Lecture: Scheduling

Nicole Megow

Technische Universität München
Zentrum für Mathematik

WS 2015/16

Organisatorial issues

- Lectures: Wed 8:30 – 10:00 in Room 00.06.011 (MI Hörsaal 3)
- Exercises (by Roman Rischke)
  - Wed 10:15 – 11:45 in Room 02.04.011 (seminar room)
  - begin next week
  - two groups in alternating weeks
- Lecture notes: electronic blackboard
- Homepage: `http://www-m9.ma.tum.de/WS2015/Scheduling`
- Exam: oral/written (will be decided later)

Content of this course

- Classification of scheduling problems
- Focus on standard deterministic scheduling problems
- Complexity of scheduling problems
- State of the art solution methods and analysis of algorithms
- Exact and approximation algorithms

Scheduling models are motivated by real-world applications. In this course we focus on theoretical foundations.

Literature

books

- M.L. Pinedo (2012). *Scheduling. Theory, Algorithms, and Systems*. Springer.
- P. Brucker (2007). *Scheduling Algorithms*. Springer.
- F. Jaehn and E. Pesch (2014), *Ablaufplanung. Einführung in Scheduling*. Springer.

articles

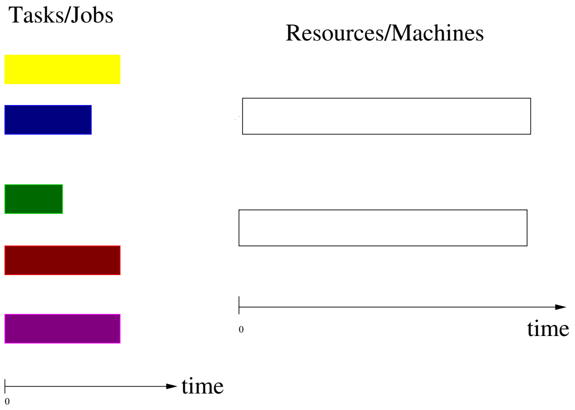
- R.L. Graham, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan (1979). Optimization and approximation in deterministic sequencing and scheduling: a survey. *Annals of Discrete Mathematics* **5**, pp. 287–326.
- E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, D.B. Shmoys (1993). Sequencing and scheduling: Algorithms and complexity. Chapter 9 in S.C. Graves, A.H.G. Rinnooy Kan, P.H. Zipkin. *Logistics of production and inventory.* Handbooks in Operations Research and Management Science **4**.
- B. Chen, C.N. Potts, G.J. Woeginger (1998). A review of machine scheduling: complexity, algorithms and approximability. In D.Z. Du and P.M. Pardalos, Handbook of Combinatorial Optimization, pp. 21–169.

Introduction and Overview

What is scheduling?

Definition (Lawler, Lenstra, Rinnooy Kan and Shmoys (1993))

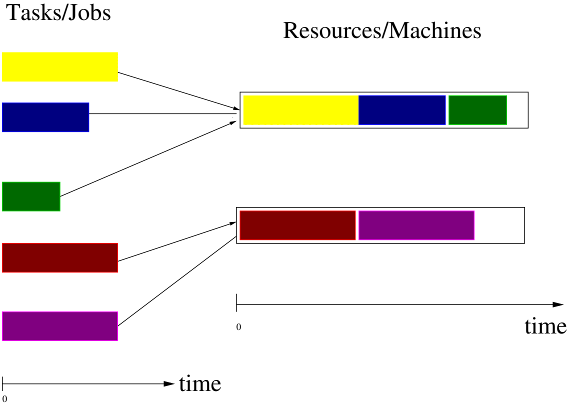
Sequencing and scheduling is concerned with the optimal allocation of scarce resources to activities over time.



What is scheduling?

Definition (Lawler, Lenstra, Rinnooy Kan and Shmoys (1993))

Sequencing and scheduling is concerned with the optimal allocation of scarce resources to activities over time.



Elements

Resources:

Machines, staff, crane, construction crews, CPUs, ...

Jobs:

Operations in a production process, stages in a construction project, execution of computer programs, ...

Objectives:

- Minimizing last completion time
- Maximize number of tasks completed on time
- ...

Example 1: Gate assignment at an airport

- Airport: dozens of gates, hundreds of planes arriving/departing per day
- Flight schedule: specifies arrival/departure time, passenger demand
- Service time at gate: dismbarking, service of the plane, boarding
- Task: assign planes to gates while minimizing objectives such as work for airline personnel and airplane delays

Example 2: Staff scheduling in hospitals

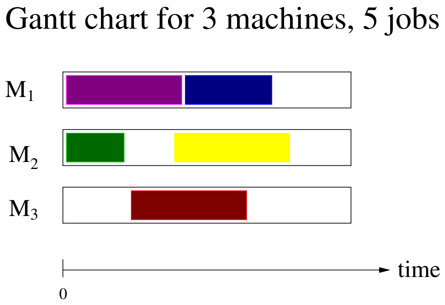
- In a hospital, staff is working night and weekend shifts
- Due to labor agreements and law, not too many consecutive night shifts can be done by one staff member
- Shifts need to be divided over staff in a fair manner
- Staff members can submit their preferences
- First goal: feasible schedule
- Second goal: fair schedule

Example 3: Scheduling tasks in a CPU

- CPU =central processing unit
- Allocate CPU time to tasks.
- Goal: Minimize average time a tasks spends in the system.

General scheduling models – representation

- m machines, $M_1, \ldots, M_m$ or $1, \ldots, m$
- n jobs, $J_1, \ldots, J_n$ or $1, \ldots, n$
- A job j has processing time p_{ij} when processing on machine i.
- Schedules represented by Gantt charts (developed by Henry Gantt in the 1910s to illustrate project schedules.)



Classification of scheduling problems

Classification of scheduling problems

Many scheduling problems can be described by a three-field notation $\alpha\,|\beta\,|\gamma$:

- α denotes machine environment
- β denotes job characteristics
- γ denotes objective function

Remark

A field may contain more than one entry, but may also be empty.

Classification – machine environment (α)

- $\alpha = 1$: single machine
- $\alpha = \mathrm{P}$: identical parallel machines
  - m identical machines
  - processing time of job is independent of machine: $p_{ij} = p_j$
- $\alpha = \mathrm{Q}$: uniform or related parallel machines
  - m parallel machines with speeds $s_1, \ldots, s_m$
  - processing time $p_{ij} = p_j / s_i$
  - if $s_1 = \ldots = s_m$ then case of identical machines
- $\alpha = \mathrm{R}$: unrelated parallel machines
  - m different parallel machines
  - processing times arbitrary p_{ij}
- $\alpha = \mathrm{P}m, \mathrm{Q}m, \mathrm{R}m$: number of machines is m (fixed)
- ...

Classification – job characteristics (β)

β may contain several values

- r_j: release dates
  - if $r_j \in \beta$ then job may not start before time r_j
  - if $r_j \notin \beta$ then job may start at any time
- *pmtn*: preemption
  - if $pmtn \in \beta$ then processing of a job may be interrupted and resumed later even on a different machine
  - if $pmtn \notin \beta$ then the processing of an operation must not be interrupted
- *prec*: precedence constraints
  - there are precedence relations (partial order) between jobs: a job may not start before all its predecessors have been finished
  - typically represented by directed acyclic graph
  - special forms of precedence relations, e.g., *chain*, *intree*, *outtree*
- d_j: jobs have deadlines by which they must finish
  $\longrightarrow$ discussion due date vs. deadline

Classification – objective function (γ)

Notation:

- C_j completion time of job j
- $F_j = C_j - r_j$ flowtime of job j
- $L_j = C_j - d_j$ lateness of job j
- $T_j = \max\{0, C_j - d_j\}$ tardiness of job j
- U_j unit penalty: if $C_j > d_j$ then $U_j = 1$, otherwise $U_j = 0$

Objective functions (min):

- $C_{\max}$: makespan $C_{\max} = \max_j C_j$
- $F_{\max}$: maximum flowtime $F_{\max} = \max_j F_j$
- $L_{\max}$: maximum lateness $L_{\max} = \max_j L_j$

- $\sum(w_j)C_j, \sum(w_j)F_j, \sum(w_j)T_j, \dots$
- $\sum(w_j)U_j$ (weighted) number of tardy jobs