode assignment problem that consists in computing all combinations of project duration and corresponding least-cost schedule can be determined by applying network flow techniques (see Kelley 1961). A survey of multi-mode resource allocation problems including different types of tradeoffs between the durations, resource requirements, and direct costs of activity execution modes can be found in Domschke and Drexl (1991). The *project duration problem* consists in scheduling the activities of a project subject to the limited availability of renewable resources like manpower or machinery such that all activities are completed within a minimum amount of time. Early approaches to solving the project duration problem include mixed-integer linear programming formulations (see, e.g., Wiest 1963) and priority-rule methods (cf. Kelley 1963, Verhines 1963, and Moder and Phillips 1964). The objective when dealing with a *resource levelling problem* is to "smooth" the utilization of renewable resources over time as much as possible, within a prescribed maximum project duration. In some cases, a desired threshold limit on the resource availability is given, and resources are to be levelled around this target. In other cases, one strives at minimizing the variance of resource utilization over time or minimizing the absolute magnitude of fluctuation in the resource profiles. The first procedures for resource levelling offered by Burgess and Killebrew (1962) and Levy et al. (1963) were based on sequentially moving in time slack activities.

In the following years, a great deal of effort has been devoted to heuristic and exact algorithms for the project duration problem. In the 1990s, project planning methods gained increasing importance from their applicability to scheduling problems arising beyond the area of proper project management, for example, in production planning, time-tabling, or investment scheduling. Different generalizations of the basic resource allocation problems have received growing attention in recent years. Those expansions include a variety of objectives as well as the presence of different kinds of resources, general temporal constraints given by prescribed minimum and maximum time lags between the start times of activities, and uncertainty with respect to activity durations. For a review of solution procedures, we refer to the survey papers by Icmeli et al. (1993), Elmaghraby (1995), Özdamar and Ulusoy (1995), Tavares (1995), Herroelen et al. (1998), Brucker et al. (1999), Kolisch and Padman (2001), and Kolisch (2001*a*). A comprehensive state-of-the-art overview of the field is given by the handbook of Demeulemeester and Herroelen (2002), with an emphasis on algorithms for project scheduling problems with precedence constraints among activities instead of temporal constraints. A detailed treatment of specific project scheduling problems of the latter type can be found in the monographs by Kolisch (1995), Schirmer (1998), Hartmann (1999*a*), Klein (2000), and Kimms (2001*a*). The book by Hajdu (1997) is mainly concerned with several types of time-cost tradeoff problems. Solution procedures for several project scheduling problems with general temporal constraints have been discussed by De Reyck (1998), De Reyck et al. (1999), Neumann and Zimmermann (1999*a*), Zimmermann (2001*a*), and Neumann et al. (2003*b*). Models and algorithms for project scheduling with stochastic activity durations are studied in the doctoral dissertations of Stork (2001) and Uetz (2002). A review of models and algorithms for projects with stochastic evolution structure can be found in Neumann (1999*b*).

Contribution. In this monograph we discuss structural issues, efficient solution methods, and applications for various types of *deterministic resource allocation problems* including general temporal constraints, different types of resource requirements, and several classes of objective functions. The diversity of the models dealt with permits us to cover many features that arise in industrial scheduling problems. Each resource allocation problem gives rise to a corresponding project scheduling model, which provides a formal statement of the resource allocation task as an optimization problem. This model may or may not include explicit resource constraints. In the latter case, limitations on the availability of resources are taken into account by the objective function. Our main focus in this book is on developing a unifying algorithmic framework within which the different kinds of project scheduling models can be treated. This framework is based on the seminal work by research groups around Rolf Möhring and Franz-Josef Radermacher, who have proposed an order-theoretic approach to stochastic and deterministic resource-constrained project scheduling (see, e.g., Radermacher 1978, Möhring 1984, Radermacher 1985, or Bartusch et al. 1988). We extend the order-theoretic approach to resource allocation problems involving so-called cumulative resources, which represent a generalization of both renewable and nonrenewable resources known thus far. Based on the results of a structural analysis of resource constraints and objective functions, we discuss two general types of resource allocation procedures. By enhancing the basic models treated with supplemental kinds of constraints we bridge the gap between issues of greatly academic interest and requirements emerging in industrial contexts.

Synopsis. The book is divided into six chapters. Chapter 1 provides an introduction to three basic project scheduling models. First we address project scheduling subject to temporal constraints and review how the temporal scheduling calculations for a project can be performed efficiently by calculating longest path lengths in project networks. We then discuss resource constraints which arise from the scarcity of renewable resources required. If the availability of a resource at some point in time depends on all previous requirements, we speak of a cumulative resource. We consider the case where cumulative resources are depleted and replenished discontinuously at certain points in time. The available project funds, depleted by disbursements and replenished by progress payments, or the residual storage space for intermediate products are examples of cumulative resources. For both kinds of resource constraints, we explain how to observe the limited resource availability by introducing precedence relationships between activities.

In Chapter 2 we discuss a relation-based characterization of feasible schedules, which is based on different types of relations in the set of activities. Each relation defines a set of precedence constraints between activities. This characterization provides two representations of the feasible region of project scheduling problems as unions of finitely many relation-induced convex sets. Whereas the first representation refers to a covering of the feasible region by relation-induced polytopes, the second representation arises from partitioning the feasible region into sets of feasible schedules for which the same precedence constraints are satisfied. Those two representations are the starting point for a classification of schedules as characteristic points like minimal or extreme points of certain relation-induced subsets of the feasible region. For different types of objective functions, we show which class of schedules has to be investigated for finding optimal schedules.

Depending on the objective function under study, we propose two different basic solution approaches. Chapter 3 is dedicated to relaxation-based algorithms, which at first delete the resource constraints and solve the resulting time-constrained project scheduling problem. Excessive resource utilization is then stepwise settled by iteratively introducing appropriate precedence relationships between activities (i.e., sequencing) and re-performing the time-constrained project scheduling. For different objective functions, we discuss efficient primal and dual methods for solving the time-constrained project scheduling problem. Those methods are used within branch-and-bound algorithms based on the relaxation approach.

If we deal with objective functions for which time-constrained project scheduling cannot be performed efficiently, we apply a constructive approach. The candidate schedules from the respective class are enumerated by constructing schedule-induced preorders in the activity set and investigating appropriate vertices of the corresponding polytopes. In Chapter 4 we treat local search algorithms operating on those sets of vertices, where the schedules are represented as spanning trees of preorder-induced expansions of the underlying project network.

Chapter 5 is concerned with several expansions of the basic project scheduling models. First we discuss the case where during certain time periods given by break calendars, resources are not available for processing activities. Certain activities may be suspended at the beginning of a break, whereas other activities must not be interrupted. Suspended activities have to be resumed immediately after the break. The second expansion consists in sequence-dependent changeover times between the activities of a project. Changeover times occur, for example, if the project is executed at distributed locations and resources have to be transferred between the different sites. Next, we review methods to discrete multi-mode project scheduling, where activities can be performed in a finite number of alternative execution modes. Finally, we consider continuous cumulative resources that are depleted and replenished continuously over time.

In Chapter 6 we discuss several applications of the models treated in Chapters 1 to 5 to scheduling problems arising outside the field of proper project planning in production planning (make-to-order and small-batch production in manufacturing, batch scheduling in the process industries) and finance (evaluation of investment projects). Finally, we propose two alternative techniques for coping with uncertainty in project scheduling, which is commonly encountered when performing real-life projects.