Abstract

Shunt planning is the capstone of the planning process at Dutch Railways Passengers (NSR). A shunt plan contains matching, parking and routing decisions for all train units on a shunt yard that are not needed for a certain amount of time.

Shunt plans are nowadays made by hand for every station and every night of the week. A shunt plan is highly sensitive to changes in previous planning pocesses such as timetabling and the planning of the rolling stock circulation. Furthermore, the capacity on the shunt yards is limited. Therefore one wants to have tools that are able to produce efficient shunt plans quickly.

In this report, after a description of the shunt problem, the tools currently in development by NSR are described and analyzed. One of these tools is the integral approach of matching and parking as described in the PhD thesis of R.M. Lentink, (see [Len06]). The disadvantages of this sequential approach of *routing* after *matching and parking* are discussed.

This leads to a description of a new integral approach of matching and parking with routing: the APT-model. The resulting mixed integer programm has many variables and constraints. Therefore a solution method is presented where variables and constraints are added in several steps without losing the strength of the integral approach. The APT-model has been implemented and tested for several instances. The results are discussed and lead to suggestions for further research.

.

Shunt planning, an integral approach of matching, parking and routing

M.R. den Hartog

August 20, 2010

Preface

NSR transports more than one million passengers daily. To do this NSR owns nearly three thousand carriages and employs more than three thousand conductors and approximately 2500 train drivers (level at 2007).

The demand for transport differs during the day. Especially during rush hours more capacity of rolling stock is needed, whereas in the night nearly all carriages are on the shunt yards. Therefore the concept of *train units* is founded: When rush hours start extra units can easily be coupled to the trains. After rush hours they can be decoupled. At the end of the day, complete ending trains can enter the shunt yard.

The units that are not needed for a certain amount of time can not all be left at the platform tracks, because the capacity of these tracks is used by train services and freight trains. therefore the shunt yards have been extended by building tracks to park train units. These tracks are mostly outside the station area because there the land is less expensive. On some shunt yards also equipment for the internal and external cleaning of units and for other processes has been built.

The process of handling units at shunt yards is called *shunting*. It includes the parking of train units, the internal and external cleaning and the routing to and from their park tracks. Shunt planning consists of making schedules for these processes. Besides that, crews have to be assigned to execute the shunt plans.

Shunt planning is the capstone of the planning process at NSR. It is subordinate to other logistical planning issues, like calculating the timetables, determining optimal train lenghts and crew scheduling. therefore a change in one of these plans nearly always influences the shunt plan. Moreover disruptions may disturb the practical execution of the shunt plan. Conversely a bad shunt plan can distrub the train services.

So the making of the shunt plan is an important part of the planning process of NSR. It is done for every station where carriages may be parked. Shunt planning is nowadays done by hand and takes lot of time. This makes that adjustments in train lengths give much work to the planners. For a more efficient usage of rolling stock, one wants to change these lenghts more times a year to match the demand at that period. Besides that, also the incrementing usage of the shunt yards by scheduling more and longer trains causes that the pressure on the capacity of the shunt yards and its planners grows. Moreover, when some parts of the infrastructure that influence the shuntplan are under construction or out of order the properties of the infrastructure with its possibilities temporarily change.

That is why algorithms are developed to deal with some aspects of the shunt plan. In this report these algorithms are treated and analyzed. Besides a new approach is presented.

I have written this report with much help of my supervisors Johann Hurink, Leo Kroon and Stefan Schuurman. I want to thank them for their guidance during this project.

Contents

1 Introduction to shunting

In this section we give an introduction to shunting. We introduce terms and we try to give an idea what makes shunt planning challenging.

1.1 Shunting

As mentioned in the preface, shunting consists of several processes to train units that are not needed for a certain amount of time. The most important ones are *parking, routing, cleaning and maintenance* and *crew scheduling.* In this report we only consider parking and routing. They are described in subsections 1.4 and 1.5. The cleaning, maintance and crew scheduling is outside the scope of this thesis.

1.2 Train units

The rolling stock is generally partitioned into several *families*, each with its own characteristics. Examples are units with faster acceleration and deceleration and double-deck units. The VIRM-family is an example of a double-deck family. It is used for intercity services. Units from the same family can be coupled and driven together. One specific family typically consists of two *types*. A specific type ψ within a specific family is discerned from the other type in the same family by its number of carriages. A unit of type VIRM 4 consists of four carriages and a unit of type VIRM 6, belonging to the same family, counts six carriages. The type of a unit u is denoted by ψ_u and Y is the set of all types of all families.

1.3 Tracks

The railway tracks have several features. The most important ones are: routing of train units, boarding and alighting of passengers, parking of train units, cleaning of rolling stock and small maintenance of rolling stock. Note that one specific track can have multiple functions. For example, a platform track can be used for parking train units at night, when it is not used for passenger or freight services. Certain escape routes out of the shunt yard have to be kept free. Tracks also have several characteristics. For shunting, the most important characteristics of a track are:

- The length of a track. This influences the amount of rolling stock that can be parked at a shunt track.

- The sides from which rolling stock can approach a track. Tracks that can be approached from both sides provide additional possibilities for parking train units as compared to tracks with a dead-end side.
- The availability of catenary. Train units with electric power can only be parked at a track with catenary.
- The availability of a railway safety system. Tracks which are not controlled by such a system, require the local traffic control organisation to avoid collisions. In some exceptions the driver relies on his sight.
- The availability of several types of equipment along the track. Examples are a battery charger, which is needed for parking diesel powered train units, and equipment for filling the water tank of toilets.

1.4 Parking

The parking of train units is far from trivial because in general parking capacity is scarce. In addition, the choice to park a train unit on a particular shunt track has several implications.

Firstly, if the train units at a shunt track are of different types, then the order of the train units on the park tracks is important. Obstruction of arriving or departing units by other units is not allowed. Otherwise a *crossing* occurs. For example, consider units ICM 3 and ICM 4 at a dead end track in figure 1. If unit ICM 3 has to depart first, a crossing occurs.

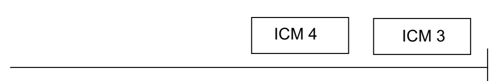

Figure 1: Example of a crossing

Secondly, the parking decision restricts the possible routes between the platforms and the shunt tracks. Thirdly, crews have to be available to carry out the resulting shunt activities within certain time intervals. Finally, certain routes and shunt tracks are preferred over others by shunt planners. Here, a track is preferred if it is located close to the platform tracks, or if it is rarely used for other purposes, e.g. for through train services or for temporary parking of rolling stock.

1.5 Routing

Routing of train units through a station takes place from specific arrival platforms to the park tracks and back to specific departure platforms or between arrival en departure platforms directly. When a platform track is available, it is possible to leave a train unit for a certain period of time on this platform after arrival or to place it there some time before departure. This introduces some flexibility with respect to the timing of the routing. Additional routing could be necessary for other processes, such as cleaning at dedicated tracks. Of course, the routes of the different train units should neither conflict with each other nor with the routes of the through train services, or other infrastructure reservations, such as track maintenance.

In this report a qualitative description of the routing and parking problem is given. It can be found in part I. Thereafter, in part II, notation is introduced which is used to give a qualitative description. After that in part III the algorithms for the routing and parking decision that are currently in development are treated and analyzed. This results in a plea for an integrated approach of these decisions. The integrated problem is modelled in part IV and its solution method is treated in part V. Finally, in part VI conclusions and suggestion for further research are proposed.

Part I
Qualitative description

In this part we give a qualitative description of the routing and parking decision. In section 2 the input for making the shunt plan is treated. Definitions about rolling stock, infrastructure and timetables are given. After that the definition of a shunt plan is presented. This is done in section 3 where also the requirements and wishes to the shunt plan are treated. In section 4 a summary of the first part is presented.

2 Input

In this section the input for a shunt problem and requirements to it are described. In succession the rolling stock circulation, infrastructure and timetables are treated.

2.1 Rolling stock circulation

2.1.1 Compositions

Before the making of the shunt plan, the *rolling stock circulation* is fixed. In this process to every train service in the timetables a number of units is assigned. These units are from the same family and in a certain order, called a *composition*. The several units of a composition do not have to be of the same type. An example of a composition is (ICM 4, ICM 3, ICM 3). Note that this composition is considered different from composition (ICM 3, ICM 4, ICM 3), because of the different order of the types.

2.1.2 Composition changes

A train service has a start and an end station according to the timetables. The composition of a train service may change between these stations. At dedicated stations, units may coupled or decoupled. The remaining of the composition has to stay the same. The composition of the example above (ICM 4, ICM 3, ICM 3) could after coupling a ICM 4 look like (ICM 4, ICM 3, ICM 3, ICM 4) or like (ICM 4, ICM 4, ICM 3, ICM 3), depending on the side at which it is coupled. However, it is impossible to end up with (ICM 4, ICM 3, ICM 4, ICM 3), because the ICM 4 can not be placed somewhere in the middle of the composition.

2.1.3 Supply

The units that are decoupled enter the *supply* of the station. The supply of a station is specified by the rolling stock circulation and contains at every moment for every type $\psi \in Y$ a non negative number of units. The term 'supply' should not be confused with the expression 'supply and demand' in which it has another meaning. Intiutively the *supply* contains all units that are not needed for a certain amount of time. During rush hours the supply typically is empty.

A unit that goes into the supply is only allowed to go out after a certain amount of time T. Thereafter the unit may depart in a departing train service. The time T is needed to route the unit to the right departure track and possibly to split from or couple it to other units.

2.1.4 Transition

At the end station of a train service some units of the composition may be assigned to a train service departing within a short time from the same platform track. In this case the arriving train service has a *transition* to the departing train service, which is called its *successor*. Units may be coupled or decoupled, but again the remaining of the composition has to stay the same. If an arriving train has no successor, all units of its composition enter the supply.

The rolling stock circulation does not prescribe what to do with units when they are in the supply. For units in the supply it is even not decided to which departing trains they will be assigned. Making these decisions for all units is called *matching*. Matching is part of making the shunt plan.

2.1.5 Parts

A *part* is an entity of one or more adjacent train units entering or leaving the supply in the same train service. A part can be decoupled from or coupled to a composition or forms a complete ending or starting train service.

The units of a part do not have to be handled in the same way. For example an arriving part may be split and the units may depart in different train services.

2.1.6 Sides of the train

In some cases the side of the composition at which units are coupled or decoupled is restricted. We distinguish between the front side and the back side of a train service. These sides are determined by **departure**: the *front side* is the side where the driver sits when he drives the departing train and the *back side* of a train is the rear of the train when it departs. At the end station of a train service the sides of the train are deduced from the departure of its successor. If it has no transition, we do not define a front or back side.

At a station between the start and end station coupling units is done at the front of the train and decoupling units is done at the back. For trains that change direction at the station, coupling may not be done at the front side, but at the back side of the train only. It is not allowed both to couple and to decouple from the same train at a station.

If there is a transition at the end station the side where units may be coupled or decoupled is not restricted. But also at the end station it is not allowed to both couple and decouple units.

2.2 Infrastructure

To make a shunt plan for a specific station the properties of the shunt yard of that station have to be known. In figure 2 an example of a shuntyard is presented. The names of the tracks are depicted above the tracks.

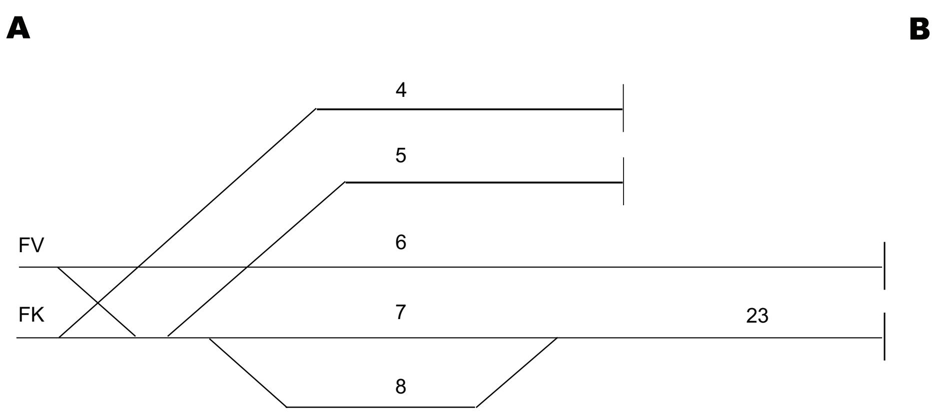

Figure 2: Part of the shuntyard at Lelystad

2.2.1 Tracks

At a shunt yard we have several tracks. For each track we know whether it is along a platform or not. We also know whether it may be used to park units or not. Further the physical lenght, whether it has catenary, a battery charger and other types of equipment is part of the input.

2.2.2 Routes

On the shunt yard there are several *routes*. A route is a feasible way to drive between two tracks. For each route we know its origin and destination track. Also the switches and possibly other tracks that are lying on the route are part of the input.

We distinguish several types of routes. A *single route* has the property that apart the two specific tracks which form the origin and destination of the route no tracks are passed when driving the route. In the graph of figure 2 there is a single route between track `FK` and `7`. A *composed route* passes other tracks on the way between its origin and destination. The route between `FK` and `23` via track `7` is an example of a composed route. The tracks that are passed by a composed route are called *via tracks*.

To introduce the next type of route, consider the following example: As one can see in figure 2 it is not possible to get from track `7` to track `8` in one forward movement. Out of track `7` one first has to drive in one direction, for example to track `23` at the right. At that track the driver walks to the other side of the train and drives in the other direction, to track `8`. Routes that need such a change of direction are called *sawing routes*. The track at which the direction changes is called the *saw track*. In the example, track `23` is the *saw track* of the route.

There are in general several routes available between two tracks. One is indicated as the *preferred route*, the others are called *alternatives*. They are numbered in decreasing preference.

2.2.3 Sides

Each station has an A-side and a B-side. Given these sides, we define the A-side of a track as the side which is closest to the A-side of the station, and similarly the B-side of a track. A track can be accessed from the A-side, the B-side or both sides. A side at which a track can be entered or left is called an *open side* of the track. In the figure 2 tracks `4, 5, 6` and `23` are open at the A-side and track `7` and `8` are open at both sides.

Moreover, we introduce the A-side of a train as the side of the train which is closest to the A-side of the station, whenever the train is within the boundaries of the station.

2.3 Timetables

Besides all information about the train services concerning units in the supply and their successors, we also need information about other passenger and possibly freight trains that have an influence on the shunt process.

For all such trains we know their planned times, the routes they take and the possible dwelling times on platform tracks. We also know their lenght at arrival and departure.

3 Output

Given a planning period and the input as described in last section a shunt plan contains decisions for all units that are in the supply for some time during that planning period. In this section we give a definition of a shunt plan and we treat restrictions and wishes to a shunt plan. These wishes come from outside and especially from downstream processes like crew scheduling.

3.1 Three ways to handle units in the supply

A unit enters the supply at the arrival time of its train service. After arrival it first occupies the platform track of this train service. We now treat several ways to handle a train unit that is in the supply:

1. The common way is to bring it to a non-platform track. Such tracks are mostly outside the platform area of a station, but also tracks parallel to platform tracks can be used. When the unit is needed again, it is brought to the platform track from where it departs. We assume that units do not change from park track meanwhile.

2. Another way is to hold the unit at its arrival platform track until it can be brought directly to its departure track.

3. Moreover, if its arrival and departure tracks are equal, it does not have to move between arrival and departure time if other activities on the platform track admit this.

The making of decisions at which track to park is called *parking*. Besides that it has to be decided at what time a unit drives from or to its park track and which route it takes. The making of these decisions is called *routing*.

If two units of a part are parked at different park tracks, we need a so called *part split*. There are two options: One option is to drive the whole part to the park track of one of the two units. Thereafter the other unit is decoupled and brought to its own park track. The other option is to split the part already at the platform track and bring both units at different times to their park tracks. We assume that the first option is not allowed, because it is in conflict with the assumption that units do not change from park track.

The decision which units are routed together is also (implicitly) contained in the routing decision.

Definition. *A shunt plan of a specific planning period consists of matching, parking and routing decisions for all units in the supply during the planning period.*

A typical planning period is between the rush hour of the evening and the morning rush hours of the next day.

3.2 Restrictions to the shunt plan

A shunt planner does not have unlimited freedom to make the matching, parking and routing decisions. There are some restrictions:

3.2.1 Matching

- For both arriving and departing units the type is prescribed. This makes that we can only match units that are of the same type.
- A unit must have enough time to drive between its arrival and departure platform track. We may only match units for which this time is large enough.
- In cases of matching parts, the units in both parts must have the same order.

To explain the last restriction consider an arriving composition (VIRM 4,VIRM 6). If at a later time composition (VIRM 6,VIRM 4) is scheduled to depart, then we can not simply match those parts, because their order is different. Only with certain additional shunt movements those units could be matched.

3.2.2 Parking

- The lenght of all units parked at a track may not exceed the lenght of the track.

- On every point in time, the unit on a track that has to depart first has to stand closest to that side of the track along which it leaves. Otherwise it is obstructed by other units and a crossing occurs.

The first restriction has to be extended in the cases where tracks are open at both sides and units can enter and leave via different sides. Then it is not always enough to check the free lenght on a track only: If a unit is parked on a track and another unit enters via one specific side, the free room on the track could be at the other side of the parked unit than the where the entering unit enters the track. The already parked unit then has to be repositioned on its track to let the other unit enter.

3.2.3 Routing

- The dwelling on the platform track before and after routing to/from the park track may not stress other train services in the timetables departing or arriving at that track. This gives restrictions to the time of routing. More about the dwelling time on the platform track can be found in section 7.

- If two units of a part are routed to/from different park tracks, then the order in which they are routed and the side they leave the platform track must be such that no crossings on the platform track occur.

- The route a unit takes may not conflict with the routes of trains from the timetables or with other shunting units. More about conflicts can be found in section 6.

- The via tracks of the route must be empty when the route is driven.

For non sawing routes there is a fixed duration of the routing time. This routing duration increases if a saw movement takes place, because by the change in direction the driver has to walk to the other side of the composition. At every saw track of the route some time is added to the total routing duration. This extra time depends on the lenght of the composition that is routed. A planner may also hold a composition for a longer time on a saw track for example to avoid conflicts on the second route part.

Maintenance of the infrastructure can influence the possibilities for shunt planner to use tracks and routes. However we do not consider the time intervals that switches or tracks are out of service and assume that they are available the whole planning period.

3.3 Wishes to the shunt plan

Some preferences of planners have to be taken into account. They are treated below.

3.3.1 Efficiency

In general we want to minimize the number of movements of compositions (shunt movements), because that saves driver and infrastructure capacity. This principle is reflected in the following ways:

- It is prefered to bring/retrieve the units of the same part together to/from the same park track.
- It is prefered to route units from different trains using the same platform track together to/from the same park track.
- It is prefered to route a unit directly from its arrival to its departure platform track above using a park track.
- It is prefered to match units arriving and departing at/from the same platformtrack if it is allowed to stand there meantime.

3.3.2 Robustness

Another important aspect to the objective is the flexibility of the plan and the resistance to small delays. It is better that a shunt plan acts as a buffer for delays and that delays are not passed on to subsequent trains.

A first way to achieve this is to maximize the minimal time the units stay on the shunt yard. Besides this, it is nice to have a plan that does not significantly change if some units arrive later then their planned time, for example by widely seperating the arrivals at and departures from each track. Otherwise small delays can cause a swap of train units on a track what can result in a crossing. Another way to prevent from crossings during executing of the shunt plan is to assign only one type of units to certain tracks, so that the order on these tracks is not important any more.

3.3.3 Other

There are also some other wishes.

- If a composition enters a park track where already units are parked, the driver has to pass a red-light sign. These situations are called *occupied* *track enterings*. They are contra-intuitive for drivers and cause that the train has to drive slowly. Therefore we try to avoid this situations.

- While chosing the park tracks and routes, we want to minimize:
  - The traveled distance by shunting units
  - The number of saw tracks on the route
  - The number of switches on the route to prevent from wear
  - The number of routes operated simultaneously in time

  The number of simultaneous routes in time can be used as a proxy for the minimum number of drivers needed. Solutions with less simultaneous routes increase the chance of finding a good solution for the crew planning problem.

- As described on page 17 it is allowed to park at platform tracks. This is only possible if the other train services admit this. The advantage of this is that it saves one or two shunt movements. However, for some platform tracks planners do not like to park units on it at daytime, because platform tracks can also be used to handle delayed or redirected trains. Thus during daytime one wants some tracks to be empty as much as possible.

3.4 Dependencies of the shunt plan

The matching, parking and routing decision are interrelated parts of the overall planning problem that a shunt planner faces. In practise, planners are mostly unaware of such a decomposition because of these relations. The most important relations are:

- Matching and parking. The result of the matching determines when physical train units are available for parking, and when these should leave the station again. Moreover, if the time difference between arrival and departure of a train unit is sufficiently small, parking is not required.

- Matching and routing. The minimum time difference in a matching of an arriving unit to a departing unit is among others determined by the routing time from the arrival platform to the departing platform.

- Routing and parking. The routing effort (avoiding conflicts) influences preferences for certain tracks over other tracks for parking train units.

- Matching, parking and routing: avoiding crossings. The arrival times on a specific park track define the order in which physical units are parked there. This determines the order in which they can leave their park track, which is restricted by the matchings that are made.

4 Summary part I

The goal is to model the shunt problem, which consists of making a shunt plan for a specific planning period at a specific station given the timetables, the rolling stock circulation and infrastructural lay-out. After a shunt plan is made we know

- for every arriving unit to which departing unit it is matched
- for every unit where it parks
- for every unit at what time it drives between its platform track and park track and which route it takes

So we know for every unit in the supply at every minute where it is located.

A shunt plan is feasible if

- no crossings occur
- no conflicts occur

A shunt plan is good if

- the number of shunt movements of compositions is small
- it is resistant to small delays
- in daytime some platform tracks are free as much as possible
- the route effort (the number of switches and changes in direction) is small
- the number of drivers that are needed to carry out the shunt movements is minimized
- the number of occupied track enterings is small

The assumptions we made in the last sections are:

- Units do not change from park track during their time in the supply (remind that the dwelling times on the platform track just before departure or just before arrival do not count as parking).
- No units are cleaned.
- Parking on a platform track is only allowed if the unit departs from that platform track.

Part II
Quantitative description

In this part we introduce formal notations for the terms introduced in last part. This gives us the possibility to give a more extensive description of the concept of conflicts, see section 6. Besides that, in section 7, the shunt window is introduced and calculated.

5 Notations

This section is dedicated to the introduction of notations.

5.1 Graph

A *shunt yard* of a station consists of a set of *tracks* $S = \{s_1, s_2, ..., s_n\}$ and a set of *switches* $W = \{w_1, w_2, ..., w_m\}$ that are connected to each other. The set of all platform tracks is denoted by S_{pl}, the set of tracks where trains can be parked by S_p. Note that in general $S_p \cap S_{pl} \neq \emptyset$. The lenght of a track $s \in S$ is denoted by l^s.

In figure 3 a part of the shuntyard of Lelystad is represented as a graph (V, E). Dots represent switches and open circles represent tracks. In this graph every switch and track appears twice as a vertex ($V = S \cup S' \cup W \cup W'$). The first vertex represents arriving from the 'left' and departing to the 'right' side of the vertex. Vice versa, the second vertex represents arriving from the 'right' and departing to the 'left' side of the vertex. Both vertices representing a track are connected with two directed arcs whenever the track they represent can be used as saw track. For all vertices $v_i, v_j \in V$ for wich a train can drive between v_i and v_j in that direction an edge (v_i, v_j) is added to the set E of edges.

5.2 Routes

A route R in the shunt yard can be expressed as a path in the graph mentioned in subsection 5.1. Given tracks s_i and s_j, a path in graph (V, E) between track s_i and track s_j is a sequence of nodes $(v_1, v_2, \ldots v_k)$, $v_1, \ldots, v_k \in V$ for which v_1 equals $s_i \in S$ or its duplication $s'_i \in S'$, v_k equals $s_j \in S$ or $s'_j \in S'$ and $(v_m, v_{m+1}) \in E$ for all $m = 1, \ldots, k-1$. Note that the direction of the edges present in the graph has to be respected.

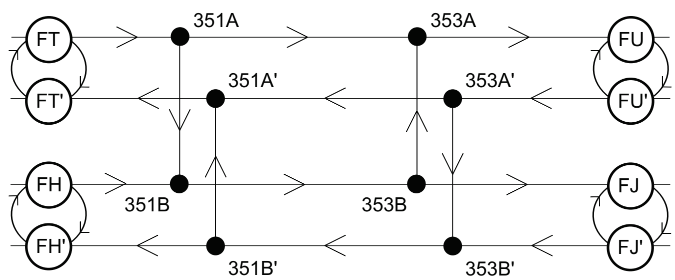

Figure 3: Graph

As a consequence single routes use paths for which hold $v_i \in W \cup W'$ for $i = 2, \ldots, k-1$. For composed routes, the associated paths have one or more vertices $v_i \in V$ $(i = 2, \ldots, k-1)$ with $v_i \in S \cup S'$. The paths $(v_1, \ldots v_k)$ for which all $v_1, \ldots, v_k$ are either in $S \cup W$ or in $S' \cup W'$ represent non-sawing routes.

For a route R we denote by S_R the via tracks of R and by W_R the set of switches of R. The origin and destination of route R are denoted by o_R and d_R respectively. The set of all non sawing routes is denoted by $\mathcal{R}$.

5.3 Relations

An ordered pair of tracks (s_i, s_j), $s_i, s_j \in S$ is called a *relation*. When a train drives a route R with origin $o_R = s_i$ and destination $d_R = s_j$ it is said to 'serve relation (s_i, s_j)'. There are in general several routes available to serve a relation (s_i, s_j). One is indicated as the *preferred route* the others are called *alternatives*. They are numbered in order of decreasing preference.

Two routes R^k and R^l in $\mathcal{R}$ are *disjunct* if both intersections $S_{R^k} \cap S_{R^l}$ and $W_{R^k} \cap W_{R^l}$ are empty.

Example. *Consider the relations* `(FT,FU)` *and* `(FH, FJ)` *in figure* 3. *Route* R *with* $S_R =$ `(FT, 351A, 353A, FU)` *serves relation* `(FT,FU)` *and route* R' *with* $S_{R'} =$ `(FJ, 353B, 351B, FH)` *serves relation* `(FJ,FH)`. *These routes are disjunct because*

$$S_R \cap S_{R'} = \{\texttt{FT, 351A, 353A, FU}\} \cap \{\texttt{FJ, 353B, 351B, FH}\} = \emptyset. \square$$

Whether or not two paths are disjunct is important when treating conflicts, see section 6.

5.4 Supply

The *supply* is represented by a set U^+ containing (abstract) *arriving* units and a set U^- containing (abstract) *departing* units. These sets are deduced from the rolling stock circulation and do not contain units that have a transition (see 2.1.4). The set U is the union of all arriving and departing units $U^+ \cup U^-$. For each unit $u \in U$ its type is denoted by ψ_u, its length by l_u, and the train service it belongs to by t_u. Furthermore, the platform track train unit u arrives or departs is denoted by pl_u. The time at which this train service is planned at pl_u is denoted by τ_{t_u}. For arriving units u this time represents the time at which t_u arrives at pl_u and for departing units v this time represents the time at which t_v departs from pl_u.

These times are used to determine conflicts between train services or between train services and routed units.

5.5 Matching

Given the sets U^+ and U^- a matching is an assignment of units of U^+ to units of U^-. The departing unit to which a $u \in U^+$ is matched is called $d(u)$.

After a matching is made, we know for an arriving unit in which train service it departs. In other words: if u and v are matched, then the physical unit that is represented by u, entering the station in train service t_u, will also get unit v and leaves the station in train service t_v.

Remind that units u and v may only be matched if $\psi_u = \psi_v$ and $\tau_{t_u} + T \leq \tau_{t_v}$. This leads to the set Q of all possible matchings:

$$Q := \{(u, v) | u \in U^+, v \in U^-, \tau_{t_u} + T \leq \tau_{t_v}, \psi_u = \psi_v\} \tag{1}$$

5.6 Parking

For each unit $u \in U^+$ we call the track where it gets parked $pt(u)$, the park track of u. For departing units $v \in U^-$ the track $pt(v)$ represents the track from which v is retrieved.

If units u and v are matched, then $pt(u) = pt(v)$ must hold so that the physical unit can drive the successive relations $(pl_u, pt(u))$ and $(pt(v), pl_v)$.

Remark: As described at page 17 it is also allowed to directly bring a unit u directly to the departure track of the train service of the unit v to which it is matched. Then $pt(u) = pl_v$ and we get the relation (pl_u, pl_v). Because $pt(u) = pt(v)$ must hold, for convenience $pt(v)$ must also be equal to pl_v in this case.

5.7 Routing

We denote by $R(u)$ the route which a unit $u \in U$ has to drive to its park track. The vector $T(u)$ contains the departure times of the non sawing route parts of route $R(u)$. For $u \in U^+$ the first element of $T(u)$ is called the *shunt time* $\tau(u)$. This is the departure time from pl_u to $pt(u)$. We denote the arrival time on $pt(u)$ by $\theta(u)$. For $v \in U^-$ the notation is the other way around: The departure time from $pt(v)$ is denoted by $\theta(v)$ and the arrival time on pl_v, the shunt time, by $\tau(v)$.

5.8 Timetables

We define the set $\mathcal{T}$ of all train services in the timetables of the station under consideration. For every train service $t \in \mathcal{T}$ we denote its lenght by l_t and its platform track by pl_t. The times τ_t and τ'_t represent the arrival and departure times to and from pl_t. The routes to and from pl_t are denoted by R_t and R'_t.

Given an ending train service t that has a transition, the successor of t is denoted by s_t. On the other hand, given a starting train service t, if there is a transition to t, the predecessor of t is denoted by p_t.

Given an arriving unit u the successor train service s_{t_u} is important to determine the possible shunt times for unit u. Section 7 is about the calculation of the set of possible shunt times, the shunt window. Before that, in section 6 more detailed information about conflicts is given.

6 Time Interval Restrictions and Conflicts

In this section the concept of time interval restrictions and conflicts is explained. Time interval restrictions are minimal time differences between two events, like driving two non-disjunct routes, but they also appear as minimal time durations for processes, like coupling of units. They are explained in the first subsection. In 6.2 an extensive example is presented where the time interval restrictions are 'in action'.

6.1 Time interval restrictions

The time interval restrictions, represented in rounded minutes, may differ per station. Below a list of all sort of time interval restrictions is given for one specific station. These values are used in the examples throughout this report.

This list is split up in minimal time durations for processes, safety restrictions for two trains visiting the same platform track and safety restrictions for two trains driving two non-disjunct routes.

The first column is a Dutch description about for which process the time interval restriction holds, the second contains the translation to English and the last collumn gives the value of the minimal time duration at a specific station.

Uitstaptijd	Passengers stepping from the train	3
Instaptijd	Passengers stepping into the train	3
Combineren	Coupling of units	3
Splitsen	Decoupling of units	2
OmbouwtijdZaagbewegingen	Change of direction on a sawing route	4

In the next table minimal times between the planned times (measured at the platform track) of two trains at the same platform track are presented.

AnaAUitDezelfdeRiRR	Two arrivals via the same side	3
VnaVNaarDezelfdeRiRR	Two departures via the same side	3
AnaVOpZelfdeSpoorRR	Arrival after departure from different sides	3
OverkruisAnaVRR	Arrival after departure at the same side	4

In the next table minimal times between the planned times of two trains at different platform tracks using two non-disjunct routes are presented.

OverkruisAnaARR	Arrival after arrival	3
OverkruisVnaVRR	Departure after departure	3
OverkruisAnaVRR	Arrival after departure	4
OverkruisVnaARR	Departure after arrival	0

Note that for ‘arrival after departure’, it does not make a difference whether the departing and arriving trains use the same platform track or not. Both are characterized by ‘OverkruisAnaVRR’.

These restrictions do not only hold for train services, but also for shunting units. In the remainder they are notated with $tir(\dots)$, for example the last row of the list above results in: $tir(OverkruisVnaARR) = 0$. For freight trains the lenghts of the intervals are in general higher.

Given a departure time t of a non sawing route part, the associated arrival time used in the calculations always equals $t + 2$.

6.2 Conflicts

If a shunt plan contains a routing composition that violates the time interval restrictions with an already scheduled train or another routing compositions, then the plan contains a *conflict*. In this subsection we give examples of conflicts.

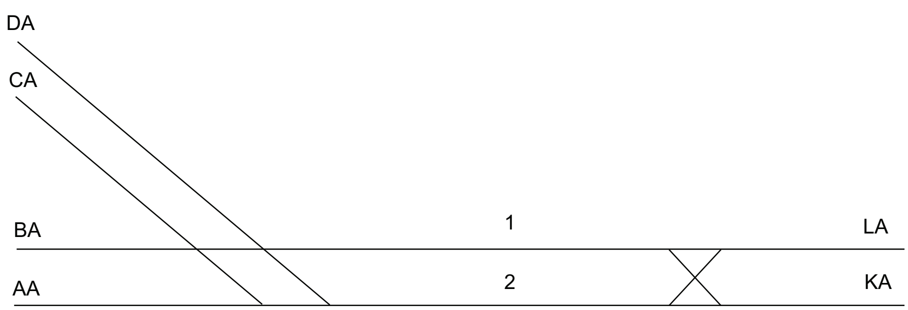

Figure 4: Conflict example

Example. *Consider relations* `(ca,2)` *and* `(aa,2)` *in figure 4. Assume that train services t_1 and t_2 serve these relations. If τ_{t_1} equals 12.01, then train t_2 may not be planned between 12.01 - $tir(AnaAUitDezelfdeRiRR)$ and 12.01 + $tir(AnaAUitDezelfdeRiRR)$. That means: only at 12.04 and later, or at 11.58 and earlier. If t_2 would be planned between these times, it has a conflict with t_1.*

Assume that we choose τ_{t_2}=12.04 and that the trains get coupled. The coupled train may only leave at 12.04 + $max(tir(Combineren), tir(Instaptijd))$, *because coupling and stepping into the train can take place simultaneously. Thus at 12.07 or later the coupled train may depart.*

Assume that we choose to depart at 12.07 from track `2` *and that the next train entering the station* t_3 *will serve relation* `(la,2)`*. This is allowed from 12.07 +* $tir(OverkruisAnaVRR)$ *on. If* t_3 *would have come from* `aa`*, then 12.07 +* $tir(AnaVOpZelfdeSpoorRR)$ *is the earliest arriving time.*

Assume that t_3 *departs at* τ'_{t_3}*=12.11 and serves relation* `(2,da)`*. Then train* t_4 *serving* `(ba,1)` *may only be planned at 12.11-*$tir(OverkruisVnaARR)$ *and earlier or at 12.11+*$tir(OverkruisAnaVRR)$ *and later. A train* t_5 *departing from* `1` *to* `ba` *may only depart at 12.11-*$tir(OverkruisVnaVRR)$ *and earlier or at 12.11+*$tir(OverkruisVnaVRR)$ *and later.* □

7 Shunt window

In general, platform tracks are highly used at daytime. In the overall schedule, passenger and freight trains passing through the station are also planned at several platform tracks. This gives restrictions to the occupation of platform tracks by shunt units. In this section we treat the derivation of the possible shunt times of a unit entering the supply in more detail.

The most important aspect that influences the possible shunt time is the side at which unit u leaves or enters pl_u. Therefore for every u we calculate in general two shunt windows, for both leaving sides of pl_u one. First we repeat the most important definitions.

7.1 Definitions on the platform track

Remind that the train in which a unit $u \in U^+$ arrives or a unit $u \in U^-$ departs is called t_u. For a $u \in U^+$ the time at which t_u arrives is denoted by τ_{t_u}. For units $u \in U^-$ this number represents the departure time of t_u.

For a unit $u \in U^+$ there are two cases: The first one is that all units of t_u are members of U^+. In the second case other units of the same train are already planned to depart in a train from the same platform track. The train in which these units depart is the *successor* of t_u and denoted by s_{t_u}. In the same way the precedessor p_{t_u} of t_u for a $u \in U^-$ is defined as the train in which the units that are scheduled in t_u and do not belong to U^- arrive.

Consider a unit $u \in U^+$. The trains arriving at pl_u after t_u arrived are denoted by $r^0_{t_u}, r^1_{t_u}, \ldots$, or simply $r^0, r^1, \ldots$. They are numbered in the order in which they arrive: $\tau_{r^i} < \tau_{r^{i+1}}$. For units in U^- train services $r^0, r^1, \ldots$

represent the trains departing from pl_u before the arrival of p_{t_u} where r^0 is the latest train that departed before the arrival of p_{t_u}: $\tau_{r^i} > \tau_{r^{i+1}}$.

7.2 Definition shunt window

The shunt window of a unit u at side S is defined as the set of all possible shunt times for leaving or departing pl_u via side S given the train services $t \in \mathcal{T}$ with $pl_t = pl_u$, neglecting the train services from the timetable with $pl_t \neq pl_u$ and decisions made during the making of the shunt plan.

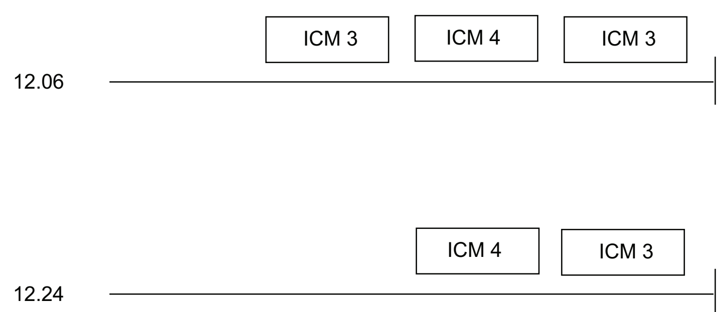

Figure 5: Example 1

Example 1. *Consider composition* (ICM 3, ICM 4, ICM 3) *entering a dead-end track as train t at 12.06 (see figure 5). Assume that the rolling stock circulation prescribes that the left most unit has to depart at 12.24 from the same platform track ($\tau_{s_{t_u}} = 12.24$). The beginning of the shunt window is at 12.24+$tir(VnaVNaarDezelfdeRichtingRR)$=12.27.* □

7.3 Determination of the shunt window

As mentioned, the shunt window of u depends on several aspects. We now focus on arriving units. For departing units, the ideas are the same, but other terms and time interval restrictions are needed.

- The lower bound of the shunt window depends on whether there is a successor s_{t_u} or not (if all units of t_u enter the supply).

- If there is a successor, the side of the train at which unit u is decoupled becomes important. This determines the lowerbound of the shunt window and in some cases also the upperbound.
- If the upperbound is not determined from s_{t_u} or there is no s_{t_u}, then the upperbound of the shunt window can be determined from $r^0_{t_u}, r^1_{t_u}, \ldots$.

7.3.1 Lowerbound

I. If all units of t_u enter the supply, the shunt window for both sides starts at $\tau_{t_u} + tir(Uitstaptijd)$.

II. If there is a successor s_{t_u} and unit u is decoupled from the back side of s_{t_u}, its shunt window at that side starts at

$$\tau_{t_u} + max(tir(Uitstaptijd), tir(Splitsen))$$

and at the other side at

$$\tau_{s_{t_u}} + tir(VnaVNaarDezelfdeRiRR).$$

If u is decoupled from the front side of s_{t_u}, the shunt window for that side starts at

$$\tau_{t_u} + max(tir(Uitstaptijd), tir(Splitsen)).$$

In this case also an upperbound exists:

$$\tau_{s_{t_u}} - tir(VnaVNaarDezelfdeRiRR)$$

It is not possible to leave via the other side in this case, so the shunt window at that side is empty.

Example 2. *Consider an arriving train service t_u with $\tau_{t_u} = 11.46$. Assume that s_{t_u} departs from the A-side at 12.14. Taking $tir(VnaVNaarDezelfdeRiRR) = 3$ the lowerbound of the shunt window of u for the A-side equals 12.17. If s_{t_u} departs from the B-side the shunt window of u equals*

$$[\tau_{t_u} + max(tir(Uitstaptijd), tir(Splitsen)),$$

$$\tau_{s_{t_u}} - tir(VnaVNaarDezelfdeRiRR)] =$$

$$[11.49, 12.11]$$

In this case for the A-side there are no feasible shunt times. □

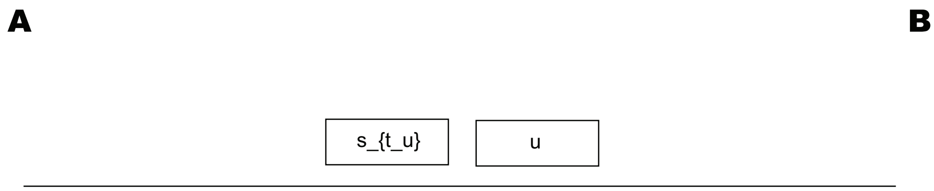

Figure 6: Example 2

7.3.2 Upperbound

For all cases the **lowerbound** of the shunt window has been specified. If the units are not decoupled from the front side, the **upperbound** is influenced by the next trains arriving at pl_u. We consider their arrival side, leaving side and lenght. If those trains do not give an upperbound or there are no next trains on pl_u, the upperbound is defined as the end of the planning period.

$\mathcal{I}$. If all $r^i_{t_u}$, (in short: r^i) enter and leave via the same side as s_{t_u} has left (say A-side) then unit u may stand at the B-side if its lenght admits. Let r' be the first train service that by its length forces unit u to leave the platform track. The upperbound for the leaving via the A-side equals

$$\tau_{r'} - tir(OverkruisAnaVRR)$$

and via the B-side equals

$$\tau_{r'} - tir(AnaVOpZelfdeSpoorRR).$$

Note that if u is the only unit that is decoupled from t_u, train service r' is the first r^i with $length(r^i) + l_u > length(pl_u)$. If there are more units from t_u that have to be shunted, then also the length of the units u' with $t_u = t_{u'}$ that are more close to the B-side than u has to be taken into account. This is explained in the next example.

Example 3. *Consider units u and u' arriving in the same train (see figure 7). Assume that s_{t_u} leaves via the A-side and that r^0 enters and leaves via the A-side. Whether or not unit u must have gone before the arrival of $r^0_{t_u}$ depends besides the lenghts of r^0 and u and the physical lenght of pl_u also on the lenght of u', even if it has been driven away already.* $\square$

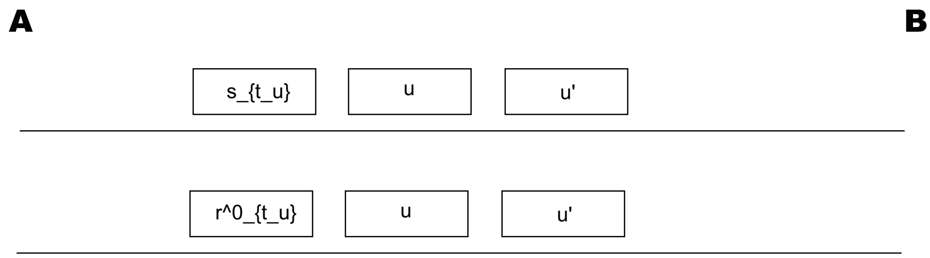

Figure 7: Example 3

If there is no s_{t_u} the side of the platform track at which u stands until $\tau(u)$ may be chosen. We choose the side of the platform track different from the side at which the train service r^0 arrives.

Example 4. *Consider complete ending train t_u without successor. Assume that*

- r^0 *enters and leaves via the B-side.*
- r^1 *enters and leaves via the A-side.*

The assumption in the first bullet point makes that u stands at the A-side. Therefore before the arrival of r^1 (via the A-side), u has to leave. □

$\mathcal{II}$. If all r^i leave and enter via the same side but those sides are not the same for all trains, then the upperbound is determined from the first train that enters via the other side than s_{t_u} has left. If u does not have a successor, it is determined from the first train that enters via the other side than r^0 entered.'

$\mathcal{III}$. If there is a r^j that leaves via the other side then it entered, this r^j gives an upperbound, because r^j passes the whole track and this would cause a crossing if u would have stand there.

Example 3 continued. *If r^0 would enter via the B-side, then the upperbound of the shunt window equals $\tau_{r^0} - tir(OverkruisAnaVRR)$ for u leaving via the A-side and $\tau_{r^0} - tir(AnaVOpZelfdeSpoorRR)$ for u leaving at the B-side.* □

Part III

Shunt planning at NSR

In section 1 we gave an idea what makes shunt planning difficult and in subsection 3.4 we showed that the three subdecisions matching, parking and routing are strongly related. This argues for an integrated approach.

In this part a broader foundation for the aim of an integrated approach is given. After a treatment of the algorithms already in development by NSR to make the shunt plan in sections 8-10 we analyze this current approach in section 11. That section treats the added value of an integrated approach to the algorithms described. A summary of this part is presented in section 12.

8 Making the shunt plan

The algorithms for making the shunt plan that are currently in development work in the following way: Firstly the matching and parking are decided integrally. Thereafter the decision of the routes and shunt times is made. For both steps there is a method to determine the correct decisions.

The first method is called OPG. This is a mixed integer program (MIP) that decides the parking and matching for every unit. The plan resulting from these decisions is called the *emplacing plan*. The plan is calculated under the assumption that arriving units enter the park tracks in the order they enter the station in the train service they belong to and that departing units leave the park tracks in the order they leave the station in the train service they belong to. The process between entering the station and entering the park track is neglected in this stage. One of the most important issues of OPG is to avoid crossings on park tracks.

From the emplacing plan several relations follow: the arriving unit $u \in U^+$ has to drive from pl_u to $pt(u)$ and at a later time from $pt(u) = pt(d(u))$ to $pl_{d(u)}$. Sometimes it can also be routed from pl_u to $pl_{d(u)}$ directly, or stay at pl_u if it is the same as $pl_{d(u)}$.

Routes and shunt times within shunt windows have to be found which can serve these relations. This is done afterwards by the second method, called TIMEFIXER. This model treats the relations given by OPG as fixed. The main decision that is made in this model is the shunt time. Within OPG this time was assumed at a specific point in time. By deciding the real shunt times we have to take care of the other activities on the platform tracks. Also we have to ensure that two units assigned to the same park track preserve the order in which they are planned in OPG if they are not of the same type. The main issue of TIMEFIXER is to avoid conflicts. The output of TIMEFIXER gives a shunt plan.

9 OPG

In this section we explain the first method, OPG. This model is currently in development. We treat a basic version as described in the PhD-thesis of R.M. Lentink, see [Len06]. Some extensions and additions made during further development are treated in subsection 11.3.

In 9.1 the mixed integer program of OPG for tracks open at only one side is treated. We treat the in- and output, the variables, sets, paramters and constraints. This model has been extended for tracks open at both sides. The description of this model can be found in subsection 9.2.

9.1 Basic model for tracks open at one side

In this subsection it is assumed that a park track is always entered via one side, the A-side. In the next subsection we will treat cases with tracks open at both sides. Given the times τ_{t_u} of all units $u \in U$ from the timetables this assumption gives rise to a partial ordering $<_A$ on the units in U:

We define that for two units u_i and u_j we have $u_i <_A u_j$ if and only if one of the following conditions is satisfied:

1. Unit u_i arrives or departs in a train with an earlier planned time than the train to which unit u_j belongs ($\tau_{t_{u_i}} < \tau_{t_{u_j}}$).

2. Arriving units u_i and u_j arrive in the same train ($t_{u_i} = t_{u_j}$) and u_j is closer to the A-side of the train than u_i.

3. Departing units u_i and u_j depart in the same train ($t_{u_i} = t_{u_j}$) and u_i is closer to the A-side of the train than u_j.

Example. *Consider an arriving train composition* $(u_1, \ldots, u_k)$*, where* u_1 *is closest to the A-side of the train and* u_k *is farthest from it. Case 2 states that it is ordered as* $u_k <_A \cdots <_A u_1$ *in* U*. Note that this is the order in which the units arrive at a park track open at the A-side. When this composition leaves the park track, the order of the physical units changes: Consider departing composition* $(v_1, \ldots, v_k)$ *with* $d(u_i) = v_i$ $(i = 1 \ldots k)$*. Then the* v_i *are ordered as* $v_1 <_A \cdots <_A v_k$*.* □

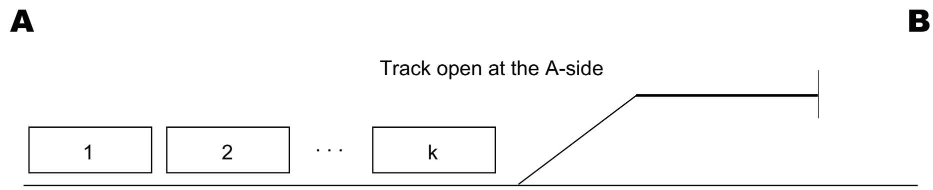

Figure 8: Ordering example

Input

The input for a specific station consists of the set of tracks S_p at which it is allowed to park and a set of platform tracks S_{pl}. These sets can overlap. The set of all possible relations Z equals $\{(s_1, s_2) \mid s_1 \in S_{pl}, s_2 \in S_p\}$. For each track s its length l^s is part of the input. Note that other features of the shunt yard (as switches and routes) are not part of the input. It is assumed that at the beginning and at the end of a planning period the tracks in S_p are empty. Note that this assumption can be relaxed easily by adding dummy arriving units at the beginning of the planning period and dummy departing units at the end of the planning period.

Next to this the input consists of the sets of units U^+ and U^-. For each unit u in U the length l_u and type ψ_u is known. Moreover the platform track pl_u and τ_{t_u} is known. From these times the partial ordering $<_A$ is derived. The sets U^+ and U^- are such that on every moment in time for every type no more units have departed than have arrived until twenty minutes ago.

Output

The output is a matching for every unit in U^+ to a unit in U^- and an assignment to a track in S_p such that for every track there is no overusage. The outcome prevents that crossings occur on the tracks in S_p. This is called the *emplacing plan*. The goal is to make an optimal emplacing plan by minimizing the number of part splittings and the number of different types at each track. The output is given by the values of the decision variables. These decision variables are restricted and connected to each other by linear constraints which are treated below.

M-variables

The main variabele of the model is $M_{u,v,s}$. It is a decision variable that is one if unit $u \in U^+$ is matched to unit $v \in U^-$ and parked meanwhile at track $s \in S_p$ and zero otherwise. Thus $M_{u,v,s} = 1$ iff $d(u) = v$ and $pt(u) = pt(v) = s$.

To avoid crossings on tracks open at the A-side only, one has to follow last-in-first-out discipline. Two physical units that do not meet this discipline may not be parked together at the same track. Remind that Q is the set of pairs of train units that can be matched. The set of pairs of matchings that would cause a crossing if they are parked on the same track, called $\mathcal{A}$, is defined as

$$\mathcal{A} = \{((u,v),(u',v')) \mid (u,v),(u',v') \in Q, u <_A u' <_A v <_A v'\}$$

The statement $u' <_A v$ between arriving unit u' and departing unit v means that $\tau_{t_{u'}} < \tau_{t_v}$.

Using this set the following constraints to prohibit crossings on tracks open at the A-side only are made:

$$M_{u,v,s} + M_{u',v',s} \leq 1 \quad \forall s \in S_p, \forall[(u,v),(u',v')] \in \mathcal{A}$$

P-variables

The variable $P_{u,s}$ is linked to the M-variables. It is one if unit u is parked at or retrieved from track s ($pt(u) = s$). Otherwise it equals zero. It is linked to the M-variables in the following way:

$$\sum_{v:(u,v)\in Q} M_{u,v,s} = P_{u,s} \quad \forall u \in U^+, s \in S_p$$

$$\sum_{u:(u,v)\in Q} M_{u,v,s} = P_{u,s} \quad \forall v \in U^-, s \in S_p$$

The constraints to ensure that every unit is assigned to exactly one track are:

$$\sum_{s\in S_p} P_{u,s} = 1 \quad \forall u \in U$$

L-variables

To ensure that no track is overused in length the variable $L_{u,s}$ is linked to the P-variables. It is the length of the units at track s after the departure or arrival of unit u. It is administrated after every arrival and departure of a unit. It is linked to the P-variables in the following way:

$$L_{u,s} = L_{u-1,s} + l_u P_{u,s} \quad \forall u \in U^+, s \in S_p$$
$$L_{u,s} = L_{u-1,s} - l_u P_{u,s} \quad \forall u \in U^-, s \in S_p$$

where unit $u-1$ is a unit with $u-1 <_A u$ and $\tau_{t_{u-1}}$ maximal.

The constraints needed are:

$$L_{u,s} \leq l^s \quad \forall u \in U^+, s \in S_p$$

Note that these restrictions only need to be checked after arrival of units.

Number of types on a track

In 3.3.2 we concluded that we want to minimize the number of different types at each track. Therefore the variable E_s is derived. It is the number of different types in excess of one parked at track s. To link E_s to the other variables, the variabele $O_{\psi,s}$ is needed. It equals one if at least one unit of type ψ is on track s. Otherwise it equals zero. The O-variables are linked to the P-variables in the following way:

$$P_{u,s} \leq O_{\psi_u,s} \quad \forall s \in S_p, u \in U$$

Then the E-variables can be derived from the O-variables:

$$\sum_{\psi \in Y} O_{\psi_u,s} \leq E_s + 1 \quad \forall s \in S_p$$

Part splittings

In 3.3.1 we concluded that we want to minimize the number of part splittings. Therefore we introduce the variable K_u that is one if units u and u' belong to the same part and are parked at or retrieved from different tracks. Otherwise it equals zero. Therefore we need to know which units belong to the same train. The set I is defined as the set of pairs of adjacent units $\{(u,u') \mid t_u = t_{u'},\ u,u' \in U\}$ that arrive or depart in the same train service. Now we can formulate the constraints:

$$K_u \geq P_{u,s} - P_{u',s} \quad \forall s \in S_p, (u, u') \in I$$

Objective

We have seen two aspects that influence the preferences of the output. It is possible to give each aspect its own weight w_i. We then get the objective:

$$w_1 \sum_{t \in U} K_t + w_2 \sum_{s \in S_p} E_s$$

Model summary

We now present the whole MIP model in a summary:

$$P_{u,s} = \begin{cases} 1 & \text{if pt(u) = s} \\ 0 & \text{otherwise.} \end{cases}$$

$$L_{u,s} = \begin{cases} \text{The length of the units at track } s \text{ after the departure or} \\ \text{arrival of unit } u. \end{cases}$$

$$M_{u,v,s} = \begin{cases} 1 & \text{if } u \in U^+ \text{ has } d(u) = v \text{ and } pt(u) = pt(v) = s \\ 0 & \text{otherwise.} \end{cases}$$

$$K_u = \begin{cases} 1 & \text{if units } u \text{ and } u' \text{ are related tot the same train and} \\ & \text{are parked at or retrieved from different tracks;} \\ 0 & \text{otherwise.} \end{cases}$$

$$O_{\psi,s} = \begin{cases} 1 & \text{if at least one unit of type } \psi \text{ is parked at track } s; \\ 0 & \text{otherwise.} \end{cases}$$

$$E_s = \text{the number of types in excess of one parked at track } s.$$

Objective and constraints:

$$\text{Minimize } w_1 \sum_{u \in U} K_t + w_2 \sum_{s \in S_p} E_s \tag{9.2}$$

subject to

$$\sum_{s \in S_p} P_{u,s} = 1 \qquad \forall u \in U \tag{9.3}$$

$$\sum_{v:(u,v)\in Q} M_{u,v,s} = P_{u,s} \qquad \forall u \in U^+, s \in S_p \tag{9.4}$$

$$\sum_{u:(u,v)\in Q} M_{u,v,s} = P_{v,s} \qquad \forall v \in U^-, s \in S_p \tag{9.5}$$

$$M_{u,v,s} + M_{u',v',s} \leq 1 \qquad \forall s \in S_p, \{(u,v),(u',v')\} \in \mathcal{A} \tag{9.6}$$

$$L_{u,s} = L_{u-1,s} + l_u P_{u,s} \qquad \forall u \in U^+, s \in S_p \tag{9.7}$$

$$L_{v,s} = L_{v-1,s} - l_v P_{v,s} \qquad \forall v \in U^-, s \in S_p \tag{9.8}$$

$$L_{u,s} \leq l^s \qquad \forall u \in U^+, s \in S_p \tag{9.9}$$

$$K_u \geq P_{u,s} - P_{u',s} \qquad \forall s \in S_p, (u,u') \in I \tag{9.10}$$

$$P_{u,s} \leq O_{\psi_u,s} \qquad \forall s \in S_p, u \in U \tag{9.11}$$

$$\sum_{\psi \in Y} O_{\psi_u,s} \leq E_s + 1 \qquad \forall s \in S_p \tag{9.12}$$

The same can be done for tracks open at the B-side only. The ordering $<_B$ is made:

We define that for two units u_i and u_j we have $u_i <_B u_j$ if and only if one of the following conditions is satisfied:

1. Unit u_i arrives or departs in a train with an earlier planned time than the train to which unit u_j belongs.
2. Arriving units u_i and u_j arrive in the same train and u_j is closer to the B-side of the train than u_i.
3. Departing units u_i and u_j depart in the same train and u_i is closer to the B-side of the train than u_j.

Using this ordering the set $\mathcal{B}$ of possible crossings at tracks open at the B-side is made, $\mathcal{B} = \{((u,v),(u',v')) \mid (u,v),(u',v') \in Q, u <_B u' <_B v <_B v'\}$. Again, the statement $u' <_B v$ between arriving unit u' and departing unit v means that $\tau_{u'} < \tau_v$. The constraints of type (9.6) holding for tracks open at the A-side are supplemented by constraints

$$M_{u,v,s} + M_{u',v',s} \leq 1 \quad \forall \{(u,v),(u',v')\} \in \mathcal{B} \tag{9.13}$$

for tracks $s \in S_p$ that are open at the B-side only.

9.2 Basic model with tracks open at both sides

For tracks open at both the A-side as the B-side an extra decision about the entering side of the track is included:

$$S_u = \begin{cases} 0 & \text{if train unit } u \text{ arrives or departs via the } A\text{-side} \\ 1 & \text{if train unit } u \text{ arrives or departs via the } B\text{-side} \end{cases}$$

Because for a unit u we do not know in advance whether $pt(u)$ is open at the A or B side only, or at both sides, this variable is introduced for all units. This causes that extra constraints are needed to forbid units to enter a track at a side at which it is not open:

$$P_{u,s} + S_u \leq 1 \text{ for all units } u \text{ and tracks } s \text{ open at the } A \text{ side only} \quad (9.14)$$

$$P_{u,s} - S_u \leq 0 \text{ for all units } u \text{ and tracks } s \text{ open at the } B \text{ side only} \quad (9.15)$$

Parts of one unit

For ease of discussion we treat the case that all parts consits of one unit. In this case the orderings $<_A$ and $<_B$ are equivalent. Consider units u_i and u_j with $\tau_{t_{u_i}} < \tau_{t_{u_j}}$. The definition of $<_A$ at page 36 gives that u_i and u_j are ordered in the following way: $u_i < u_j$. The definition of $<_B$ at page 41 gives that also $u_i <_B u_j$. So in the following constraint $<_A$ can be replaced by $<_B$.

For tracks open at both sides the crossing constraints for parts consisting of exactly one unit are derived from the following rules:

$$\text{if } M_{u,v,s} = 1 \text{ and } M_{u',v',s} = 1 \text{ and } u <_A u' <_A v <_A v', \text{ then } S_{u'} \neq S_v$$

$$\text{if } M_{u,v,s} = 1 \text{ and } M_{u',v',s} = 1 \text{ and } u' <_A u <_A v <_A v', \text{ then } S_u = S_v$$

for all $(u, v), (u', v') \in Q$. The first one avoids that a departing unit v departs from a track at the same side as arriving unit u' enters in cases that v departs later than u' arrives. The second one avoids that the physical unit represented by u and v passes the whole track while the physical unit represented by u' and v' is parked at the same track.

The first constraint rewritten in linear form results in:

$$M_{u,v,s} + M_{u',v',s} \leq 3 - S_v - S_{u'} \text{ if } u <_A u' <_A v <_A v'$$

$$M_{u,v,s} + M_{u',v',s} \leq 1 + S_v + S_{u'} \text{ if } u <_A u' <_A v <_A v'$$

The second constraint is also replaced by two constraints:

$$M_{u,v,s} + M_{u',v',s} \leq 2 - S_u + S_v \text{ if } u' <_A u <_A v <_A v'$$

$$M_{u,v,s} + M_{u',v',s} \leq 2 + S_u - S_v \text{ if } u' <_A u <_A v <_A v'$$

Generalization to parts with more units

For two units in the same part, the ordering $<_A$ is not the same as $<_B$: Consider the same arriving composition on page 37 ordered as $u_k <_A \cdots <_A u_1$. The unit closest to the A-side is farthest from the B-side. So the $<_B$ ordering of this composition becomes $u_1 <_B \cdots <_B u_k$. So two units u_i and u_j with $u_i <_A u_j <_B u_i$ belong to the same part and u_j is closer to the A-side then u_i.

This gives that the last four constraints have to be rewritten:

$$M_{u,v,s} + M_{u',v',s} \leq 3 - S_v - S_{u'} \qquad \text{if } u <_B u' <_B v <_B v' \qquad (9.16)$$

$$M_{u,v,s} + M_{u',v',s} \leq 1 + S_v + S_{u'} \qquad \text{if } u <_A u' <_A v <_A v' \qquad (9.17)$$

$$M_{u,v,s} + M_{u',v',s} \leq 2 - S_u + S_v \qquad \text{if } v <_A v', u' <_B u \qquad (9.18)$$

$$M_{u,v,s} + M_{u',v',s} \leq 2 + S_u - S_v \qquad \text{if } u' <_A u, v <_B v' \qquad (9.19)$$

such that 'later arriving' at a track is generalized from 'arriving in a later train' to also 'arriving in the same train but more close to the rear when entering the track' which are both modeled in the introduced orderings.

Finally five constraints are needed for units of the same part that are parked at two different sides of the same track:

$$M_{u,v,s} + M_{u',v',s} \leq 3 + S_u - S_v - S_{u'} \qquad \text{if } u <_A u' <_B u, v <_B v' \tag{9.20}$$

$$M_{u,v,s} + M_{u',v',s} \leq 3 + S_{v'} - S_v - S_{u'} \qquad \text{if } u <_B u', v <_A v' <_B v \tag{9.21}$$

$$M_{u,v,s} + M_{u',v',s} \leq 2 - S_u + S_v + S_{u'} \qquad \text{if } u' <_A u <_B u', v <_A v' \tag{9.22}$$

$$M_{u,v,s} + M_{u',v',s} \leq 2 - S_{v'} + S_v + S_{u'} \qquad \text{if } u <_A u', v' <_A v <_B v' \tag{9.23}$$

$$M_{u,v,s} + M_{u',v',s} \leq 3 + S_u - S_v - S_{u'} + S_{u'} \qquad \text{if } u <_A u' <_B u, v <_A v' <_B v \tag{9.24}$$

Let us explain one type. A restriction (9.21) is only restrictive if $S_v = S_{u'} = 1$ and $S_{v'} = 0$ since otherwise the righthand side is at least two, and the restriction is trivially fulfilled because $M_{u,v,s}$ and $M_{u',v',s}$ are binary. We know that units v and v' will leave in the same train. In addition, we know that $u' <_B v$ because otherwise $v' <_B v <_B u'$ which contradicts $(u', v') \in Q$. This causes a crossing, and therefore the right-hand-side is one so that this pair of matchings on the same track is not possible given this constraint.

For all nine types of constraints sets are made with pairs of matchings $(u, v), (u', v') \in Q$ for which the statement after the 'if' holds. Note that these sets can be made in advance because we assumed a fixed order of units where out we can derive $<_A$ and $<_B$ for all pairs of units. If for every pair of matchings these constraints hold, then crossings are avoided in the shunt plan.

From now on OPG refers to the model with objective (9.2) and restrictions (9.3)-(9.12), excluding (9.6) and supplemented by restrictions (9.14)-(9.24).

10 Timefixer

After OPG has made an emplacing plan, the TIMEFIXER determines for every unit the route it takes and the shunt time at which it is driven. If a sawing route is used, also the start times of the non-sawing route parts are decided by the TIMEFIXER.

As already mentioned, the main issue of the TIMEFIXER is to avoid conflicts between train services and shunting compositions and between two shunting compositions. We do not treat the model that is used. An description of a basic version can be found in the PhD thesis of J. Van den Broek [vdBro09].

11 Reflection

In this section the way of making the shunt plan as described in last sections is analyzed.

Although the methods OPG and timefixer take little account of each other, this sequential approach works quite well, in practise. Mostly, the *timefixer* has much freedom because of the large amount of possible shunt times and routes. However there are situations in which this approach does not give feasible outcomes. In the first two subsections we treat such situations. Thereafter some current adaptions to the process are presented. In the last subsections we present examples showing that there are cases for which only an integral approach of routing with matching and parking gives an optimal solution.

11.1 Change of order

Firstly, note that the arrival and departure times of the units at their park tracks as assumed by OPG are not realized in practise. The real shunt times (and thereby the real leaving and entering times of the units at their park tracks) thus change by the proposed solution of the TIMEFIXER compared to the assumption of OPG. For an arriving unit the arrival time at the park track is assumed by OPG to be equal to the arrival time of the train service it belongs to. For a departing unit the departure time from the park track is assumed to be equal to the departure time of the train service it belongs to. In practise these units have some dwelling time on the platform track. This is not a problem as long as units leaving/entering the same park track leave/enter it in the same **order** as assumed by OPG.

However, this can not always be fully guaranteed: The TIMEFIXER may not be able to find shunt times for the units such that no orders are changed and no conflicts occur whereas, if we had another emplacing plan, this could perhaps be avoided.

11.2 Blocking conflicts

Secondly, there is another way how the emplacing plan limits the possibilities for the routing decision: A route to or from a park track that passes other park tracks is unavailable when units are parked at these tracks.

A route is *blocked* if at one or more via tracks of the route units are parked. If all routes on a relation are blocked at the same time, the relation can not be served. If at that time a composition is planned to serve this relation, no route can be found and a *blocking conflict* occurs.

OPG does not take into account routes, and certainly not the via tracks of routes. So it can happen that OPG decides that a composition c has to serve a relation (s_1, s_2) as well as that for each route $R^i_{s_1,s_2}$ on this relation a unit has to park at a via track of the route. This is only a problem, if at all shunt times for c all routes on (s_1, s_2) are blocked.

Example. *Consider the infrastructure at figure 9. Assume that a composition c has to serve relation (25,8) and that this is only possible between 12.03h and 12.13h. Also assume that a physical unit represented by arriving unit u and departing unit v is scheduled to park at 22. If u can only arrive at 22 before 12.03 and v can only leave after 12.13 a blocking conflict occurs. Either c can not be scheduled in this way or u and v can not be matched and parked at track 22.* □

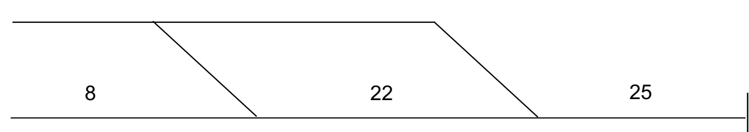

Figure 9: Blocking conflict

While making the emplacing plan we do not know the shunt times of the units exactly, so we can not enterily prevent blocking conflicts.

11.3 Infrastructural input in OPG

In this subsection we treat some current additions to the version of OPG as described in section 9. These additions are intended to prevent the situation that the TIMEFIXER can not find for every unit a feasible shunt time in situations as described in last subsections. This is done by inserting infrastructural input in OPG. By doing this the first stage OPG is able to take into account routing effort.

11.3.1 Relation costs

To trigger OPG not to use relations which are difficult to plan, costs are assigned to each relation. The more difficult to plan a relation, the higher the costs of this relation. The calculation of these costs is done by a procedure running before the OPG-run. It determines for each relation the shortest path in the infrastructure, if it exists. The costs of a relation are defined as the costs of this path. If no path exists, it returns ∞.

The *relation costs* of an emplacing plan are the total costs of all shunt movements implied by the emplacing plan. Minimizing these costs results in a better shunt plan. Below we give an expression for the relation costs C in the MIP of OPG.

Let Z be the set of all relations. Let relation $(s_1, s_2) \in Z$. The cost of (s_1, s_2) is denoted by C_{s_1,s_2} and equals the cost of relation (s_2, s_1). If $s_1 = s_2$ the cost C_{s_1,s_2} equals zero. Relations with cost zero follow if OPG chooses to park a unit u on pl_u.

Let $u \in U$ and $s \in S_p$. Remind that $P_{u,s}$ is the decision variable that is one if unit u is assigned to track s. Now we can express the relations costs C:

$$C = \sum_{u \in U^+} \sum_{(s_1,s_2) \in Z, s_1 = pl_u} C_{s_1,s_2} P_{u,s_2} + \sum_{u \in U^-} \sum_{(s_1,s_2) \in Z, s_2 = pl_u} C_{s_1,s_2} P_{u,s_1}$$

These costs are added with a certain weight w_3 to the objective of OPG (9.1.1). They make it more attractive to park a unit u at its platform track, because then the relation for u has cost zero.

11.3.2 Determining conflicts in advance

The second addition is applied in cases with fixed shunt times. Then the TIMEFIXER has no time freedom which increases the risk of infeasibility. In these cases more information is available when running OPG, namely the real fixed shunt times. For these cases extra information about which relations will give conflicts is given to OPG. This input is in the form of constraints.

This procedure also produces constraints to avoid blocking conflicts if the shunt times are fixed in advance. These constraints can also be applied if only the **order** in which units are shunted is fixed.

11.3.3 Shifting

The last addition that we treat tries to avoid situations as described in subsection 11.1. As mentioned earlier the orderings $<_A$ and $<_B$ are derived from the times τ_{t_u} of units $u \in U$. Sometimes this order can not be realized. This is especially the case if the earliest possible shunt time for an arriving unit is much later than the arriving time of its train service or if the latest possible shunt time for a departing unit is much earlier than the departure time of its train service. We give an example:

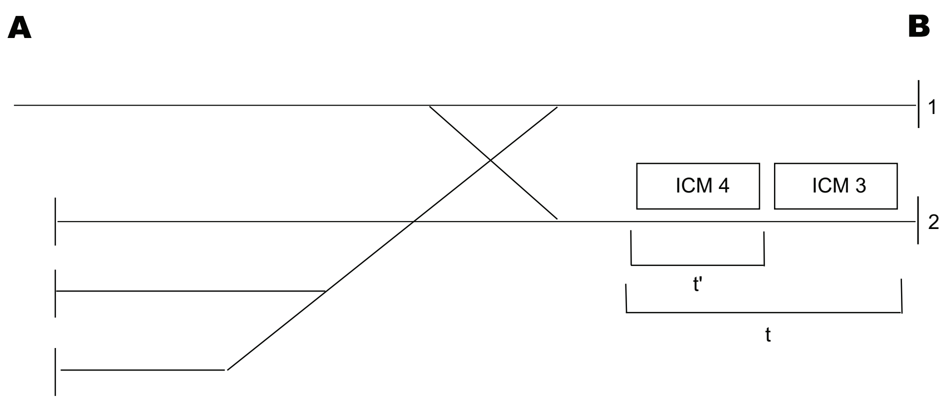

Figure 10: Shifting example

Example. *Consider arriving train service t with composition* ICM 4 $<_B$ ICM 3 *in figure 10. Assume that it enters a dead end track at* $\tau_t = 9.01$. *Assume that the rolling stock circulation prescribes that* ICM 4 *has to depart in train service t' with* $\tau_{t'} = 9.29$ *from the same track, in other words: t has a transition to t'. Unit* ICM 3 *enters the supply, but it can only be shunted after the departure of t'. In* OPG *for these cases the time of unit* ICM 3 *where out the $<_A$ and $<_B$ ordering is changed from τ_t into $\tau_{t'}$.*

This is useful if between 9.01 and 9.29, say 9.16, other units enter the station. If the units of this train service have to be shunted before 9.29, then they enter their park track before unit ICM 3. *If both are parked on the same* *track this gives another order on this track. Using $\tau_{t'}$ in* OPG *as shunt time for* ICM 3 *in stead of τ_t makes that* OPG *can take this order into account.* □

11.4 Time freedom

In last subsection we saw that some adaptions to OPG can make the co-operation with TIMEFIXER better. However, in some cases it is not known in advance whether an 'intervention' in the input of OPG is useful.

Example continued. *The example of last subsection is continued. Assume that the train service entering at track* 1 *has composition* (ICM 4) *and that this unit has a shunt window with upperbound at 9.43. The set U^+ then is:*

Unit u	type	τ_{t_u}	pl_u	$sw(u)$
u_1	ICM 3	9.01	2	[9.32, 9.43]
u_2	ICM 4	9.16	1	[9.19, 9.43]

Assume that the rolling stock circulation prescribes that these units have to depart in the same train service. The set U^- is then:

Unit v	type	τ_{t_v}	pl_v	$sw(u)$
v_1	ICM 3	15.29	1	[15.02, 15.26]
v_2	ICM 4	15.29	1	[15.02, 15.26]

Because $t_{v_1} = t_{v_2}$ we want to route v_1 and v_2 together. This is only possible if they come from the same park track. Note that by the types the only possible matching is $d(u_1) = v_1$ and $d(u_2) = v_2$, thus that this input implies that we want to put u_1 and u_2 at the same park track.

If the input contains the requirement $v_1 <_B v_2$, then we want that for the arriving units holds: $u_2 <_B u_1$, because giving shunt times to OPG *so that u_2 is shunted later than u_1, would give an emplacing plan with u_1 and u_2 at the same track so that v_1 and v_2 can be routed together.*

If the input prescribes that $v_2 <_B v_1$, then u_1 should enter the park track ***before*** *u_2 and we would insert a later shunt time for u_2 than for u_1 in* OPG.
□

In this example we had a very small set of units, so that we could immediately determine which units are matched and which shunt times to give to these units to save shunt movements. Therefore it became clear from these sets how to manipulate the input of OPG. However, if these sets would consist of much more units, then it is not sure that u_1 has to be matched to

v_1 and u_2 to v_2 what makes that it is not clear how to manipulate the input of OPG.

Below we present another case where we can not see in advance whether manipulating the input is useful or not.

Shunting together

In this subsection we show that the possibility of routing units from different train services together can lead to another order of these units on the park tracks than assumed by OPG.

Remind that the number of part splittings is included in the objective of OPG. This is done because if a part is split, both compositions have to be routed apart. That needs extra infrastructural and driver capacity. Another way to save routings, is to shunt two arriving units u and u' from different train services $t_u \neq t_{u'}$ together to the same park track. If both arrive at the same platform track, this can be done by holding u on pl_u until $t_{u'}$ arrives. After that u and u' can be coupled and shunted together to their park track.

However, this can lead to a different order of u and u' on $pt(u)$ than assumed by OPG, thus TIMEFIXER can not always use this opportunity. This is illustrated by an example:

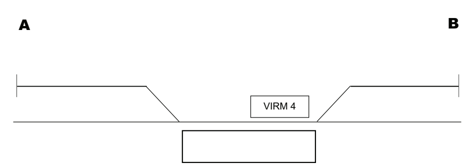

Figure 11: `virm 4` on the platform track

Example. *Consider unit* VIRM 4 *entering the platform track, figure 11. One possibility is to shunt it to the left most park track. Next, unit* VIRM 6 *enters the station via the A-side, like the situation as in figure 12 left. If this unit is shunted to the same park track, we end up with the order* VIRM 4 $<_B$ VIRM 6 *(unit* VIRM 6 *more close to the B-side than unit* VIRM 4*), as can be seen at the right side of figure 12. This is the order as assumed in* OPG.

If we choose to hold VIRM 4 *at the platform track, the situation after* VIRM 6 *has entered the station is as depicted at the left side of figure 13. Now, both units can drive together to the left most park track and we end up with the situation as the right side of figure 13. The issue now is that the*

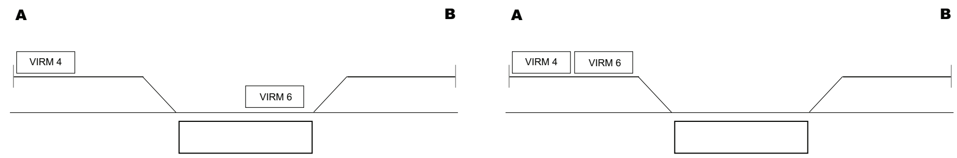

Figure 12: Option 1

units are ordered as VIRM $6 <_B$ VIRM 4*, which is the opposite way as in the first situation.* □

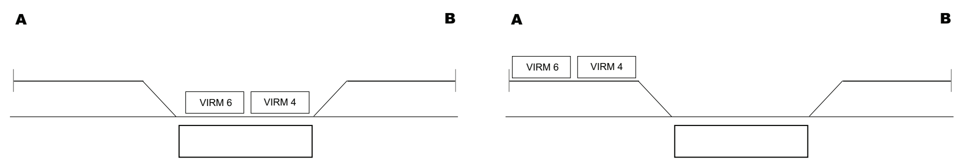

Figure 13: Option 2

Changing the input of OPG is only useful if this new order results in a better shunt plan. However, this is in general not known in advance. OPG has been built under the assumption that such a joint movement is never done. Therefore, within this framework we can not collect the profits from shunting together.

12 Summary part III

In this part we described the current two-stage approach of the shunt plan. We saw that in the second stage infeasibility can occur caused by decisions in the first stage. Also some profit in the second stage could not be collected by choices in the first stage.

Two stages

In the first stage, OPG, only matching and parking is decided. This is done under the assumption of fixed shunt times. The fixed shunt times gave rise to the partial orderings $<_A$ and $<_B$ on all units in U assuming that units from different trains are shunted separately. Given these orderings crossings can be avoided smoothly because the sets of possible crossings can be determined in advance for all tracks.

The real shunt times are decided in the second stage by the TIMEFIXER. In this second stage some things can go wrong:

- The real decided shunt times can be such that two units on the same park track arrive in different order than assumed by OPG.
- The route to/from the park track of a certain unit may be blocked by other parking units during all shunt times in the shunt window.

Some extra functionality is added to the basic version of OPG. These methods add infrastructural input to OPG so that it makes a better emplacing plan.

Besides it can happen that two units can be routed in two different orders to the same park track and that one of the orders is better. In some cases it is possible to save shunt movements by shunting units from different trains that arrive/depart at/from the same platform track together. This can also lead to another order of units on the park tracks. However, we have to insert fixed shunt times in OPG. In small cases it may be possible to determine how to manipulate the input of the first stage to get a feasible/better solution in the second stage, but in bigger cases this is not always possible. In those cases an integral approach is better.

The conclusion of this section is that the problems that arrise by the sequential approach with OPG and TIMEFIXER can only be fully banned by an integral approach of routing with matching and parking and that an integral approach is able to meet the optimal number of shunt movements in more cases.

Part IV

Problem approach

In section 3 we presented a problem description. We want to model this problem. In part III we argued that an integral approach of matching, parking and routing leads to optimality in more cases. In this part a MIP formulation is presented for making these three decisions integral. It is called 'the APT model'.

In section 13 we introduce the decision variables that fix the routing and parking decision for each unit. In section 14 the decision variables for the matching decision are described. Also the connection between these two types of decision variables is treated in that section. Section 15 shows how the APT model is able to realize a better emplacing plan by using the freedom in time.

For the proposed solution method it is referred to part V.

13 APT-model

In this section the decision variables for the routing and parking decision are treated. Besides the constraints between these variables to avoid conflicts, blocking conflicts, overusage of park tracks and crossings on platform tracks are described.

The goal is to assign every unit $u \in U^+$ to a park track, denoted by $pt(u)$ and a route $R(u)$. Moreover, for every route part i of $R(u)$ a departure time $T^i(u)$ has to be assigned. Note that $T^1(u)$ equals the shunt time $\tau(u)$.

Given a unit $u \in U^+$, the set of all combinations of possible choices is denoted by APT_u. Apt is the abbreviation for 'arrival on park track'. An element $apt \in APT_u$ contains a feasible, (unique and) full description of the way unit u is handled between arrival of t_u and arrival at its park track $pt(u)$. In subsection 13.3 we give an idea how sets APT_u can be generated.

An apt *apt* has among others the following fields:

- c_{apt}, the composition of units for which *apt* is made. It represents that the units $u \in c_{apt}$ are routed together.
- R_{apt}, the route. The origin of this route $o_{R_{apt}}$ equals pl_u for each $u \in c_{apt}$.
- pt_{apt}, the park track which is equal to $d_{R_{apt}}$.

- T_{apt}, a vector of departure times of the route parts of R_{apt}.
- θ_{apt}, the arrival time on pt_{apt}, that follows from the departure time of the last route part.

Every unit $u \in U^+$ can appear in several apt's: In apt's where u shunts alone, but also in the apt's where u shunts together with other units u' of the same part. We give an example how the set $APT = \bigcup_{u \in U^+} APT_u$ may look like in a small case.

13.1 Example

Consider units (VIRM 4, VIRM 6) arriving at platform track p, see figure 14. Assume that the arrival time is 12.00, that VIRM 4 is more close to the B-side than VIRM 6 and that both units enter the supply. The units can only leave p via the B-side and at 12.10 the next train arrives via the A-side and it forces by its length that both VIRM 4 and VIRM 6 have to leave before that arrival.

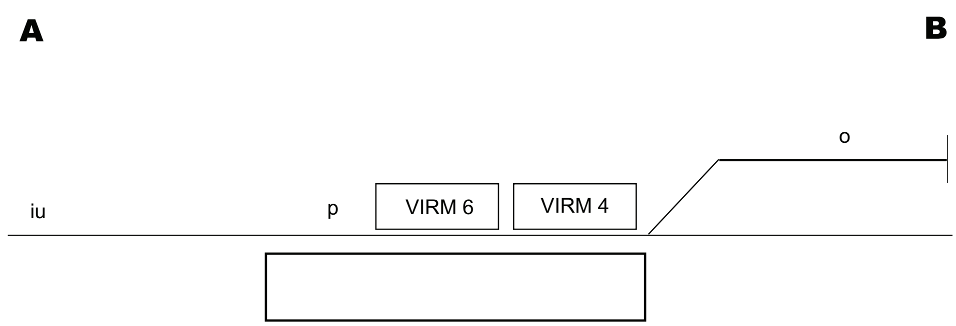

Figure 14: Apt example

There will be apt's for shunting VIRM 4 and VIRM 6 together and for shunting them apart. The shunt window for the composition (VIRM 4, VIRM 6) equals $[12.00+tir(Uitstaptijd), 12.10-tir(AnaVOpZelfdeSpoorRR)] =$ $= [12.03, 12.07]$. If there is only one park track o and one route $R_{p,o}$ leading to it, the following apt's are included in APT:

c_{apt}	τ_{apt}	R_{apt}	number
(VIRM 4, VIRM 6)	12.03	$R_{p,o}$	1
(VIRM 4, VIRM 6)	12.04	$R_{p,o}$	2
(VIRM 4, VIRM 6)	12.05	$R_{p,o}$	3
(VIRM 4, VIRM 6)	12.06	$R_{p,o}$	4
(VIRM 4, VIRM 6)	12.07	$R_{p,o}$	5

(The last column is to identify an apt in later text).

Besides there are apt's for VIRM 4 and VIRM 6 shunting apart from each other. The shunt window for composition (VIRM 4) and composition (VIRM 6) equals

$$[12.00+min(tir(Uitstaptijd), tir(Splitsen)), 12.10-tir(AnaVOpZelfdeSpoorRR)]$$

$$= [12.03, 12.07]$$

So we include the following apt's:

c_{apt}	τ_{apt}	R_{apt}	number
VIRM 4	12.03	$R_{p,o}$	6
VIRM 4	12.04	$R_{p,o}$	7
VIRM 4	12.05	$R_{p,o}$	8
VIRM 4	12.06	$R_{p,o}$	9
VIRM 4	12.07	$R_{p,o}$	10
VIRM 6	12.03	$R_{p,o}$	11
VIRM 6	12.04	$R_{p,o}$	12
VIRM 6	12.05	$R_{p,o}$	13
VIRM 6	12.06	$R_{p,o}$	14
VIRM 6	12.07	$R_{p,o}$	15

13.2 Characteristics of an apt

If a certain $apt \in APT$ is chosen to be part of the shunt plan, the following decisions result:

- The units of composition c_{apt} are shunted together.
- For all units $u \in c_{apt}$ the route $R(u)$ equals R_{apt}.
- The departure times of the routeparts $T(u)$ equal T_{apt}.

Based on these decisions, also the following information can be derived:

- From R_{apt} we know along which side the platform track pl_u ($u \in c_{apt}$) is left. We call this $plsi_{apt}$. It attains the value A or B.
- From the timetables we know which trains enter and leave pl_u during the dwell time of c_{apt}. This dwell time ends at T^1_{apt}. That determines the side of the platform track where c_{apt} stands between τ_{t_u} and τ_{apt}. We denote this by ds_{apt}, the dwell side, which attains A or B.
- From R_{apt} we also know the via tracks that are passed during driving the route and at which time they are occupied by c_{apt}. We deduce from *apt* a list of pairs $((track^1, time^1), (track^2, time^2), \dots)_{apt}$ meaning that $track^i$ is passed at $time^i$ if *apt* is chosen.
- From R_{apt} we know what the park track $pt(u)$ of the units $u \in c_{apt}$ is. We call this pt_{apt}. Also the side of pt_{apt} along which this track is entered, called si_{apt}, is deduced from R_{apt}.
- From the departure times of the several route parts, the time θ_{apt} at which composition c_{apt} arrives on its park track is deduced.

Note that the information about c_{apt} also includes the internal ordering of the units. Finally some costs C_{apt} may be assigned to the choice of *apt*. The way costs may be specified is discussed in later parts.

13.3 The set APT

For every $apt \in APT$ we introduce the decision variable x_{apt} which attains the value one if *apt* is chosen and zero otherwise.

$$x_{apt} = \begin{cases} 1 & \text{if } apt \text{ is chosen} \\ 0 & \text{otherwise} \end{cases}$$

The set APT, containing all apt's for all units in a specific planning period that have no conflicts with already scheduled trains is calculated in a preprocessing step. Thereby the infrastructural properties and the timetables have to be taken into account. In the following subsection we give an idea how this set can be generated.

Generating the set APT

Given the timetables with all trains passing, starting or ending at the station and the infrastructure, for every unit $u \in U^+$ the set of possible apt's can be calculated. In this section an idea is given how this can be done.

We consider a unit u arriving in train service t_u at time τ_{t_u} on platform track pl_u. We make a list of all possible compositions in which u may take part. For every such composition c we do the following:

- The shunt windows, one for each leaving side, are calculated.
- For every minute in the shunt window and all routes leading to a park track an apt is made if the time and route do not cause a conflict with already scheduled trains.
- If there are possibilities for sawing, the apt's are extended per route part. At the saw tracks for every feasible departure time and route new apt's are made.

The calculation of the shunt window is done as in section 7. Below we describe the other two steps:

- For both leaving sides $plsi$, for every minute m in the shunt window at side $plsi$ we make a list $\mathcal{L}$ of all routes $R_m \in \mathcal{R}$ (the set of non sawing routes) that
  - have origin at the platform track ($o_{R_m} = pl_u$),
  - are in the direction $plsi$,
  - have destinations that are a park tracks or tracks where it is allowed to saw,
  - have destinations that are long enough to handle the composition ($l_c \leq l_{d_{R_m}}$) and
  - have no conflict with already scheduled trains at minute m.
- For all $R_m \in \mathcal{L}$ we check whether it can be extended by adding a route part in the opposite direction. This procedure can be repeated a maximum allowed number of saw movements.

  Let st be the minimal saw time for composition c. It is calculated from its total length l_c. Let $tir(Nietzagenderoute)$ be the driving time of a non sawing route. A route R_m can be extended if track d_{R_m} is not used by trains from the timetables between the arrival time on this track ($m + tir(Nietzagenderoute)$) and the earliest leaving time ($m + tir(Nietzagenderoute) + st$).

  Given a route R_m for all minutes $m' \geq m + tir(Nietzagenderoute) + st$ we list all non sawing routes $R_{m'} \in \mathcal{R}$ that

– have origin at the saw track $o_{R_{m'}} = d_{R_m}$,
– are in the direction opposite from R_m,
– have destinations that are park tracks or tracks where it is allowed to saw,
– have no destination or via tracks that are equal to a via track or origin of R_m,
– have destinations that are long enough to handle the composition ($l_c \leq l_{d_{R_{m'}}}$) and
– have no conflict with already scheduled trains at minute m'.

This is done for every route R_m from $\mathcal{L}$. After the maximum number of saw movements has been reached, for all routes that have destination at a park track, we make an apt and add it to the set APT_u.

Note that within this procedure also apt's for compositions with units from different train services can be made. If within a shunt window of a unit a unit from another train service of the same family enters the platform track, apt's for the composition with both units can be made. More about this can be found in section 15.1.

Concluding, we need:

- a tool to calculate the shunt windows for a given composition
- a tool to determine conflicts
- a maximal number of saw movements

This results in a set APT_u for every unit u that contains all possible apt's that do not conflict with already scheduled trains.

Departing units

Until now we only focussed on arriving units. Of course also for departing units $v \in U^-$ a routing and parking decision has to be made. For them we define dpt's, 'departures from park tracks'. A $dpt \in DPT_v$ contains a full description of handling a departing unit v between departure from its park track until departure in t_v. A dpt is somehow the symetric to apt and can be handled similar.

When choosing dpt, we decide that the units of composition c_{dpt}

- are shunted together and retrieved from track pt_{dpt},
- depart from that track at θ_{dpt} via side si_{dpt},
- take route R_{dpt},
- have departure times equal to T_{dpt},
- enter the platform track at τ_{dpt} via side $plsi_{dpt}$ and
- dwells at pl_u at side ds_{dpt}. This side is determined from p_{t_u} $(u \in c_{dpt})$, the predecessor of t_u, or earlier arriving and departing trains from pl_u.

Note that θ for both apt's as dpt's indicates the planned time at the **park** track and that τ always indicates the time at the **platform** track.

13.4 Constraints per unit

In the following subsections we present situations between apt's and dpt's that need to be avoided and present constraints which ensure this. We treat the following situations:

- overusage of park tracks,
- conflicts implied by the departure times and routes of the apt's and dpt's,
- crossing on platform tracks and
- blocking conflicts

But next to avoiding conflicts, we have to avoid that two different decisions can be made for one unit. This is ensured by the constraints

$$\sum_{apt \in APT_u} x_{apt} = 1 \quad \forall u \in U^+ \tag{13.1}$$

Example. *In the example of subsection 13.1 this results in:*

$$x_i \in \{0,1\} \quad (i = 1, \ldots, 15)$$
$$\sum_{apt \in APT_{\text{VIRM 4}}} x_{apt} = 1$$

$$\sum_{apt \in APT_{\text{VIRM 6}}} x_{apt} = 1$$

Thus:

$$x_1 + \cdots + x_{10} = 1$$

$$x_1 + \cdots + x_5 + x_{11} + \cdots + x_{15} = 1$$

In the same way we have to guarantee that:

$$\sum_{dpt \in DPT_v} x_{dpt} = 1 \quad \forall v \in V^- \tag{13.2}$$

This construction makes that also the decision to bring u directly to $pl_{d(u)}$ counts as an apt. In this case, no dpt for $d(u)$ is needed, because it is already on its platform track. However, the requirement that every departing unit gets a dpt has been inserted, so we have to make a dummy dpt that has to be chosen in this case. More about this can be found in subsection 15.2.

13.5 Constraints avoiding overusage of park tracks

In this subsection we deduce the constraints that avoid an overusage of park tracks. We first define the variables that administrate the length on a track. Using these variables the constraint can easily be formulated.

In the beginning of the planning period the tracks $s \in S_p$ are assumed to be empty. During the planning period several compositions enter and leave the park tracks. To avoid overusage of the lenght of park tracks we introduce the variable $L_{s,\theta}$ as the total length of the units parked at track s at time θ. The total length at time θ is the sum of all compositions that have entered before θ substracting the sum of all compositions that have left before θ.

For a track $s \in S_p/S_{pl}$ and θ in the planning horizon, $L_{s,\theta}$ can be linked to the apt's and dpt's in the following way:

$$L_{s,\theta} = \sum_{apt \in APT |\, \theta_{apt} \leq \theta,\, pt_{apt} = s} l_{apt} x_{apt} - \sum_{dpt \in DPT |\, \theta_{dpt} \leq \theta,\, pt_{dpt} = s} l_{dpt} x_{dpt} \tag{13.3}$$

For platform tracks, the situation is completely different. The lenght of the units on it has to be deduced from the chosen **shunt** times of all units. Besides also the trains from the timetables may make use of these tracks. The total lenght on $s \in S_{pl}$ at time θ equals the sum of

- the lenght of all train services t entering track s with $\tau_t < \theta$,
- substracting the lenght of the train services t departing from track s with $\tau_t < \theta$,
- adding the lenght of units $u \in U^-$ with $\tau(u) < \theta$ and
- substracting the arriving units $u \in U^+$ with $\tau(u) < \theta$.

So for a platform track $s \in S_{pl}$ the total used lenght at time θ is linked to the apt's and dpt's like:

$$L_{s,\theta} = \sum_{t \in \mathcal{T} |\, pl_t = s,\, \tau_t \leq \theta,\, t \text{ arriving}} l_t - \sum_{t \in \mathcal{T} |\, pl_t = s,\, \tau_t \leq \theta,\, t \text{ departing}} l_t + \tag{13.4}$$

$$+ \sum_{dpt \in DPT_u |\, pl_{dpt} = s,\, \tau_{dpt} < \theta} l_{dpt} x_{dpt} - \sum_{apt \in APT_u |\, pl_{apt} = s,\, \tau_{apt} < \theta} l_{apt} x_{apt}$$

Let l^s be the phycical lenght of track s. The constraints then become, for all time θ in the planning horizon:

$$L_{s,\theta} \leq l^s \quad \forall s \in S_p \tag{13.5}$$

13.6 Constraints avoiding conflicts

Within the construction of apt's and dpt's, we already ensured that they have no conflicts with already scheduled trains. However, two apt's or dpt's can have a conflict with each other.

In an apt both the departure/arrival times and routes of every route part are known. So given two elements of $APT \cup DPT$, i and j, we know whether they have a conflict. If at least one route part of i has a conflict with a route part of j we make the constraint:

$$x_i + x_j \leq 1 \tag{13.6}$$

The way how such conflicts can be detected, follows from the explanation in section 6. We conclude with an example:

Example. *Consider a unit u_1 arriving on platform track* `1` *and a unit u_2 arriving on track* `2`*, see figure 15. Consider the following apt's:*

$APT_{u_1} =$

τ_{u_1}	R^1	T^2	R^2	
12.01	$R_{1,4}$	12.07	$R_{4,3}$	1
12.01	$R_{1,4}$	12.08	$R_{4,3}$	2
12.01	$R_{1,4}$	12.09	$R_{4,3}$	3
12.02	$R_{1,4}$	12.08	$R_{4,3}$	4
12.02	$R_{1,4}$	12.09	$R_{4,3}$	5

$APT_{u_2} =$

τ_{u_2}	R^1	
12.05	$R_{2,5}$	6
12.06	$R_{2,5}$	7
12.07	$R_{2,5}$	8
12.08	$R_{2,5}$	9
12.09	$R_{2,5}$	10

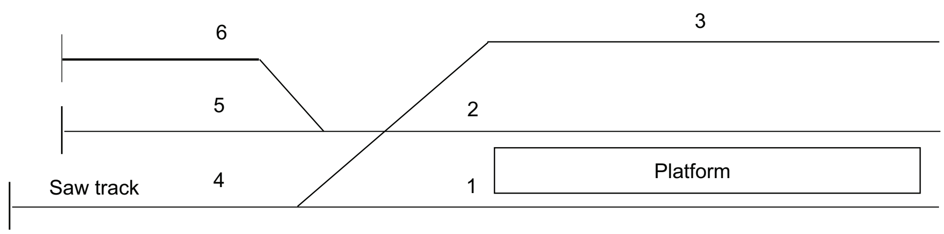

Figure 15: Example avoiding conflicts

Some constraints needed in this example are:

$$x_{apt_1} + x_{apt_7} \leq 1$$

$$x_{apt_1} + x_{apt8} \leq 1$$

$$x_{apt_1} + x_{apt9} \leq 1$$

□

13.7 Crossings on the platform track

If two units from the same part choose different apt's, the time and side that they leave the platform track can cause a crossing. A crossing takes place if a composition has to leave from its track at a moment that other units are standing more close to the leaving side of the track than the composition that has to leave. In this subsection we present the constraints that avoid this.

Just like in the preceding subsection all elements to know whether two apt's $i \in APT_u$ and $j \in APT_{u'}$ with $t_u = t_{u'}$ cause a crossing on their platform track are known: the departure times (τ_i and τ_j) and the leaving sides ($plsi_i$ and $plsi_j$). The order in which the units of i and j stand on the platfrom track is deduced from the composition of the train service in which they enter the station.

If choosing both i and j would cause a crossing, we add the following constraint:

$$x_i + x_j \leq 1 \tag{13.7}$$

For departing parts, the same sort of constraints can be derived.

Note that these constraints are only made for two units from the same part. Crossings on the platform track between units out of different parts are part of next subsection (case 2).

13.8 Blocking conflicts

A *blocking conflict* takes place if a shunting composition uses a composed route that is not available because units are parked at a via track when the composition is planned to pass that track. In this subsection we describe how these situations can be avoided.

We identify three types of blocking conflicts and we explain them below:

1. Two shunting compositions make use of the same saw track
2. A shunting composition leaves/enters its platform track via a side that is occupied by units of another part
3. A shunting composition passes via tracks where units are parked

We now give the constraints needed in these three cases.

1. If R_{apt} is sawing, then it dwells some time at its saw track s. The time that c_{apt} is present at s are the minutes between the departure time of the route part with destination s increased by $tir(Nietzagenderoute)$ and the departure time of the route part with origin s. If this interval has an overlap with the interval that an apt' uses track s as saw track, we add the constraint

$$x_{apt} + x_{apt'} \leq 1 \tag{13.8}$$

2. The constraints of subsection 13.7 do not avoid crossings between units from different train services that are present on the same platform track. Given two apt's apt and apt' with $pl_{apt} = pl_{apt'}$ and $ds_{apt} = plsi_{apt'}$. If $\tau_{apt'}$, its leaving time from the platform track, is between τ_{t_u} ($u \in c_{apt}$) and τ_{apt}, then a crossing occurs and we add the constraint.

$$x_{apt} + x_{apt'} \leq 1 \tag{13.9}$$

   Similar constraints are needed for dpt's.

3. During driving a composed route, the via tracks can be used by parking units. In this case the apt or dpt implying this composed route may not be chosen. Given an apt/dpt, we know the tracks s^i the composition passes and the times θ^i at which it passes those tracks. To determine whether a certain track s^i is empty at time θ^i, we deduce the binary variable B_{s^i,θ^i} that is one if something is parked at s^i at time θ^i and zero if s^i is empty at that time. The B-variable can be derived from the lenght variable L_{s^i,θ^i} by $B_{s,\theta} \geq \frac{L_{s,\theta}}{l^s}$ which results in:

$$B_{s,\theta} l^s \geq L_{s,\theta} \tag{13.10}$$

   The constraints to ensure that track s^i is free at the time at which apt/dpt uses it, θ^i, can be formulated by

$$B_{s^i,\theta^i} + x_{apt/dpt} \leq 1 \tag{13.11}$$

14 Matchings in the APT-model

Once chosen apt's we know for each arriving unit the process between arrival in its train service and arrival on its park track. Equivalently, if we have chosen dpt's for each departing unit the process between departure from its park track and departure time of its train service is known. Until now we did not insert anything in the MIP to guarantee that the set of chosen apt's fits with the set of chosen dpt's.

In this section we describe the structure to align the choices for apt- and dpt-variables. This is done by defining matching variables, linking them to the apt- and dpt-variables and formulate constraints to them.

14.1 Conditions to matchings

There are several conditions that have to hold if a matching between an arriving unit and a departing unit can be made. Units $u \in U^+$ and $v \in U^-$ with $\psi_u = \psi_v$ can be matched ($d(u) = v$) iff

- the park track of u is the same as the park track of v, ($\rightarrow pt(u) = pt(v)$).
- the arrival time of u on $pt(u)$ is earlier than the departing time of v, ($\theta_u < \theta_v$).
- at time θ_v the physical unit represented by u and v stands closest to the side at $pt(v)$ along which v is planned to leave. Otherwise a crossing occurs.

For each condition we dedicate a subsection to formulate corresponding constraints. But before that, we introduce matching variables.

14.2 Matching variables

We introduce the decision variable $M_{u,v}$ that takes the value one if u is matched to v and zero otherwise. This decision variable is made for all pairs of arriving and departing units (u, v) in Q. This set, as defined at page 26, contains the pairs (u, v) that are of the same type ($\psi_u = \psi_v$) and have $\tau_{t_u} + T \leq \tau_{t_v}$.

We have to ensure that every unit u is matched to exactly one departing unit v. This is done by the following two types of constraints:

$$\sum_{v \in U^- : (u,v) \in Q} M_{u,v} = 1 \quad \forall u \in U^+ \tag{14.1}$$

$$\sum_{u \in U^+ : (u,v) \in Q} M_{u,v} = 1 \quad \forall v \in U^- \tag{14.2}$$

14.3 On the same track

We have to ensure that u and v are assigned to the same track if they are matched. Thus $M_{u,v} = 1$ is only allowed if $pt(u) = pt(v)$. In other words: if u is parked at s and u and v are matched, then v may not be parked at another track:

For all matchings $(u,v) \in Q$, for all $s \in S_p$:

$$\sum_{apt \in APT_u, pt_{apt}=s} x_{apt} + \sum_{dpt \in DPT_v, pt_{dpt} \neq s} x_{dpt} \leq 2 - M_{u,v} \tag{14.3}$$

Note that the first summation only contains apt's of APT_u. This summation can not be larger than one by constraint (13.1) applied to u. The same holds for the second summation and constraint (13.2) applied to v. Therefore, constraints (14.3) are only restrictive if $M_{u,v}$ is one.

14.4 Time difference

To ensure that an arriving unit u is not matched to a departing unit v that has to leave $pt(v) = pt(u)$ before u arrives at that track, we link the matching variables to the apt- and dpt-variables via some extra variables. These variables are introducted and linked to apt- and dpt-variables. At the end of this subsection we present the main constraint to the matching variables.

We firstly derive variables $\theta(u)$ for all $u \in U^+$ that represent the arrival time of u on $pt(u)$.

$$\theta(u) = \sum_{apt \in APT_u} \theta_{apt} x_{apt} \quad \forall u \in U^+ \tag{14.4}$$

For departing units $v \in U^-$ the variable θ_v represents the departure time from its park track.

$$\theta(v) = \sum_{dpt \in DPT_v} \theta_{dpt} x_{dpt} \quad \forall v \in U^- \tag{14.5}$$

Secondly, from these θ-variables we derive for each $(u,v), u \in U^+, v \in U^-$ the 'time difference variable' $T_{u,v}$ that is zero if $\theta(v) < \theta(u)$ and one if $\theta(v) > \theta(u)$. If $\theta(u) = \theta(v)$, then $T_{u,v}$ may attain both values.

$$T_{u,v} = \begin{cases} 0 & \text{if } \theta(v) < \theta(u) \\ 1 & \text{if } \theta(v) > \theta(u) \end{cases}$$

Let l_{ph} be the length of the planning horizon in minutes. The constraints to link the time difference variables $T_{u,v}$ to the θ-variables are:

$$l_{ph}(T_{u,v} - 1) \leq \theta(v) - \theta(u) \quad \forall u \in U^+,\ v \in U^- \tag{14.6}$$

$$l_{ph} T_{u,v} \geq \theta(v) - \theta(u) \quad \forall u \in U^+,\ v \in U^- \tag{14.7}$$

Let us take a closer look to these constraints.

- If $\theta(v) < \theta(u)$ for both constraints the right hand side is strictly negative. If $T_{u,v} = 1$, the expression at the left hand side of (14.6) equals zero, thus this constraint forces $T_{u,v}$ to be zero in this case. Constraint (14.7) is trivially satisfied.
- If $\theta(v) > \theta(u)$, the right hand sides are strictly positive and constraint (14.7) forces $T_{u,v}$ to be one. Constraint (14.6) is trivially satisfied in this case.
- If $\theta(v) = \theta(u)$, the right hand side equals zero and both constraints are not restrictive.

The main constraints to prohibit matchings between u and v that can not be made because of a negative time difference between $\theta(u)$ and $\theta(v)$ are:

$$M_{u,v} \leq T_{u,v} \quad \forall (u,v) \in Q \tag{14.8}$$

The usage of the time difference variables in constraints of type (14.8) only requires the constraints of type (14.6) that bound these variables from above. Later on, in 14.5.5, we will use $T_{u,v}$ variables in constraints that require that they are bounded from below too.

14.5 Crossings

To avoid crossings between units on a park track we have to know the order they are on that track. That gives restrictions to the way they can be retrieved from the track.

To know the order of two arriving units u and u' we have to encounter the following aspects:

- The arrival times $\theta(u)$ and $\theta(u')$.
- If their park track is open at both sides: the arrival sides $S(u)$ and $S(u')$.
- If u and u' arrive in the same composition: the order of u and u' in which they are in the composition.

This subsection is built in the following way: Firstly ordering variables $a(u, u')$ are defined that indicate whether u or u' is more close to the A-side of the park track if they would enter the same track. These ordering variables are completely determined from the apt-variables by formulating several constraints. Using these ordering variables we make constraints avoiding crossings by excluding combinations of dpt-variables.

14.5.1 Ordering variables

We now define the variables that indicate the order of two units on their park track. Given units $u, u' \in U^+$ we define $a(u, u')$ to be zero if u is more close to the A-side on its park track than u' and one if u' is more close to the A-side on its park track than u **if they would enter the same track**. For the units in figure 16 the ordering variable $a(u, u')$ equals zero.

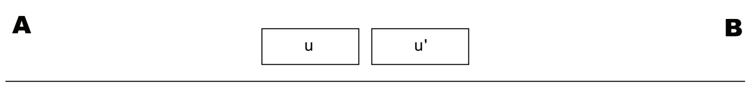

Figure 16: Ordering variable $a(u, u') = 0$

$$a_{u,u'} = \begin{cases} 0 & \text{if } u \text{ is more close to the } A\text{-side then } u' \\ & \text{if they would enter the same track.} \\ 1 & \text{if } u' \text{ is more close to the } A\text{-side then } u \\ & \text{if they would enter the same track.} \end{cases}$$

If u and u' do not enter the same park track, the variable $a(u, u')$ may attain both values. However, by the constraints that will be presented below, some of these variables will get one specific value, but they do not influence the MIP.

For consistency we insert the following constraints:

$$a(u, u') + a(u', u) = 1 \quad \forall u, u' \in U^+ \tag{14.9}$$

Below we present how these variables can be linked to respectively the arrival times, arrival sides and compositions.

14.5.2 Ordering variables and arrival sides

As mentioned in the beginning of this subsection, the order in which physical units are on the park track depends among others on the sides at which the arriving units enter the track. Let us define the variable $S(u)$ in the following way:

$$S(u) = \begin{cases} 0 & \text{if train unit } u \text{ arrives via the } A\text{-side} \\ 1 & \text{if train unit } u \text{ arrives via the } B\text{-side} \end{cases}$$

This variable has to be linked to the apt-variables:

For all *apt* where c_{apt} enters pt_{apt} via the A-side ($si_{apt} = A$) we add the constraint:

$$S(u) \leq 1 - x_{apt} \quad \forall u \in c_{apt} \tag{14.10}$$

For all *apt* where c_{apt} enters pt_{apt} via the B-side ($si_{apt} = B$) we add the constraint:

$$S(u) \geq x_{apt} \quad \forall u \in c_{apt} \tag{14.11}$$

Note that these constraints are only restrictive if $x_{apt} = 1$.

Given two arriving units u and u' and a track open at both sides there are four combinations of arrival sides for these units.

- If u enters via the A-side and u' via the B-side, then $a(u, u')$ has to be zero, because if they enter the same track, u is more close the A side than u'. The constraint that realizes this is:

$$a(u, u') \leq 1 + S(u) - S(u') \quad \forall u, u' \in U^+ \tag{14.12}$$

  This constraint is only restrictive if $S(u) = 0$ and $S(u') = 1$. Note that also the value of the ordering variable of units that do not enter the same park track is fixed. Later on we will see that this does not matter.

- If u enters via the B-side and u' via the A-side, then $a(u, u')$ has to be one, by the same argumentation. We could insert the constraint $a(u, u') \geq S(u) - S(u')$ to fix the ordering variable in this case, but that is not needed: The constraints (14.9) in combination with $a(u', u) \leq 1 + S(u') - S(u)$ (= constraint of type (14.12) with the units in different order) are enough to let $a(u, u')$ be one if u' arrives via the A-side and u via the B-side.

- If both units enter via the same side, the value of the ordering variable is determined by the arrival**times** $\theta(u)$ and $\theta(u')$ at their park tracks. This is treated below.

14.5.3 Ordering variables and arrival times

We now restrict the ordering variables for units that enter via the same side. Given units u and u', there are two options:

- If both u and u' arrive via the A-side, then $a(u,u') = 0$ iff $\theta(u) > \theta(u')$.
- If both u and u' arrive via the B-side, then $a(u,u') = 0$ iff $\theta(u) < \theta(u')$.

Both options can be represented by the following constraints:

$$a(u,u') \leq \frac{\theta(u') - \theta(u)}{l_{ph}} + 1 + S(u) + S(u') \quad \forall u, u' \in U^{+} \tag{14.13}$$

$$a(u,u') \leq \frac{\theta(u) - \theta(u')}{l_{ph}} + 3 - S(u) - S(u') \quad \forall u, u' \in U^{+} \tag{14.14}$$

Let us explain these constraints.

- If $\theta(u) > \theta(u')$ the value of the first term of the right hand side of (14.13) lies in the interval $[-1, 0)$, because l_{ph} is the length of the planning period. This makes that if both u and u' arrive via the A-side ($S(u) = S(u') = 0$) the value of the right hand side lies in $[0, 1)$. Because of the integrality condition of $a(u,u')$ its value will be equal to zero. Note that constraint (14.14) is not restrictive in this case.
- If $\theta(u) < \theta(u')$ the value of the first term of (14.14) lies in the interval $[-1, 0)$. The other terms make the right hand side lying in $[0, 1)$ if both units enter via the B-side. The integrality condition of $a(u,u')$ forces its value to be zero. In this case constraint (14.13) is not restrictive.

Note that if both $\theta(u) = \theta(u')$ and $S(u) = S(u')$, then $a(u,u')$ is not fixed with the constraints presented until now. If two units enter the same park track at the same side at the same time, then by the constraints avoiding conflicts (type (13.6)) we are guaranteed that in this case the units are routed in the same composition. In this case the order is deduced from the internal ordering of the composition of the apt-variables. For these cases the constraints fixing $a(u,u')$ are presented below.

14.5.4 Ordering variables and compositions

If u and u' are shunted together, then the situation is quite easy. For every $apt \in APT$ where c_{apt} consists of two or more units, we know the ordering in the composition. The constraints are:

For $apt \in APT$ with c_{apt} consists of two or more units, for $(u, u') \in c_{apt}$ with u more close to the A-side then u':

$$a(u, u') \leq 1 - x_{apt} \tag{14.15}$$

14.5.5 Forbidden combinations of dpt's

Once we have fixed the value of the ordering variables we turn to the restrictions that guarantee that the units are not retrieved from the tracks in a wrong order. In these restrictions we exclude the dpt's in which the units are in a wrong order in the composition. Besides we exclude combinations of dpt's that cause a crossing. They are constructed below step by step: We first make constraints that have to be restrictive if some assumptions hold. Thereafter the assumptions are incorporated in the constraint.

Consider arriving units u and u' and departing units v and v'. Assume that u is matched to v and u' to v'. If there is a moment in time that both physical units are present on the same park track, then given the value of the ordering variable, the choice for a specific $dpt \in DPT_v$ makes that some $dpt' \in DPT_{v'}$ are prohibited.

Example. *Consider the situation in figure 16 on page 68. Assume that $d(u) = v$ and that a specific $dpt \in DPT_v$ is chosen. If si_{dpt}, the leaving side of unit v, equals A, then the prohibited dpt's are the ones for which the other unit leaves via the A-side at a time earlier than θ_{dpt}. If si_{dpt} equals B, then the other unit may not leave via the B-side at a later time than θ_{dpt}. Departing via the A-side is totally prohibited in this case.* □

So given matched units u and v and matched units u' and v', for each $dpt \in DPT_v$ we make constraints that exclude the combination of choosing both dpt and one of the forbidden dpt's from $DPT_{v'}$. Given a $dpt \in DPT_v$, let $P_{dpt} \subset DPT_{v'}$ be the set of such prohibited dpt's. This set is made under the assumption $a(u, u') = 0$. The constraints therefore have to be restrictive if the assumption $a(u, u') = 0$ holds and they may not be restrictive if $a(u, u') = 1$.

To exclude every combination of dpt with a dpt' from P_{dpt} only in cases of $a_{u,u'} = 0$ we make the following constraint:

$$x_{dpt} + \sum_{dpt' \in P_{dpt}} x_{dpt'} \leq 1 + a(u, u') \quad (14.16)$$

The left hand side can not be larger than two because the summation only contains dpt's from the set $DPT_{v'}$ from which at most one dpt can be chosen by the constraints of type (13.2). Thus this this constraint is only restrictive if $a(u, u') = 0$.

Note that P_{dpt} only contains dpt's with park track equal to pt_{dpt}. So only combinations of two dpt's with the same park track are excluded. This makes that it does not matter that some ordering variables are fixed for two units that do not enter the same park track: The departing units to which these units are matched will by constraints (14.3) also be retrieved from different tracks and such combinations are not excluded by the constraints of type (14.16).

We want this constraint to be restrictive only if the assumptions as made above hold. These assumptions are:

- Unit u is matched to v and unit u' to v'.
- There is a moment in time that both physical units are present on their park track.

These assumptions can be incorporated in the constraint. The first one by adding to the right hand side the expression: $2 - M_{u,v} - M_{u',v'}$ that is zero if both matchings are made and larger otherwise. The second one by adding also $2 - T_{u,v'} - T_{u',v'}$, because this expression is zero if $T_{u,v'} = T_{u',v} = 1$: unit v' leaves after u has entered and v leaves after u' has entered. Together with the assumption that u is matched to v and u' to v', that is by constraints of type (14.8) only possible if v leaves after u entered and v' leaves after u' entered, it follows that there is a moment in time that both units are present on the track if $M_{u,v} = M_{u',v'} = T_{u,v'} = T_{u',v} = 1$. After adding these terms we end up with:

For all $(u, v), (u', v') \in Q$, for all $dpt \in DPT_v$:

$$x_{dpt} + \sum_{dpt' \in P_{dpt}} x_{dpt'} \leq 5 + a(u, u') - M_{u,v} - M_{u',v'} - T_{u,v'} - T_{u',v} \quad (14.17)$$

The following example explains the need of expression $2 - T_{u,v'} - T_{u',v'}$ in the constraint above.

Example. *Consider two units u and u′. If unit u′ enters at 12.01 via the A-side and u enters at 14.01 also via the A-side, the value of* $a_{u,u'}$ *is zero. Consider departing units v and v′ with* $(u, v), (u', v') \in Q$*. Combinations of dpt's* $dpt \in DPT_v$*,* $dpt' \in DPT_{v'}$ *with* $si_{dpt} = si_{dpt'} = A$ *and* $\theta_{dpt'} < \theta_{dpt}$ *are prohibited,* ***except for dpt's with*** $\theta_{dpt'} < 14.01$*, because in that case the physical unit reprensted by u′ and v′ is gone before unit u arrives. The value of* $T_{u',v}$ *is zero in this case (see constraint* (14.6) *at page 66) what makes that the right hand side is at least two and the constraint is not restrictive any more.* □

Besides constraints of type (14.17) we need constraints that prohibit that units are shunted in a composition that is of the wrong order. If $dpt \in DPT_v \cap DPT_{v'}$ and v and v' are not in the right order, then we can directly insert the constraint

$$x_{dpt} \leq 3 + a(u, u') - M_{u,v} - M_{u',v'} \tag{14.18}$$

Note that the time difference variables are not needed here, because if $x_{dpt} = 1$ for a *dpt* with $v, v' \in c_{dpt}$, by constraints of type (14.8), (14.6) and (14.7) we know that both physical units represented by v and v' are present on the track before θ_{dpt}.

14.6 Example

We conclude this section with an example. In this example we show how the constraints of this section work together to avoid infeasible shunt plans. Consider a track open at the A-side where first composition (`y,z`) (physical unit `y` more close to the A-side than physical unit `z`) arrives and thereafter composition (`x`) so that we get the situation as presented in figure 17.

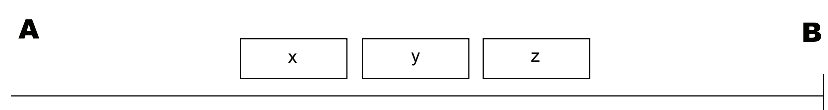

Figure 17: Situation after composition (y, z) and (x) have arrived

We show which combination of constraints avoids that composition (y) departs before (x) departs or that (x, z) or (y, x) (in the wrong order) leaves.

- Composition (y) can not leave earlier than (x). This is avoided by a constraint of type (14.17), see page 72. We treat the variables that are part of this constraint.

- The value of $a(x, y)$ equals zero because of constraints of type (14.13). Both units enter via the A-side and $\theta(y) < \theta(x)$.
- The matching variables are both one.
- Both time difference variables are one, by constraints of type (14.7), because unit x arrives before y leaves and unit y arrives before x leaves.
- The combination of dpt's where unit y leaves earlier than x is prohibited if $a(x, y) = 0$, so such dpt of unit y is part of the prohibited set of a dpt of unit x.

So the right hand side equals one in this case and the left hand side contains the combination of dpt's where composition (y) departs earlier than (x), so that this combination can not be chosen.

- If composition (x, z) leaves, then by constraints avoiding conflicts (type (13.6)), composition (y) has to leave $tir(VnaVNaarDezelfdeRiRR)$ minutes earlier or later.

  - If (y) leaves earlier than (x, z), then a crossing between unit x and y occurs. This is prohibited by constraints of type (14.17). This is avoided in the same way as described in the preceding bullet point.
  - If (y) leaves later than (x, z) we have a crossing between units y and z. The value of $a(y, z)$ is zero, by constraints of type (14.15) that ensure that ordering variables for two units that enter in the same composition are set to the right value. The remaining of the argumentation goes the same as the situation above.

  Thus composition (x, z) can not depart from the track in this situation.

- Composition (y, x) (y more close to the A-side than x) can not leave because constraints of type (14.18) forbid every dpt where y and x are in this order if $a(x, y) = 0$. This value of $a(x, y)$ is indeed zero, because of the constraints of type (14.13): y arrived earlier than x and via the A-side.

15 Freedom in time

In this section we describe how the APT-model is able to realize profit by using freedom in time. We distinquish between three ways in which the freedom in time can be used.

- Changing the order in which compositions are routed: If a composition is holding on its platform track for some time, it is possible to realize a different order of units on a park track. This can save shunt movements for departing units.

- Routing compositions with units from different trains: If some units are holded on their platform track until the next arriving units enter the same track, a composition with units from different trains can be made. This saves shunt movements for arriving units. Moreover, this can imply that these units stand in a different order on the track what can also save shunt movements for departing units.

- Parking on platform tracks: If a composition is routed from its arrival platform track to its departure platform track directly, we save one shunt move. Moreover, when a unit departs from the platform track at which it is arrived two shunt movements are saved.

The first option is present in the APT-model: If the whole set of apt's and dpt's is inserted, all possible orders of units on park tracks can be made. However this whole set is far to large to use. To be able to take advantage of the exchange of units, we present a two-stage solution method. This method has been implemented and is described section 17.

As mentioned in subsections 13.3 and 13.4 respectively the other two options can be inserted in the apt model easily. The APT-model is well suited to allow parking on platform tracks and routing compositions with units from different train services. In this section we describe how these extensions can be inserted.

15.1 Routing mixed compositions

To realize profit by allowing routing compositions with units from different trains some small adaptions to the APT-model have to be made. These adaptions are described in this subsection. We distinguish two types of adaptions:

- Additions to the sets APT and DPT

- Additions to the constraints

Firstly, the possibility of routing units from different train services together must be detected. Secondly extra apt's/dpt's have to be made to make these 'mixed compositions' possible. Moreover, some constraints are needed to avoid crossings on platform tracks.

The detecting of the possibility to route mixed compositions is done during calculating the shunt windows. According to section 7, the upperbound of the shunt window of unit $u \in U^+$ is determined by the first train service entering pl_u that

- enters via the other side than it leaves *or*
- enters via the other side than s_{t_u} has left (or if there is no s_{t_u}: the other side than $r^0_{t_u}$ entered) *or*
- has length greater than the length of pl_u minus l_u.

Assume that before the end of the shunt window of a unit $u \in U^+$ a train service $t \in \mathcal{T}$ enters pl_u. Assume that from t a unit u' is decoupled that is of the same *family* as u. According to subsection 1.2 unit u and u' can be coupled and driven together. Note that this is only possible if unit u' is decoupled from the *front* side of train service t.

Example. *Consider a unit of type* VIRM 4 *at its platform track (figure 18). During its shunt window train service t with composition (*VIRM 6, VIRM 4*) enters this track, see figure 19. Assume that the rolling stock circulation prescribes that the right most unit departs as train service s_t. The units that remain on the platform track are of the same family* VIRM *and can be routed together.* □

Figure 18: Unit `virm 4` on its platform track

Concluding, the assumption that t enters pl_u **during the shunt window of** u implies that apt's with composition (u, u') can be made and added

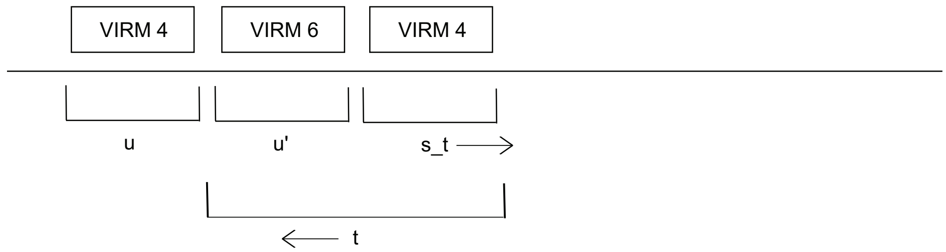

Figure 19: Mixed composition (`virm 4, virm 6`) remains

to APT. During making the set of apt's and generating constraints, this composition can be handled in the same way as a normal composition. The internal ordering is deduced from the arrival side of train service t.

Note that it is not known in advance whether it is profitable to shunt u and u' together. Therefore also apt's for u' shunting single are made. To fully use freedom in time, also unit u can get apt's with shunt time after arrival of t. Constraints of type 13.9 avoid some possible crossings on pl_u between unit u and u'.

15.2 Parking on plaftform tracks

To realize profit by allowing parking on platform tracks some small adaptions to the APT-model have to be made. These adaptions are described in this subsection. Besides the additions to the apt and dpt set also extra constraints are needed.

We now focus on the case that an arriving unit is directly routed to the platform track from which it departs, as described in option two of subsection 3.1. Thereafter, in 15.2.4, we discuss the additions for the case that these two platform tracks are equal and no routing is needed. As mentioned in 3.1, routing units directly to a platform track is only allowed if the units also depart from that platform track. In other words, for every unit of a composition that is routed to a platform track holds that it has to arrive there in the shunt window of a departing unit to which it can be matched. Otherwise it can not depart from that platfrom track in a train service.

Units that are directly routed to their depature platform track give some difficulties. Consider a unit u that can directly be routed to a platform track where it arrives in the shunt window of unit v to which it can be matched. We have to encounter the following aspects:

- Both u and v must have an apt resp. dpt assigned with $pt(u) = pt(v)$. Otherwise they may not be matched (see constraints of type (14.3)). We choose the convention that the park track of such units is always equal to the platform track of the **departing** unit, so $pt(u) = pt(v) = pl_v$.
- During parking at pl_v the length of this track may not be exceeded.
- During parking at pl_v no crossings may occur.

15.2.1 Special dpt's

If we allow for a unit $u \in U^+$ an apt which brings u directly to the platform track of the unit $v \in U^-$ to which it is matched, we have the problem that we can not assign a dpt for v, because it is already on its platform track after executing the route implied by this apt. Therefore, we have to introduce a dummy dpt that v gets if such an apt is chosen. This special dpt does not represent a route, but the other characteristics have to be such that the remaining of the framework of the APT-model may stay the same.

Given a unit u and an *apt* in which u routes directly to a platform track ($pt_{apt} \in S_{pl}$) and enters there in the shunt window of a departing unit ($\theta_{apt} \in sw(v)$) of a $v \in U^-$ to which it can be matched ($(u, v) \in Q$ and $pl_v = pt_{apt}$). Then in DPT there has to be present a special *dpt* for unit v with $pt_{dpt} = pl_v$. The τ_{dpt} and θ_{dpt} of this dpt have to be chosen so that

- the length on pl_v is well administrated and
- no crossings on pl_v occur.

This leads to the choice $\tau_{dpt} = \theta_{apt}$ and $\theta_{dpt} = \tau_{t_v}$. These choices are explained below:

- Its shunt time τ_{dpt} has to be the time that v enters its platform track, θ_{apt}. That is because of
  - the procedure of making constraints of type 13.4 administrating the total length of the units on a platform track. They count the length of unit v only between the shunt time $\tau(v)$ and departure time in service t_v.

- the procedure of making constraints of subsection 13.7. In making these constraints, we use the shunt time of v to ensure that units departing in the same train t_v enter pl_v in the right order.
- the procedure of making constraints in subsection 13.8 of type (13.9) avoiding that the side at which v is parked is used by another apt or dpt between $\tau(v)$ and τ_{t_v}.

- Its time of leaving the park track θ_{dpt} equals the time that t_v departs, namely τ_{t_v}. That is because of the constraints of type (14.17) avoiding crossings on park tracks. These constraints use the ordering variables that are linked to the θ of the chosen apt's and dpt's (see 14.13 and 14.14). These constraints are needed to ensure that
  - the composition of t_v has the units that come from c_{apt} in the right order
  - if the composition of apt also contains units that depart in other train services, that no crossing between the units of c_{apt} occur on pl_v.

Taking into account that crossings have to be avoided, also the entering and leaving sides of pl_v have to be known. Therefore si_{dpt} and $plsi_{dpt}$ have to be chosen properly:

- The side at which v enters its park track pl_v equals the side at which apt enters this track, si_{apt}.
- The side at which v leaves its park track equals the side at which t_v departs.

If apt is put in APT, for every unit in c_{apt} a dpt with the properties described above is added to DPT. These are called the *associated dpt*'s of *apt*. The requirements $\tau_{dpt} = \theta_{apt}$ and $plsi_{dpt} = si_{apt}$ introduce an extra relation between the choice of an apt with $pt_{apt} \in S_{pl}$ and the choice of a dpt. Below we describe how this relation is inserted in the APT-model.

15.2.2 Extra constraints

Given an *apt* and for every unit of c_{apt} an associated *dpt*, a constraint has to be inserted forcing to choose *dpt* if *apt* is chosen. For all *apt* with $pt_{apt} \in S_{pl}$, for all associated *dpt* of *apt*:

$$x_{dpt} \geq x_{apt} \tag{15.1}$$

This makes that if $x_{apt} = 1$, then $x_{dpt} = 1$.

15.2.3 Example

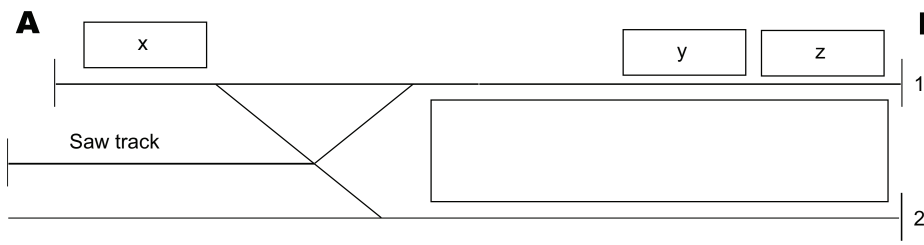

Figure 20: Situation at 11h48

Consider figure 20.

- Consider arriving train service t_1, with $pl_{t_1} = \texttt{1}$, $\tau_{t_1} = 11.48$ and composition (y,z)
- Consider departing train service t_2, with $pl_{t_2} = \texttt{2}$, $\tau_{t_2} = 12.12$ and composition (x,y,z)

Assume that the start of the shunt window at the A-side of track 2 of departing units y and z is before 11.48. This implies that these units after arrival can be routed to track 2 directly. Therefore an *apt* for the composition (y,z) is made with $pt_{apt} = \texttt{2}$. It is routed via track *Saw track* and has departure times 11.51 and 11.57. This implies that $\theta_{apt} = 11.59$.

The special associated *dpt* with $c_{dpt} = (\texttt{y,z})$ has to be made with the following fields: $\tau_{dpt} = 11.59$, $pt_{dpt} = \texttt{2}$, $plsi_{dpt} = A$, $\theta_{dpt} = 12.12$ and $si_{dpt} = A$, the leaving side of t_2.

- The $\tau_{dpt} = 11.59$ makes sure that unit x enters track 2 **after** 11.59 via the A-side if constraints of type (13.7) are applied. Moreover, this ensures that the lenght of units y and z is counted between 11.59 and 12.12, see constraints of type (13.4).
- The $\theta_{dpt} = 12.12$ allows the arriving units of t_1 to match with the departing units of t_2, because then $\theta_{apt} < \theta_{dpt}$ what makes that if *apt* and *dpt* are chosen constraints of type (14.8) are not restrictive. Moreover, this ensures that units y and z can not depart as composition (z,y), because if $\theta(\texttt{y}) = \theta(\texttt{z})$ only the constraints of type (14.15) are restrictive. The ordering variables of the arriving units take the internal ordering of the composition of *apt* which is (y,z) and the constraint of type (14.18) forbid then each *dpt* with internal ordering (z,y).

Constraint $x_{dpt} \geq x_{apt}$ is inserted.

To clarify the choice of the value of θ in the special dpt's we change this example a little bit. Assume

- that the composition of departing train service t_2 is (x,y) and
- that there is another departing train service, t_3, with $pl_{t_3} = 2$, $\tau_{t_3} = 12.42$ and composition z.

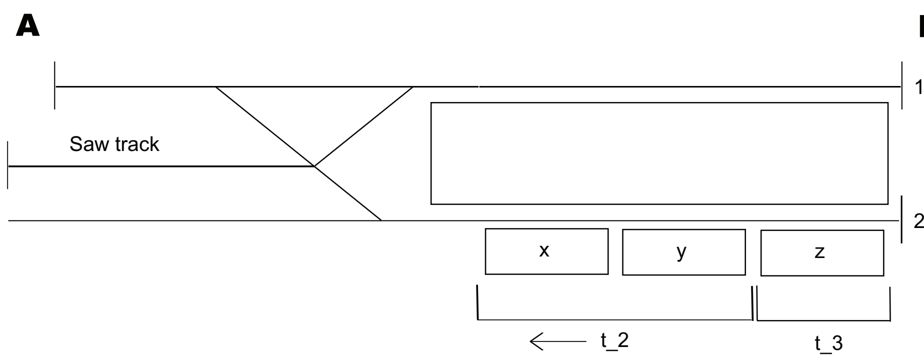

Figure 21: Situation at 12h12

Consider figure 21. If the shunt window of departing unit z still starts before 11.48, the same *apt* as above can be used to route both y and z to track 2. Now two associated dpt's have to be made, for both one.

1. dpt_1, with $c_{dpt_1} = (\mathtt{y})$, $\tau_{dpt_1} = 11.59$, $pt_{dpt_1} = 2$, $plsi_{dpt_1} = A$, $si_{dpt_1} = A$ and $\theta_{dpt_1} = 12.12$.
2. dpt_2, with $c_{dpt_2} = (\mathtt{z})$, $\tau_{dpt_2} = 11.59$, $pt_{dpt_2} = 2$, $plsi_{dpt_2} = A$, $si_{dpt_2} = A$ and $\theta_{dpt_2} = 12.42$.

Both dpt_1 and dpt_2 can be chosen if *apt* is chosen, because $\theta_{dpt_1} < \theta_{dpt_2}$, both leave via the A-side and the ordering variable $a(\mathtt{y},\mathtt{z})$ is zero by constraints of type (14.15). This makes that dpt_1 and dpt_2 do not appear together in the left hand side of a constraint of type (14.17).

Both constraint $x_{dpt_1} \geq x_{apt}$ and $x_{dpt_2} \geq x_{apt}$ are inserted.

Mind that in this case some extra constraints are needed to avoid crossings between unit z and other units departing from track 2 between 12.12

and 12.42 that not belong to t_3: If for example the composition of t_2 is `y,x`, then unit `x` enters track `2` before 11.59, θ_{apt}, see figure 22, because then it is more close to the A-side than unit `y`. Then unit `z` and `x` have a crossing at 12.12. □

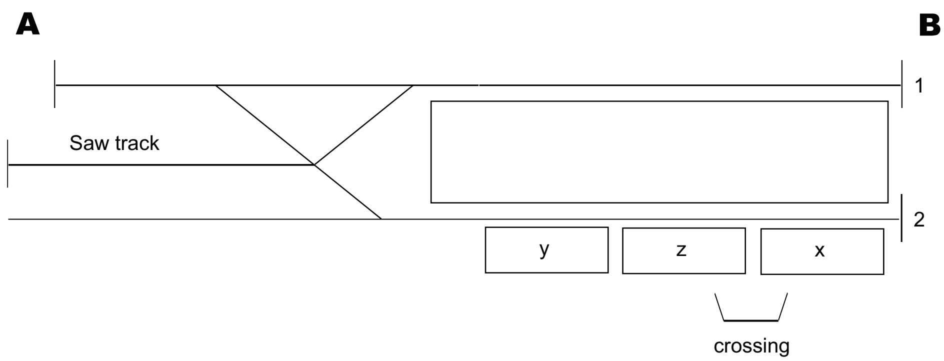

Figure 22: Crossing between `x` and `z`

15.2.4 Without routing

If an arriving unit $u \in U^+$ can be matched to a departing unit $v \in U^-$ that departs from the same platform track ($pl_v = pl_u$), also a special apt has to be made. The physical unit represented by u and v does not have to move between arrival and departure of the station. Again, detecting this possibility is done by considering the shunt windows: If τ_{t_v} is in the shunt window of u, an *apt* for unit u with $pt_{apt} = pl_u$ can be made and inserted to APT. Again also an associated *dpt* for v has to be added to DPT.

The *apt* has the following characteristics:

- $pt_{apt} = pl_u$
- $\tau_{apt} = \tau_{t_v}$
- $plsi_{apt}$ equals the side t_v leaves the track
- $\theta_{apt} = \tau_{t_u}$
- si_{apt} equals the side t_u enters the track
- The total lenght of the units in *apt* has to be equal to zero

Given this *apt* the associated *dpt* is constructed in the same way as described above. Then

- unit u and v may be matched,
- the length of pl_v is not exceeded by the units parked there and
- no crossings occur on pl_v.

We explain the choices for this *apt*:

- The time τ_{apt} and side $plsi_{apt}$ are used to make the constraints avoiding crossings on track pl_v between units of train service t_u (subsection 13.7). The departure time and side of t_v are the real time and side unit u leaves pl_v if *apt* is chosen, so are suitable to detect crossings.
- The time τ_{dpt} and side $plsi_{dpt} = si_{apt}$ are used to make constraints avoiding crossings on track pl_v between units of train service t_v. The real time unit v enters pl_v equals the entering time train of service t_u.
- The tracks pt_{apt} and pt_{dpt} are equal and $\theta_{apt} < \theta_{dpt}$, thus units u and v may be matched
- By calculating the length of the units on pl_v the units of c_{apt} would be counted twice: unit u between τ_{t_u} and $\tau(u)$ and unit v between $\tau(v)$ and τ_{t_v}. Therefore we defined the lenght of units in *apt* to be zero.
- Crossings between other units parked at pl_v are prohibited, because the entering time and side θ_{apt}, si_{apt}, θ_{dpt} and si_{dpt} are chosen as the real time and side unit u enters and unit v leaves if this *apt* and this *dpt* are chosen.

The choice of *dpt* for unit v in case *apt* is chosen is ensured by constraints 15.1.

Part V
Solution method

In part IV we described a MIP for the shunt problem. Given sets APT and DPT this MIP chooses the best subset of these sets, given an objective. Ideally we would generate all possible apt's and dpt's and insert them in the MIP. This would give the best possible solution. However, generating these whole sets would give far to much apt's and dpt's to build the MIP and solve it.

Therefore we start with two smaller and in practise manageable sets APT_0 and DPT_0, called the *start sets*. The MIP working on these sets is denoted by $MIP(APT_0, DPT_0)$ and described in section 16.

In subsection 16.2, we describe how these start sets are made. Compared to the whole sets APT and DPT, in the startsets we do not have flexibility in shunt time any more. For every composition and parktrack only the apt/dpt with smallest dwelling time on the platform track is inserted. However, the flexibility in the choice of compositions and park tracks is present in the start sets.

In section 17 we introduce a solution method that consists of two stages. The first stage uses the same start sets, but in the second stage we extend these sets such that the freedom in time is back for compositions for which it is profitable to stay longer on the platform track.

For both solution methods we describe test cases and some approaches with the aim of reducing the calculation time.

16 One-stage solution method

In this section we describe a solution method that consists of one stage. We describe test cases and a real life case at station Enkhuizen in the Netherlands. We also report calculation times and the effect of some approaches to reduce the calculation time.

16.1 Cost function

In this subsection we propose a cost function. We only assign costs to the apt- and dpt-variables, because the routing effort is mainly present in these choices. As mentioned in the beginning of section 3 the scheduling of train drivers that drive the compositions takes place when the shunt plan has been fixed. By minimizing the route effort, we try to ease this process.

For every apt a driver has to be scheduled, except for apt's with $pl_{apt} = pt_{apt}$. These apt's do not represent a movement and get costs zero. The other apt's get fixed costs C_{main}. Besides these fixed costs, for every minute between departure and arrival costs are added with weigth $C_{drivingtime}$. The driving time of an *apt* equals $\theta_{apt} - \tau_{apt}$ and represents how cumbersome *apt* is: if R_{apt} contains saw movements, the driving time is higher and thus the costs are higher. For a *dpt*, the driving time equals $\tau_{dpt} - \theta_{dpt}$.

This leads to the following cost function:

$$\sum_{apt \in APT, pt_{apt} \neq pl_{apt}} C_{apt} x_{apt} + \sum_{dpt \in DPT, pt_{dpt} \neq pl_{dpt}} C_{dpt} x_{dpt} \qquad (16.1)$$

where

$$C_{apt} = C_{main} + C_{drivingtime}(\theta_{apt} - \tau_{apt})$$
$$C_{dpt} = C_{main} + C_{drivingtime}(\tau_{dpt} - \theta_{dpt})$$

Introducing part splits can easily be discouraged by assigning a high value to C_{main} in comparison with $C_{drivingtime}$. Routing directly to the departure platform tracks is also encouraged by a high C_{main}. If the units from two different train services are shunted together the cost function is C_{main} lower. Also, if a unit is directly brought to its departure platform track we save C_{main} and if the departure platform track is equal to the arrival platform track we again save C_{main}. It costs C_{main} extra to introduce a part split and bring two units from the same part to different park track or retrieve two departing units from different park tracks.

Note that many aspects that influence how good a shunt plan is (see section 3) are left out. For example, reducing the dependency of the shunt plan to the timetables or improving the buffering effect of the shunt plan of delays in the execution of the timetables can be done afterwards by shifting some minutes without inducing a conflict or a change of order. By the choice of the cost function as described above we make these aspects inferior to the routing effort.

16.2 Making the start set

In this subsection we propose subsets of apt's and dpt's that form the start sets APT_0 and DPT_0 of the solution methods described in this part. The way the start set is made is based on the procedure as described in subsection 13.3. We again focus on the set of apt's. In the procedure described in 13.3

for every composition and park track an apt is inserted **for every possible minute.** In the start sets for every composition and park track **only one** apt/dpt is inserted. In order to avoid conflicts between compositions of the same part in the apt's and dpt's of APT_0 and DPT_0 the shunt windows of some compositions are changed. Below, we treat the differences with the procedure of subsection 13.3 in more detail.

- The first difference comes up in the calculation of the shunt window. This calculation is slightly different for compositions that are result of a part split. This is done to avoid conflicts between compositions out of the same part.

  Consider an arriving part with two units a and b where unit a is more close to the A-side. The lowerbound of the shunt window for b at the A-side equals the start of the shunt window for unit a at the A-side, increased with $tir(VnaVNaarDezelfdeRiRR)$ minutes. This makes that both compositions can leave without conflicts via the A-side. To be able to leave the platform track via the B-side without conflicts, the lower bound of the shunt window for unit a at the B-side equals the start time of the shunt window for unit b at the A-side, increased with the same value. For departing parts, the upperbound of the shunt window for compositions that are result of a part split are descreased with the value $tir(AnaAUitDezelfdeRiRR)$ if there are other units departing in the same train service that are more close to the entering side of the platform track.

  Example. *Applying these shunt windows to the units of example 13.1 at page 54 results in the absence of the apt's with numbers 11 t/m 13.* □

- The most important difference where the making of the start set differs from the description in 13.3 is that for every composition and for every route at most one apt is made. Given a composition and a route, once a feasible shunt time has been found for that route, the procedure turns to the next route.

  The search for feasible shunt times for routes starts for apt's at the beginning of the shunt window. Also at the saw tracks, every apt is extended for at most one departure time for every route. For dpt's the search for feasible shunt times starts at the **end** of the shunt window. For every composition there is searched for routes from park- or saw tracks that lead to the platform track of the units of the composition.

Thereafter the dpt's are extended by working **backwards**. Every apt and dpt may contain at most one saw movement.

As a consequence, for every composition of arriving units and every park track the first feasible apt to reach that park track (given the timetables and infrastructure) is in the start set and for every composition of departing units and every park track the latest feasible dpt to retrieve this composition from that park track is in the start set. If more than one route leads to the same park track s, more apt's/dpt's to/from track s are inserted in the start set. This is especially useful if these routes have different via tracks, because of avoiding blocking conflicts.

Example. *The start set of apt's of the units of example 13.1 contains only three apt's: 1, 6 and 14.* □

16.3 One-stage solution method

The Mixed Integer Program (MIP) consisting of

- the variables introduced in part IV for apt's and dpt's in the start sets APT_0 and DPT_0 as constructed in subsection 16.2,
- the constraints introduced in part IV except for (14.16) and
- the objective (16.1) as described in subsection 16.1

is reffered to by the 'one-stage solution method' and is denoted by

$$MIP(APT_0, DPT_0)$$

16.4 Example

In this subsection we present an example to show how the one-stage solution method works in practise. We firstly describe a fictive shunt yard as depicted in figure 23. This shunt yard is also used in the following subsections where we report calculation times for some cases. Thereafter we describe one specific case. The aim of this case is to show that the one-stage solution method is able to solve cases to optimality by considering routing, parking and matching integral.

16.4.1 Shunt yard

Consider figure 23. Tracks `1` and `2` are platform tracks. Tracks `1m` ... `5m` are park tracks. Track `xx` is used for sawing, but is available for parking too. Tracks `AA`, `BA`, `1` and `2` are not open for parking.

This shunt yard has some properties that are typical for the stations at which an integrated approach would be useful:

- The routes between the platform tracks and park track `2m` and `3m` are composed routes: Via track `1m` has to be empty if units are routed to and from track `2m` and `3m`.
- Track `xx` is a via track for the routes between the platform tracks and track `4m` and `5m`. However, when units are parked at `xx`, other compositions can saw at track `xx` if the length of `xx` admits.

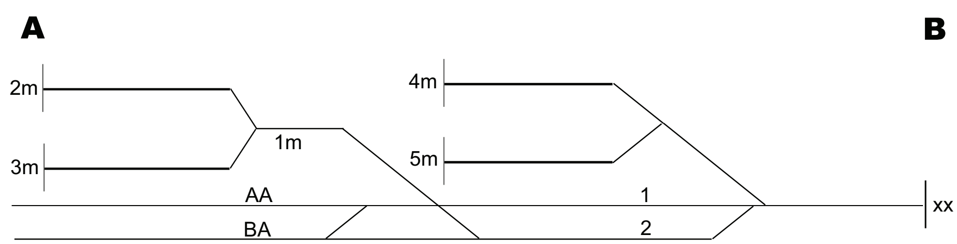

Figure 23: Lay out of the case

Below a table with the lengths of each park track is presented:

Track	Length
`1m`	150
`2m`	320
`3m`	300
`xx`	300
`4m`	500
`5m`	350

16.4.2 Fictive case

We present a case of arriving and departing units arriving on the fictive shunt yard. Below we present the set of arriving units. The *position* of a unit indicates the internal ordering of the units in a composition. The position of u is zero if t_u consists of one unit. If t_u consists of more units, a unit i is more close to the A-side then a unit j iff $i < j$. Besides, for every unit, its platform track, type, length and the planned time of its train service are displayed.

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
4267	0	2	LB	105	19h01
1771	1	2	OC	110	19h43
1771	2	2	OC	110	19h43
4281	0	2	LE	120	22h31
1783	1	2	OH	80	22h43
1783	2	2	OC	110	22h43
4283	0	2	LC	90	23h01
1785	1	2	OC	110	23h13
1785	2	2	OH	80	23h13
4285	0	2	LB	105	23h31
4287	0	2	LA	130	24h01

The set of departing units is:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
4214	1	1	LC	90	29h29
4214	2	1	LE	120	29h29
1718	1	1	OC	110	29h47
1718	2	1	OC	110	29h47
1718	3	1	OH	80	29h47
4216	0	1	LA	130	29h59
1720	1	1	OH	80	30h17
1720	2	1	OC	110	30h17
1720	3	1	OC	110	30h17
4218	1	1	LB	105	30h29
4218	2	1	LB	105	30h29

For this case the set APT_0 consists of 81 apt's and DPT_0 contains 102 dpt's. We set the parameter C_{main} to 100 and $C_{drivingtime}$ to 1, because we want to discourage $MIP(APT_0, DPT_0)$ to introduce part splits.

16.4.3 Output

Below the plan resulting from running the one-stage solution method is presented. In the first collumn we see the units that are shunted together in a composition. In the second collumn we find the park track and in the last collumn the shunt time that is chosen.

For example, in the second row, we see that both units of train service 1771 are shunted together to track `4m` at 19h46. In the sixth and seventh row we see that for the units of train service 1785 a part split is introduced: Unit [1785, 1] is brought to track `2m` and unit [1785, 2] to track `4m`. Both at 23h16.

Composition						*pt*	τ
4267	0					3m	19h4
1771	1	1771	2			4m	19h46
4281	0					xx	22h34
1783	1	1783	2			2m	22h46
4283	0					xx	23h4
1785	1					2m	23h16
1785	2					4m	23h16
4285	0					3m	23h34
4287	0					1m	24h4
4214	1	4214	2			xx	29h26
1718	1	1718	2	1718	3	4m	29h44
4216	0					1m	29h56
1720	1	1720	2	1720	3	2m	30h14
4218	1	4218	2			3m	30h26

The calculation time for this case is equal to 3.67 seconds and the costs are 1450. An overview of the situation after all arriving units have entered their park tracks is given in figure 24.

- Note that units [1785, 1] and [1785, 2] can depart at the same time. Unit [1785, 1] leaves track `2` via the *A*-side and unit [1785, 2] via the *B*-side.

- Note that while track `xx` is used for parking, unit [1785, 2] is routed to `4m` via saw track `xx`. Units [4281, 0] and [4283, 0] are parked there at that time. The length of these units plus the length of unit [1785, 2] equals 290 meter, which is smaller than the length of track `xx`, thus this routing is possible.

- Note that while track `1m` is used for parking, no routes to or from

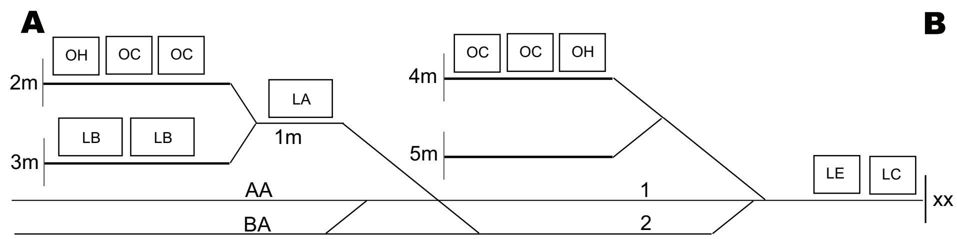

Figure 24: All units on the shunt yard

tracks `2m` and `3m` are planned.

Below we describe several other cases on this shunt yard. Also, in 16.5, a case of eleven units with longer calculation time to which we tried some methods to reduce the calculation time is presented.

16.5 Computational experiments

In this section we give an idea of the calculation time of the model as described in part IV and what we tried to reduce this time. For all calculations we use an Intel Xeon X5472, 3GHz, 3GB RAM machine. For solving the MIP we use Cplex 12.1.

16.5.1 Introduction

We use the shunt yard as described in 16.4.1. Only the length of track `xx` is changed to 200 meter, so that longer compositions can not park there anymore. The parameters C_{main} and $C_{drivingtime}$ stay on 100 respectively 1. We present a table to show how the calculation time increases if the number of units increases.

# units	calculation time
3	< 1
8	4.5
10	60

The last case with ten units consists of units of three different types. Three of type $\mathcal{I}$, three of type $\mathcal{II}$ and four of type $\mathcal{III}$. If we divide the units of the last type in two units of type $\mathcal{III}$ and two of type $\mathcal{IV}$ we end up with a calculation time of 15 seconds. The change of the instance has reduced the number of possible matchings. Therefore, it would be interesting to find out if the number of possible matchings is a better indicator for the calculation time.

By timing the two units of type $\mathcal{IV}$ such that they are the latest arriving and the first departing, we have an optimal outcome within only one second. It would be interesting to find out if cases where units arrive and depart following last in first out discipline have significantly smaller calculation time than cases where this discipline is not followed.

16.5.2 Case of eleven units

In this subsection we describe a fictive case to which we apply approaches to decrease the calculation time.

The arriving units are:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
2001	1	2	OH	80	22h01
2001	2	2	OC	110	22h01
2003	1	2	LE	120	22h31
2003	2	2	LE	120	22h31
2005	0	2	OC	110	23h01
2007	1	2	LC	90	23h31
2007	2	2	LE	120	23h31
2011	1	2	OH	80	24h31
2011	2	2	OH	80	24h31
2013	1	2	OC	110	25h01
2013	2	2	OH	80	25h01

The set of departing units is:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
2016	0	1	OH	80	31h29
2018	0	1	OC	110	31h59
2022	1	1	OH	80	32h59
2022	2	1	OC	110	32h59
2024	1	1	LE	120	33h29
2024	2	1	LC	90	33h29
2026	0	1	OH	80	33h59
2030	1	1	OH	80	34h59
2030	2	1	OC	110	34h59
2032	0	1	LE	120	35h29
2034	0	1	LE	120	35h59

Below the shunt plan resulting by running the one stage solution method is presented:

Composition				pt	τ
2001	1	2001	2	2m	22h04
2003	1			3m	22h34
2003	2			4m	22h34
2005	0			2m	23h04
2007	1			3m	23h34
2007	2			5m	23h34
2011	1	2011	2	4m	24h34
2013	1	2013	2	4m	25h04
2016	0			4m	31h26
2018	0			2m	31h56
2022	1	2022	2	4m	32h56
2024	1	2024	2	3m	33h26
2026	0			4m	33h56
2030	1	2030	2	2m	34h56
2032	0			4m	35h26
2034	0			5m	35h56

The costs of this shunt plan are 1786 and the calculation time is **50** seconds. This is larger than the calculation time of the case in subsection 16.4.2. Perhaps the increased number of possible matchings (from 27 to 32) constributes to this.

Below we present some methods to reduce the calculation time of the one stage solution method. We report the calculation times and reflect on them at the end of this subsection.

To be able to compare calculation times, we want to avoid different calculation times for identical runs of the same instance. Therefore we add small costs to parking specific units at specific tracks to avoid the occurence of two different optimal outcomes that have the same costs but different calculation times.

16.5.3 Reducing M in big-M constraints

The constraints of type 14.6 contain a 'big M'. It has the value of the length of the planning period. This is because the difference of $\theta(u)$ and $\theta(v)$ can be at most that large. But we can try to calculate the maximal difference out of the sets APT_u and DPT_v and replace the lenght of the planning period by this maximal time difference, so that we get constraints with smaller big-M's.

The value of this maximal difference is as follows: Given an arriving unit u and a departing unit v:

$$u_{max} := max_{apt \in APT_u} \theta_{apt}$$

$$v_{min} := min_{apt \in APT_v} \theta_{dpt}$$

If $v_{min} > u_{max}$, then $T_{u,v} = 1$. Otherwise we use big M with positive value $u_{max} - v_{min}$.

In the same way the constant in the constraints of type 14.7 are adjusted. If $v_{max} < u_{min}$, then $T_{u,v} = 0$, otherwise the constant is $v_{max} - u_{min}$.

The calculation time of $MIP(APT_0, DPT_0)$ with these reduced big M's increases to **82** seconds.

16.5.4 Aggregation

The constraints of type 14.10 and 14.11 can be aggregated. Given a unit $u \in U^+$, we know that at most one $apt \in APT_u$ can be chosen, so we make the constraints

$$S_u \leq 1 - \sum_{apt \in APT_u, si_{apt} = \text{A}} x_{apt}$$

and

$$S_u \geq \sum_{apt \in APT_u, si_{apt} = \text{B}} x_{apt}$$

for all u and substitute these constraints for the constraints of type 14.10 and 14.11.

By applying both aggregation and reducing big M, the calculation time again increases, to **125** seconds.

16.5.5 LP-relaxation of a and S-variables

In general the number of integer variables influences the time to solve a MIP. We take a closer look to the constraints and conclude that some variables do not need to be required integer.

The integer $S(u)$-variables that indicate at which side unit u enters its park track can be relaxed to numeric variables. The constraints that link these variables to the apt-variables, of type 14.10 and 14.11, set the value of $S(u)$ to zero or one exactly: The constraints of type 13.1 ensure that for every unit $u \in U^+$ exactly one $x_{apt},\ apt \in APT_u$ equals one so that one of the constraints of type 14.10 and 14.11 implies an equality for $S(u)$, given the lowerbound (zero) and upperbound (one) for $S(u)$.

The integer variables $a(u, u')$ that indicate the order of unit u and u' on the park track can be relaxed to numeric variables. Let us take a look at constraints of type (14.17) to explain this. For all values of $a(u, u')$ in the half open interval $[0, 1)$ this constraint is restrictive. (Perhaps this becomes more clear by looking at constraints of type (14.16).) So, if both units u and u' have to be retrieved from the same track, one of the variables $a(u, u')$ or $a(u', u)$ equals one. Constraints of type (14.9) ensure that the other value equals zero.

Applying

- Aggregation,
- Reducing big M and
- LP-relaxation S-variables

leads to a decreament of the calculation time: **86** seconds.

Applying the LP-relaxtation of the a-variables too leads again to a higher calculation time: **124** seconds. We can conclude that it is not necessarily true that solving a MIP with CPLEX takes less time if less integer variables are present.

16.5.6 Reflection

To summarize we present a table with the calculation times reported above. The adaptions are applied cumulative.

Adaptions	Calculation time
None	50
Reducing big M	82
Aggregation	125
LP-relaxation S-variables	86
LP-relaxation a-variables	124

Nearly all approaches lead to a worse calculation time, except for the LP-relaxation of the S-variables. We therefore try to only apply relaxing of S-variables and leave the other approaches out, but then we obtain a calculation time of 327 seconds, that is worse than the 50 seconds of the run in which the S-variables are not relaxed.

To know whether these approaches in general lead to a worse calculation time, we also apply them to other cases. In some cases the reducting of big M and the aggregation leads to a small improvement of the calculation time. By relaxing the a-variables the solution time increases in all cases that we tried.

In some runs where relaxing of a and S variables is applied, we see that a feasible solution was calculated earlier in the solving process than in runs where these variables are not relaxed. It would be interesting to know whether this is the case in general.

Due to historical reasons, for the computational experiments in the remaining of this report only the aggregation and reducing of big M are applied.

16.6 Case Enkhuizen

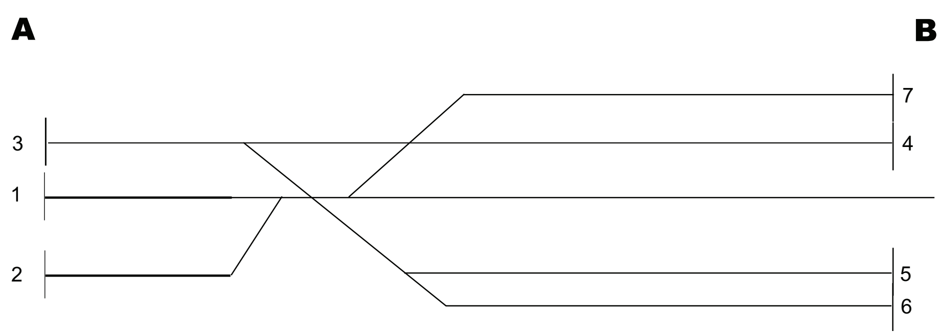

Figure 25: The shunt yard at Enkhuizen

We test the one stage solution method to a real instance at station Enkhuizen in the north-west of the Netherlands. Its shunt yard is depicted in figure 25. The lengths of the park tracks are:

Track	Length
3	641
4	603
5	623
6	623
7	600

The time period starts at the evening rush hours of january 8, 2008 and ends at the morning rush hours of january 9, 2008. In this period the arriving and departing units are nearly in balance. Only a composition $[60000, 0]$ is added.

The set of arriving units is:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
60000	0	1	LB	105	16h00
14556	0	2	L	175	17h48
14558	0	2	L	175	18h18
73460	0	3	AE	102	20h44
4568	0	1	LM	80	21h8
4576	0	1	LA	130	23h8
4580	0	1	LM	80	24h8
73497	1	1	LB	105	25h4
73497	2	1	LB	105	25h4
4586	0	1	LM	80	25h39

The set of departing units:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
4591	0	2	LM	80	23h23
3317	0	2	AE	102	28h42
4519	1	1	LM	80	29h30
4519	2	1	LM	80	29h30
4521	0	2	LA	130	29h54
4523	0	1	LB	105	30h23
14523	0	2	L	175	30h43
14525	1	2	LB	105	31h13
14525	2	2	LB	105	31h13
14527	0	2	L	175	31h43

The shunt plan resulting from $MIP(APT_0, DPT_0)$ contains the following decisions:

Composition				pt	τ
60000	0			4	16h03
14556	0			6	17h56
14558	0			5	18h26
73460	0			6	20h47
4568	0			7	21h11
4576	0			5	23h11
4580	0			4	24h11
73497	1	73497	2	7	25h07
4586	0			4	25h42
4591	0			7	23h20
3317	0			6	28h39
4519	0	4519	2	4	29h27
4521	0			5	29h51
4523	0			4	30h20
14523	0			5	30h40
14525	1	14525	2	7	31h10
14527	0			6	31h40

This plan has optimal value **1734** and is calculated in less than a second.

Also used in subsection 17.7

17 Two-stage solution method: Using the freedom in time

In this section we describe a solution method that consists of two stages. In comparison with the one stage solution method, the method described in this section uses some freedom in time by allowing longer dwell times on the platform track. In this way units are able to enter park tracks in different orders. This can save shunt movements. Both stages consist of one MIP solving and some pre-processing for generating apt's and dpt's.

17.1 Introduction

In the previous section we generated the start set APT_0 and DPT_0 such that every composition can take every route at most at one point in time: the shunt time that implies the shortest dwelling time on the platform track. One way to insert time freedom is to generate all possible apt's and dpt's as described in subsection 13.3. This method gives for every possible shunt time in the shunt window a new apt. Generating this whole set would give far too much apt's and dpt's to build the model and solve it. Therefore we start with the start sets as described in 16.2. After the first stage we extend these sets by adding apt's and dpt's with other shunt times. With these sets we solve the original MIP. In figure 26 the two state solution method is presented next to the one stage solution method in a schematic way.

In order to deduce which apt's and dpt's have to be added for the second stage, the MIP acting on the startset in the first stage is a little bit adapted compared to $MIP(APT_0, DPT_0)$. This is done by leaving out and relaxing some of the constraints. Let us call this MIP of the first stage $MIP'(APT_0, DPT_0)$. Note that an outcome of $MIP'(APT_0, DPT_0)$ does not necessary represent a feasible shunt plan, since no longer all constraints are incorporated in the MIP.

The idea is to use the outcome of $MIP'(APT_0, DPT_0)$ to deduce which apt's and dpt's should be added for the second stage. Let us denote the sets that are used for the second model by APT_1 and DPT_1. In the second MIP all constraints as described in part IV are incorporated, so that if there is an outcome, this outcome represents a feasible shunt plan. The second stage is denoted by $MIP(APT_1, DPT_1)$.

In this section the properties of the first and second stage are described in detail. But before, we explain the main idea behind this solution method.

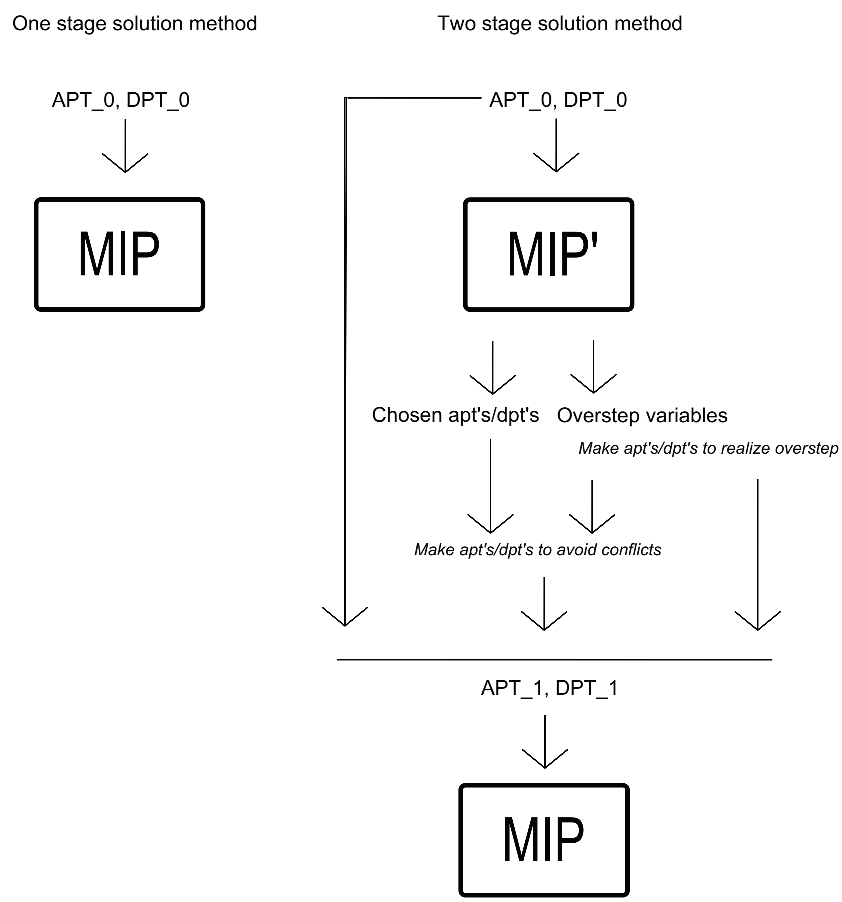

Figure 26: Schematic view on the solution methods

Also used in 17.5

17.2 Main Idea

The idea is to let $MIP'(APT_0, DPT_0)$ suggest a good **order** of the units on the park tracks. This is because a different order of the shunt units can save shunt movements. Inserting apt's and dpt's in such a way that every possible order on all park tracks can be achieved probably results in far too large sets. Therefore after the first stage, the sets APT_0 and DPT_0 are extended to APT_1 and DPT_1 by adding only apt's and dpt's that facilitate the order as suggested by the outcome of the first stage.

It remains to describe how we let $MIP'(APT_0, DPT_0)$ suggest a good order of the shunt units. Consider the ordering variables $a(u, u')$ that fix the order of units that are on the same park track. As described in subsection 14.5, besides the arrival sides and chosen compositions also the arrival **times** play an important role in the order of the units. The order in shunt time directly influences the order on the park track. However, the set APT_0 is such that only the first possible arrival time is available for every composition and every park track.

The main difference between MIP and MIP' now is that in MIP' besides the order as implied by APT_0 also other orders may be realized. This is done by introducing variables that represent how much later an apt arrives at its park track than prescribed by the arrival time θ of the apt. In the optimal solution these variables attain the values that realize the suggested order. From these variables in the optimal solution we want to know which apt's to make. They can make it possible that two compositions enter the park track in another order so that for example one dpt can be saved, but they can also make it possible that less saw movements are needed.

Adapting some constraints by inserting these variables in some way 'simulates' that there are more apt's, with later shunt times, than only the apt's in the start set APT_0. Before we turn to the precise description of both stages, we give an example.

17.3 Example

We illustrate the main idea with an example. Consider a shunt yard with two platform tracks (1 and 2) and two park tracks (3 and 4), see figure 27. Consider two units of different types arriving in two train services and leaving the station as one composition in a departing train service.

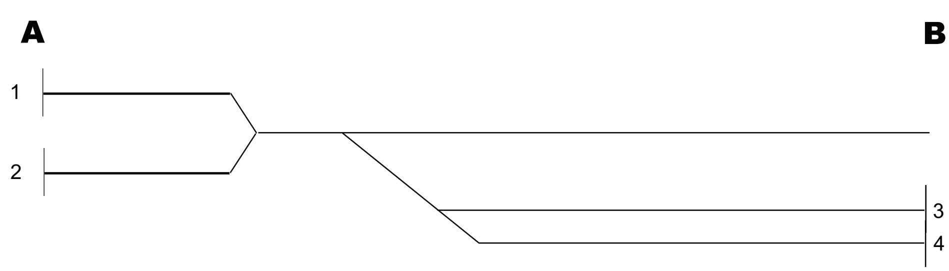

Figure 27: Shunt yard of example 17.3

Assume that

- a unit of type OA enters track 2 at 09.58,
- a unit of type AD enters track 1 at 10.28,
- a composition (OA,AD) departs from track 1 at 15.32.

Imagine that $MIP(APT_0, DPT_0)$ acts on this input. The units are planned at two different park tracks, because driving to the same track at the times as proposed by the apt's in APT_0 (at 10.01 and 10.31 resp.) gives that composition (AD,OA) (unit AD more close to the A-side than unit OA) arises, see figure 28. That is not the composition in which the units have to leave. Thus the units are parked and retrieved from different tracks, so that by retrieving them in the good order, the required composition arises. This results in four shunt movements.

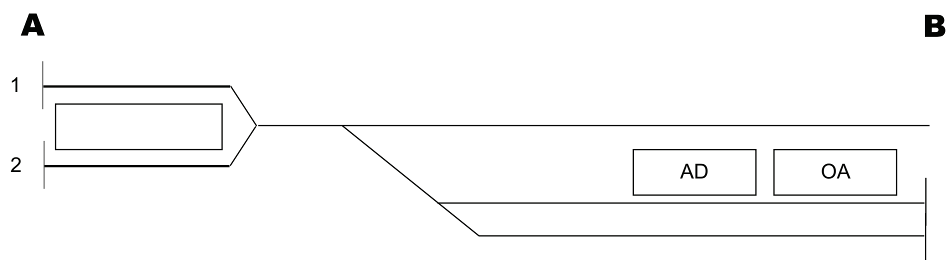

Figure 28: Units are in the wrong order

If track 2 is not used by train services after 9.58, then unit OA can dwell there for some time. If it is routed **after** unit AD, then it can enter the same track so that the right composition (OA,AD) (unit OA more close to the A-side than AD) arises. This composition can with one shunt move be retrieved from this track what saves one shunt move in comparison with the first solution of $MIP(APT_0, DPT_0)$.

The idea of the two-stage solution method is that the model of the first stage determines such possibility and suggests a later shunt time for OA, after AD arrived at its park track. An apt with this later shunt time is made and inserted in APT_1. Mind that in the first stage, there is no apt with this later shunt time: Some constraints are relaxed so that this possibility is 'suggested' by the first model.

Note that perhaps not all suggested apt's can be made. Moreover, even if all suggested apt's can be made, there is no guarantee that they can be used. More details about this can be found in subsection 17.5.2. Of course, the situation where the added apt's and dpt's are not used is tried to be avoided as much as possible.

17.4 First stage

In this subsection we treat $MIP'(APT_0, DPT_0)$ in detail.

In the first stage we use the start sets as described in 16.2. We build the model MIP' that is slightly different from MIP. We introduce three differences:

- We leave out the constraints of type (13.6) avoiding conflicts during driving.
- We also leave out the constraints of type (13.8) avoiding conflicts on saw tracks if two compositions are sawing there at the same time.
- Besides constraints of type (14.4) and (14.5), linking the θ variables to the choice of apt's and dpt's are adjusted.

The first difference is because of the limited number of shunt times in APT_0 and DPT_0: The model can choose for every composition and park track only one shunt time. Therefore we do not want to exclude a combination of apt's/dpt's that have a conflict, because probably this conflict can easily be avoided by adding small adapted apt's in the second stage. The same holds for the second difference. After the first stage, for the apt's/dpt's that are chosen by $MIP'(APT_0, DPT_0)$ extra apt's/dpt's are added if they have a conflict. This results in less apt's/dpt's then if we would add extra apt's/dpt's for all apt's/dpt's in the startset that have a conflict.

The third difference is introduced to realize the main idea as described in 17.2: facilitating freedom in time to obtain a suggestion for a good order. We now explain how the constraints linking θ to the chosen apt's are relaxed.

17.4.1 Overstep

To facilitate freedom in time, we introduce for each *apt* in APT_0 an *overstep variable* y_{apt}. It represents the number of minutes that composition c_{apt} enters its park track pt_{apt} later than θ_{apt}. We insert these variables in the model.

If a certain $apt \in APT_0$ is chosen, for all units $u \in c_{apt}$ the resulting arrival time $\theta(u)$ equals $\theta_{apt} + y_{apt}$. The variable $\theta(u)$ has to get this value because the ordering variables are linked to these arrival times to set the order of the units (see constraints of type (14.13) and (14.14)) and a different order can save shunt moves.

Note that the overstep variables may not be negative, because every apt in the start set already represents the earliest possible way to reach the park track. These variables also have an upperbound. For each $apt \in APT_0$ the 'maximal overstep' mo_{apt} is calculated. It represents the maximal number of minutes that composition c_{apt} can enter pt_{apt} later then θ_{apt}. This value is calculated using the upperbound of the shunt window. More about this can be found in 17.4.2.

From the value of these overstep variables in the optimal solution of $MIP'(APT_0, DPT_0)$ we know how to extend APT_0 and DPT_0 to APT_1 and DPT_1 for the second stage.

The constraints of type (14.4) linking the θ-variables to the choice of apt's for every arriving unit u are adapted by adding to the right hand side the value of $\sum_{apt \in APT_u} y_{apt}$. In the same way, the constraints of type (14.5) for every departing unit v are adapted by substracting $\sum_{dpt \in DPT_v} y_{dpt}$ where also $0 \leq y_{dpt} \leq mo_{dpt}$, the maximal number of minutes that composition c_{dpt} can leave its park track pt_{dpt} earlier then θ_{dpt}.

This results in:

$$\theta(u) = \sum_{apt \in APT_u} (\theta_{apt} x_{apt} + y_{apt}) \quad \forall u \in U^+ \tag{17.1}$$

and

$$\theta(v) = \sum_{dpt \in DPT_v} (\theta_{dpt} x_{dpt} - y_{dpt}) \quad \forall v \in U^- \tag{17.2}$$

We have to ensure that only the overstep variables of the selected apt's and dpt's are greater than zero. This is done by introducing the following constraints:

$$y_{apt} \leq x_{apt} mo_{apt} \quad \forall apt \in APT \tag{17.3}$$

$$y_{dpt} \leq x_{dpt} mo_{dpt} \quad \forall dpt \in DPT \tag{17.4}$$

As mentioned, an overstep can save shunt moves or saw movements. This gives that we want low costs for the overstep variables. On the other hand, an overstep for an apt or a dpt also introduces a longer dwell time at the platform track. Moreover, every chosen overstep gives work for the algorithm to add apt's and dpt's. This makes that we introduce a small cost C_o to every minute of overstep, so that the objective becomes:

$$\sum_{apt \in APT} (C_{apt} x_{apt} + C_o y_{apt}) + \sum_{dpt \in DPT} (C_{dpt} x_{dpt} + C_o y_{dpt}) \tag{17.5}$$

In the test cases we use $C_o = 0.03$ and for calculating C_{dpt} and C_{apt} we use $C_{main} = 100$ and $C_{drivingtime} = 1$. To be sure that no more minute overstep is introduced in the optimal solution than needed, the maximal absolute gap between the best possible solution and the outcome of the first stage is set to $C_o = 0.03$.

17.4.2 Maximal overstep

To calculate the maximal overstep and to explain the need for some more constraints (see 17.4.3) we first have to take a closer look at the overstep in the optimal solution of the first stage. Consider the arriving units OA and AD of example 17.3 with $pl_{\text{OA}} \neq pl_{\text{AD}}$. Assume that the start set contains among others the following apt's:

- apt with $\tau_{apt} = 10.01$, $c_{apt} = (\text{OA})$, $pt_{apt} = s$ and $\theta_{apt} = 10h03$
- apt' with $\tau_{apt'} = 10.31$, $c_{apt'} = (\text{AD})$, $pt_{apt'} = s$ and $\theta_{apt'} = 10h33$

As we saw in 17.3, the optimal solution contains a different order of units OA and AD at track s than implied by these apt's: the overstep suggested by the model of the first stage y_{apt} equals 30 minutes exactly, the smallest overstep value for which $a(\text{OA,AD})$ may become zero, so that in the optimal solution $\theta(\text{OA}) = \theta(\text{AD}) = 10.33$. That is because of two observations:

- The constraint linking the ordering variable $a(\text{OA,AD})$ to $\theta(\text{OA})$ respectively $\theta(\text{AD})$ (type (14.13)) is not restrictive for two units in different compositions that arrive via the same side at the same time.
- Thereby, we set the maximal possible gap between the optimal value and the outcome of the first stage at C_o so that always the lowest possible overstep is chosen.

To change the order of OA and AD at track s, unit OA has to depart $y_{apt} + tir(AnaAUitDezelfdeRiRR) = 33$ minutes later. This is only possible if the shunt time τ_{apt}+33 = 10.34 is in the shunt window. Therefore the value of mo_{apt} is calculated by the difference of the upperbound of the shunt window for OA and the τ_{apt} **subtracting** the value of $tir(AnaAUitDezelfdeRiRR)$.

In cases of apt's with sawing routes, there are some differences that we will not treat in detail. Also the calculation of the maximal overstep of dpt's is not treated here. We now turn to the extra constraints that are needed if the constraints (17.1) - (17.4) are in the model instead of (14.4) and (14.5)

17.4.3 Extra constraints

We treat the extra constraints as mentioned above. They have to ensure that compositions are not broken up by units with overstep. As we saw, in the optimal solution an apt *apt* gets overstep y_{apt} so that $\theta_{apt} + y_{apt}$ equals the arrival time of the composition to which it prefers to exchange. The constraints linking the ordering variables to the θ-variables (type (14.13) and (14.14)) are not restrictive for units from different compositions that arrive simultaneously at the same side of the same track.

This implies that the first model can also suggest the units of c_{apt} to take place in the **middle** of the composition to which it prefers to exchange without violating any constraint.

Example *Consider example 17.3. Now assume that the second arriving composition consists of **two** units of type* AD *instead of one and that composition (* AD,OA,AD*) has to depart. Without the constraints that we present below the model of the first stage would again suggest an overstep of 30 minutes and plan both on the same track, see figure 29.* □

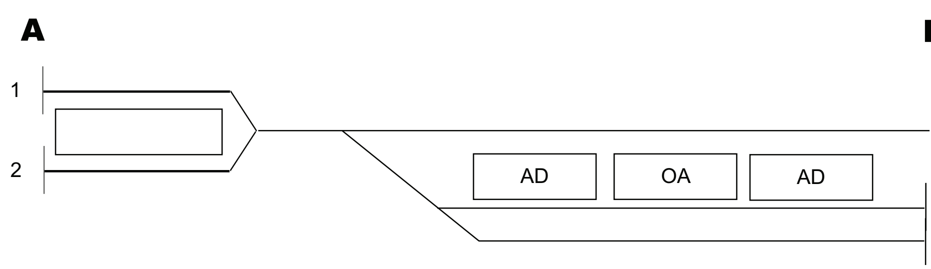

Figure 29: Unit OA can not be placed in the middle of composition (AD,AD)

Note that such an ordering costs more shunt movements, because it implies a part split of the two units of type `AD`. So we do not want the first model to suggest this. Below we present constraints that avoid such orderings that brake up a composition.

Consider an *apt* with composition $c_{apt} = (c_1, c_2, \ldots, c_n)$. Assume that c_i is more close to the A-side than c_j iff $i < j$. We want to prohibit that a unit $u \notin c_{apt}$ is placed between two units of c_{apt}. That is the case if unit u is more close to the A-side than c_j while a c_i with $i < j$ is more close to the A-side than u. This is the case if both $a(u, c_i)$ and $a(c_j, u)$ are one. So the constraints become:

For all $apt \in APT$ for which c_{apt} consists of more than one unit, for all arriving units $u \notin c_{apt}$:

$$x_{apt} + a(u, c_i) + a(c_j, u) \leq 2 \quad \forall (c_i, c_j) \in c_{apt},\ i < j \tag{17.6}$$

If units c_i and c_j are routed together, then $x_{apt} = 1$. This gives a restriction to the values of $a(u, c_i)$ and $a(c_j, u)$, stating that not both may be one. If the units are not routed together, this constraint is not restrictive.

Example. *Unit* OA *of example 17.3 above would act as unit u in this constraint and both units of type* AD *as* c_1 *and* c_2*.* □

17.4.4 Definition MIP'

The MIP consisting of

- the variables introduced in part IV for apt's and dpt's in the start sets APT_0 and DPT_0 as constructed in subsection 16.2 and the y-variables as introduced in 17.4.1,
- the constraints introduced in part IV except for constraints of type
  - (14.16) (these are also left out in MIP),
  - (13.8) avoiding double usage of saw tracks,
  - (13.6) avoiding conflicts and
  - (14.4) and (14.5) linking the θ-variables that are replaced by
    - ∗ the constraints of type (17.1) and (17.2) in 17.4.1 inserting the overstep in θ-variables
    - ∗ the constraints of type (17.3) and (17.4) bounding the overstep variables
    - ∗ the constraints of type (17.6) avoiding that compositions are broken up

- the objective (17.5) as described in 17.4.1

is called the first stage of the 'two-stage solution method' and is denoted by $MIP'(APT_0, DPT_0)$.

17.5 Second Stage

In the second stage, apt's and dpt's are added to get APT_1 and DPT_1: Firstly, the apt's and dpt's that achieve the suggested oversteps are added. Secondly, to avoid conflicts, we add apt's and dpt's that deviate a small number of minutes from the apt's and dpt's that are chosen by $MIP'(APT_0, DPT_0)$ from the start sets. On page 100 this process is represented in a schematic way.

In the remainder of this subsection we describe how this can be done.

17.5.1 Adding apt's and dpt's

The oversteps suggested by the outcome of the first run are used to add apt's and dpt's. This process is not described in all details, but we give an idea how this can be done:

- The suggested overstep has to be made a bit larger, because of avoiding conflicts with the composition to which it is exchanged.
- The new apt's and dpt's have to be such that no conflicts with the timetable occur.

Thereafter it is checked for the apt's and dpt's chosen in the first stage whether they contain conflicts. If two apt's/dpt's have a conflict, for both extra apt's/dpt's with departure times a few minutes after (in case of apt's) or before (in case of dpt's) the original times are added. It would also be a good idea to look for alternative routes that do not conflict and not change the departure time.

17.5.2 Discussion

As mentioned earlier, the outcome of the first stage does not always represent a feasible shunt plan. By adding apt's and dpt's we try to realize a feasible shunt plan close to the suggested shunt plan. However, there is no guarantee that this will succeed.

- Note that there is no guarantee that the suggested apt's and dpt's can all be made. It is perhaps not possible to find a feasible route after the suggested overstep time and before the end of the shuntwindow that has no conflict with a train from the timetable. By calculating the maximal overstep value mo_{apt} in a more clever way these situations may be avoided.

- Moreover, even if all suggested apt's and dpt's are made, there is no guarantee that they are used in the second stage. Some examples follow to explain this:

  - A new apt or dpt can cause a crossing on its platform track in the case that its composition is the result of a part split. That is because in the first stage the chosen apt with overstep did not have an adapted shunt time: only θ is adapted.
  - A new apt apt^1 that realizes an overstep can have θ_{apt^1} that is larger than suggested. This can happen because of trains from the timetables. It can happen that between $\theta_{apt} + y_{apt}$ and θ_{apt^1} the park track is entered by another composition. This results in another order on the park track and makes that the suggested shunt plan is infeasible.

- Note that both a larger θ for arriving units and a smaller θ for departing units imply a shorter stay at the parktrack. This implies that we do not have to bother whether the new apt's violate length constraints in combination with the apt's chosen in the optimal solution of the first stage. This holds both for the apt's and dpt's realizing overstep as for the apt's and dpt's that are added because of avoiding conflicts. By introducing parking on platform tracks, the suggested oversteps can cause overusage of platform tracks.

17.5.3 Second MIP model

In $MIP(APT_1, DPT_1)$ the overstep variables y do not exist. The maximal absolute gap between the best possible solution and the outcome of $MIP(APT_1, DPT_1)$ therefore may be larger: the value of the minimal extra driving time caused by a saw movement.

We want to investigate what is lost by using the two stage model compared to using $MIP(APT, DPT)$, the model that is far too large by the number of possible apt's and dpt's. We can make a statement for cases in which the outcome of the second stage of the two stage solution method is smaller than the outcome of the first stage.

Let $ov(MIP'(APT_0, DPT_0))$ be the *objective value* of the first stage of the two-stage solution method subtracting the costs for overstep. The optimal value of the second stage is denoted by $ov(MIP(APT_1, DPT_1))$.

Claim. *If $ov(MIP(APT_1, DPT_1))$ is equal to $ov(MIP'(APT_0, DPT_0))$, all proposed oversteps are realized and the resulting shunt plan of the two-stage solution method does not contain more shunt moves than the optimal shunt plan of $MIP(APT, DPT)$.*

Proof. *In the first stage of the two-stage solution method all possible orders on all park tracks can be chosen by introducing overstepts: The freedom in choosing compositions and park tracks is present in APT_0 and DPT_0. The freedom in time is introduced by the possibility to introduce overstep.*

This results in the lowest possible number of shunt moves. These oversteps are only restricted by the upperbound of the shunt window. This is also the case if the whole set of possible apt's and dpt's is made. These possible orders are neither restricted by conflicts or crossings on platform tracks, because

- *the constraints avoiding conflicts are not inserted in the first model*
- *the startset is made in such a way that no crossings on platform tracks occur (see first adaption in subsection 16.2) between units of the same part. The shunt times of the apt's and dpt's are not adapted if an overstep for the apt or dpt is introduced.*

Because the first model has all possible orders to its disposal, the outcome can not be worse than the one-stage solution method where also every possible arrival time on its park track is available. □

Remark. Note that the assumption $ov(MIP(APT_1, DPT_1))$ equals $ov(MIP'(APT_0, DPT_0))$ is in general not enough to state that also $ov(MIP(APT_1, DPT_1)) = ov(MIP(APT, DPT))$. Namely, the optimal solution of $MIP(APT_1, DPT_1)$ may contain apt's/dpt's with saw movements with saw times that are influenced by trains from the timetables. By shifting the departure times of the route parts before such saw movement, perhaps shorter dwell times on saw tracks can be realized. These apt's/dpt's are present in APT, DPT, but not in APT_1, DPT_1. This makes that the optimal value of $MIP(APT, DPT)$ could be smaller than $ov(MIP(APT_1, DPT_1))$.

If the optimal value of the second stage is larger than in de first stage it could be useful to insert a third stage. We offer time for quality: all apt's and dpt's for units of the same type as units in apt's/dpt's that had an overstep in the optimal outcome of $MIP'(APT_0, DPT_0)$ are inserted brute force. If the outcome of the third stage has again a larger objective value than $ov(MIP'(APT_0, DPT_0))$ a fourth stage could consist of adding the complete set of apt's and dpt's.

17.6 Computation experiments

In this subsection we present some calculation times for the two-stage solution method. We use the case at station Enkhuizen described in subsection 16.6 with some small adaptions. Some units are added to get a larger calculation time. This makes comparing calculation times more reliable. We tried several methods to reduce the calculation time. We again used an Intel Xeon X5472, 3GHz, 3GB RAM machine and MIP-solver Cplex 12.1.

The calculation time without any method to reduce it equals 56 seconds for the first run and 42 seconds for the second run. The time for generating and adding apt's and dpt's is negligible troughout this subsection. In general the second stage takes less calculation time. This can be explained by counting the number of variables. In comparison with $MIP(APT_0, DPT_0)$, the first stage of the two-stage solution method $MIP'(APT_0, DPT_0)$ has less constraints, but more variables (namely the overstep variables).

1. Contrary to what was stated in the introduction of the apt's and dpt's, in the model as implemented and tested the dwellside of the composition of an apt/dpt was not deduced from the next train services on the platform track. Instead, for both dwell sides an apt/dpt was made. However, the platform tracks at station Enkhuizen are all dead end tracks. This implies that we know for sure that all train services leave and enter via the B-side. That makes that we only need apt's and dpt's with $ds_{apt} = A$, the dead-end side of the track. This halves the number of apt's and dpt's. The calculation times are 19 seconds for the first run and 11 seconds for the second run, what gives roughly a three times shorter calculation time.

2. The second method is to forbid that the units of departing parts are retrieved from different tracks. Only part splits for arriving units are

allowed. This is because coupling units is a process with an uncertainty in time duration. Thus one does not want to couple just before departure as train service in the timetables. This reduces the number of dpt's. Applying this rule cumulative with the first method gives calculation time of 5 seconds for the first model and 2 seconds for the second model.

3. The last method is to forbid that departing units have overstep. Applying this in combination with the second method gives calculation times 5 seconds for the first model and 1 seconds for the second one. This is not a very big difference with only applying the second method, because the number of overstep variables was already low, because of the reducing of dpt variables. Applying this rule without the second method gives calculation times 17 for the first model and 14 for the second one. This is not smaller than the 19 and 11 seconds for the run where it was not forbidden for departing units to have overstep. One reason may be that the θ-variables of a dpt are only linked to the time difference variables, and that no ordering variables for departing units exist. This makes that the overstep variables for dpt's do not play a significant role in MIP'.

In all cases the optimal value did not change.

17.7 Case Enkhuizen

We conclude with an example. This example is based on the description in subsection 16.6. The input is a little bith adapted so that the strength of the two stage solution method appears. We compare the outcome of the one stage solution method to the outcome of the two stage solution method.

17.7.1 Input changes

Consider the optimal solution of subsection 16.6 as presented at page 98. Units $[4580, 0]$ and $[4586, 0]$ are both of type LM and depart together in train service 4519. Unit $[4580, 0]$ arrives first at track 4, so that it is matched to unit $[4519, 2]$. Unit $[4586, 0]$ arrives later and is therefore more close to the A-side of track 4. It is matched to $[4519, 1]$.

We change the input in such a way that arriving unit $[4580, 0]$ has benefit from arriving later on its park track. Its type LM is changed to type LMZ. The departing unit $[4519, 1]$ is also changed into a unit of type LMZ. Note that the matchings as made in the original input can not be made anymore on track 4, because units [4580, 0] and [4519, 2] are not longer of the same type.

If units [4580, 0] and [4586, 0] arrive in the other order on track 4, unit [4580, 0] can be matched to [4519, 1]. So unit [4580, 0] wants to arrive later on its park track than [4586, 0] arrives there. This leads to the second change in the input: We change the platform track of unit [4580, 0] into track 2. At this track no other train units arrive after [4580, 0] the remaining evening, so that it can dwell there for a longer time.

17.7.2 Output first stage

We present the output of the first stage of the two stage solution method, $MIP'(APT_0, DPT_0)$:

Composition				pt	τ
60000	0			4	16h3
14556	0			5	17h56
14558	0			6	18h26
73460	0			6	20h47
4568	0			4	21h11
4576	0			5	23h11
4580	**0**			**4**	**24h11**
73497	1	73497	2	7	25h7
4586	**0**			**4**	**25h42**
4591	0			4	23h20
3317	0			6	28h39
4519	**1**	**4519**	**2**	**4**	**29h27**
4521	0			5	29h51
4523	0			4	30h20
14523	0			6	30h40
14525	1	14525	2	7	31h10
14527	0			5	31h40

The solution value equals 1736.73. There is one y_{apt} value that is nonzero: It is the *apt* with $c_{apt} = [4580, 0]$ and $pt_{apt} = 4$, in line 7. This is like predicted above. The value of y_{apt} equals 91. Note that this is exactly the difference in minutes between the shunt time of unit [4586, 0] ($25h42$) and [4580, 0] ($24h11$).

In the set APT_1 an apt apt' for composition $c_{apt'} = [4580, 0]$, park track 4 and shunt time $\tau_{apt} + y_{apt} + tir(AnaAUitDezelfdeRiRR)$=$24h11 + 91 + 3 = 25h45$ is inserted.

The value $ov(MIP'(APT_0, DPT_0))$ is calculated form the objective value subtracting the overstep cost, so we get $ov(MIP'(APT_0, DPT_0))$ =1736.73− $y_{apt} * c_{ov}$ = 1736.73 − 91 ∗ 0.03 =1734.0.

17.7.3 Output second stage

Below we present the output of the second stage $MIP(APT_1, DPT_1)$.

Composition				pt	τ
60000	0			4	16h03
14556	0			6	17h56
14558	0			7	18h26
73460	0			7	20h47
4568	0			5	21h11
4576	0			6	23h11
73497	1	73497	2	5	25h07
4586	**0**			**4**	**25h42**
4580	**0**			**4**	**25h45**
4591	0			5	23h20
3317	0			7	28h39
4519	**1**	**4519**	**2**	**4**	**29h27**
4521	0			6	29h51
4523	0			4	30h20
14523	0			6	30h40
14525	1	14525	2	5	31h10
14527	0			7	31h40

The solution value equals 1734.0. This equals $ov(MIP'(APT_0, DPT_0))$, so the suggested overstep is realized and no third stage is needed. This can be seen in the ninth line of the output: unit [4580, 0] has shunt time $25h45$ so that it arrives three minutes later than unit [4586, 0] at track 4 so that [4580, 0] can be matched to [4519, 1] and unit [4586, 0] to [4519, 2].

17.7.4 Comparison with the one stage model

If $MIP(APT_0, DPT_0)$ acts on the same input, than units [4580, 0] and [4586, 0] are both brought to track 3. This track is open at the other side than track 4 and therefore on this track the right order of the units on this track is realized after entering at the times implied by the apt's in APT_0. However, a saw movement is needed to enter this track. The costs therefore are 23 higher because of the longer driving time of the chosen apt's and dpt's.

Note that if no track like track `3`, that is open at the B-side, would be present at the shunt yard, the units $[4580, 0]$ and $[4586, 0]$ have to be brought to different tracks. As a consequence they also have to be retrieved from different tracks, so that one extra dpt is needed. The costs of $MIP(APT_0, DPT_0)$ are then 100 higher than the costs of the outcome of the two stage solution method.

17.8 Occupancy of the shunt yard

Considering the final solution of the case in last subsection and the lengths of tracks at the shunt yard in Enkhuizen, we see that the capacity of the shunt yard is able to handle more shunt units. In this subsection we try to find out how the calculation time increases if the occupancy of the shunt yard grows.

One way to enlarge the occupancy of the shunt yard is to extend the number of arriving and departing units. From already passed subsections we know that a growing number of units and thereby a growing number of possible matchings leads to a higher calculation time anyway. This makes it difficult to compare. Instead we can also enlarge the lengths of the units or reduce the lengths of some tracks. This erlarges the occupancy too.

In this subsection we firstly investigate what happens with the calculation time if the capacity of the shunt yard decreases. These tests are done with the same input of arriving and departing units. Thereafter we change these sets so that the total lenght of the units grows. Thereafter we check wether the effect on the calculation time is the same on another shunt yard.

We compare the calculation times of the two-stage solution method with applying the first method of subsection 17.6. The case of of subsection 17.7 is modified and extended to a case of 16 units, see the tables below.

Arriving units:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
60000	1	1	LB	102	16h00
14556	1	2	L	178	17h48
14558	1	2	L	178	18h18
73460	1	3	AE	102	20h44
73460	2	3	RR	52	20h44
73460	3	3	RR	52	20h44
4568	1	1	LM	79	21h08
4568	2	1	LMY	79	21h08
4576	1	1	LA	124	23h08
4580	1	2	LMY	79	24h08
73497	1	1	LB	102	25h04
73497	2	1	LB	102	25h04
4586	1	1	LMY	79	25h39
4586	2	1	LM	79	25h39
4586	3	1	LM	79	25h39
4588	1	2	LA	124	26h09

Departing units:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
4591	1	2	LM	79	23h23
3315	1	1	RR	52	28h12
3315	2	1	RR	52	28h12
14515	1	2	LA	124	28h27
3317	1	2	AE	102	28h42
14517	1	2	LM	79	28h57
14517	2	2	LMY	79	28h57
4519	1	1	LMY	79	29h30
4519	2	1	LMY	79	29h30
4519	3	1	LM	79	29h30
4521	1	2	LA	124	29h54
4523	1	1	LB	102	30h23
14523	1	2	L	178	30h43
14525	1	2	LB	102	31h13
14525	2	2	LB	102	31h13
14527	1	2	L	178	31h43

The output is a feasible shunt plan with costs 2252. The calculation time is

56.75 seconds including both stages. We can not compare this outcome with the calculation times of subsection 17.6, because some technical parameters are changed.

In the outcome, tracks `5` and `6` are not used over their whole length.

- **The lengths of tracks `5` and `6` are halved**. The original shunt plan is still feasible then. The calculation time for this run is almost the same (54.89 seconds) and the objective value does not change (2252).
- **Again the lengths of these tracks are halved**. Now the orginal shunt plan is not feasible anymore. We get a shunt plan with costs 2268 in a calculation time of 96.79 seconds. In this shunt plan one unit extra is brought to track `3`. That needs two saw movements (one for the arriving unit and one for the departing). This explains the higher objective value.
- **Track `6` is removed. Track `5` stays at one quarter of its original length.** In this shunt yard with this input still a feasible shunt plan is possible. The costs are 2288 and calculation time is 126.47.
- **Also track `5` is removed**. A feasible shunt plan probably exists, with costs at most 2413. We say 'probably' because Cplex ran out of memory after eight minutes in the first stage and had found a solution with costs 2413.

We see that the calculation time grows if the capacity of the shunt yard becomes smaller and the arriving and departing units stay the same.

We now investigate what happens if the total lenght of the units grows. Therefore we make a case with the following properties:

- The number of units stays the same
- The number of possible matchings stays the same
- The total length of the units is larger

This case is represented in the tables below.

Arriving units:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
60000	1	1	LA	124	16h00
14554	1	2	AE	102	17h18
14556	1	2	L	178	17h48
14558	1	2	L	178	18h18
73460	1	3	AE	102	20h44
73460	2	3	RR	52	20h44
73460	3	3	RR	52	20h44
4568	2	1	LMY	79	21h08
4576	1	1	LA	124	23h08
4580	1	2	L	178	24h08
73497	1	1	LB	102	25h04
73497	2	1	LB	102	25h04
4586	1	1	LMY	79	25h39
4586	2	1	LM	79	25h39
4586	3	1	LM	79	25h39
4588	1	2	LA	124	26h09

Departing units:

t_u	position	pl_u	ψ_u	l_u	τ_{t_u}
3313	0	1	AE	102	27h42
3315	1	1	RR	52	28h12
3315	2	1	RR	52	28h12
14515	1	2	LA	124	28h27
3317	1	2	AE	102	28h42
14517	1	2	LM	79	28h57
14517	2	2	LMY	79	28h57
4519	1	1	LMY	79	29h30
4519	2	1	LM	79	29h30
4521	1	2	LA	124	29h54
4523	1	1	LA	124	30h23
14523	1	2	L	178	30h43
14525	1	2	LB	102	31h13
14525	2	2	LB	102	31h13
14527	1	2	L	178	31h43
14529	1	2	L	178	32h13

On the orginal shunt yard this case has an optimal feasible shunt plan with costs 2448 after 23.58 seconds. That is smaller than the calculation time of the case at the beginning of this subsection. From that we conclude that the changes did not make this instant more difficult.

- Just as in the first case, the lengths of tracks `5` and `6` are halved. We get a feasible shunt plan with costs 2448 after 96.37 seconds.
- Again the lengths of these tracks are halved. Now the calculation time is 684.26 and the optimal value is 2486. The optimality prove of the first stage was most time consuming: The optimal solution was found in 65 seconds. If we would break the Cplex run when the best possible solution is at most one percent smaller than the current solution this calculation time would be much smaller.

We compare the results of this case with the results of the first case.

- On the shunt yard where the track lenghts of `5` and `6` was halved, the second case has substantial larger calculation time than the first case (96.37 and 54.89). Note that in the second case two extra saw movements are needed, while in the first case these are not needed.
- On the shunt yard where the track lenghts are reduced to a quarter of the orginal lenght, the calculation time was again larger for the second case (684.26 and 96.79).

Considering that on the orginal shunt yard the second case was solved faster than the first case and that the number of possible matchings stays the same, we conclude that the higher calculation time for the second case compared to the first case can be explained by the higher occupancy (total length of the units is larger) of this shunt yard.

Fictive case

We now want to check whether we obtain the same effect on another shunt yard. We take the shunt yard and the case consisting of eleven units of subsection 16.5. To make more overstep possible the instance is changed a little bit: The platform tracks of the train services arrive over and over on track `1` and track `2`.

The two stage solution method has a calculation time of 51.22 seconds and an optimal solution value of 1705.

- **The shunt yard is extended with three tracks open at the** B**-side only. These tracks can be reached from both track `1` and `2` without a saw movement.** In this case an optimal solution with costs 1530 is obtained within 7.53 seconds.

We saw that the capacity of the shunt yard grows and the case does not change. This leads to a substantial smaller calculation time.

We now extend the case so that the number of possible matchings stays the same. We have a case with 13 units. The calculation time is 29.34 seconds and the solution value is 1938.

- Tracks `4m` and `5m` are not used in the optimal solution, so we decrease the capacity of the shunt yard: **The lengths of track `2m` and `3m` are halved**. The orginal shunt plan is not feasible anymore. An optimal solution is found within 56.80 seconds and has optimal value 1997.
- **The same tracks are again halved.** This time a solution in which these tracks are not used at all is obtained in 61.43 seconds and has optimal value 1997.
- **The lengts of track `2m` and `3m` are set back to their original length. In stead we halve the length of tracks `11m`, `12m` and `13m`.** The calculation time increases to 463.96. The optimal value is 2099.

We again see that a growing occupancy of the shunt yard leads to a higher calculation time. This is the result in two different indepedent cases. However, more research is needed to be able to make a general statement.

Part VI

Conclusions and suggestions for further research

In this report we showed that the APT-model is able to realize profit by making the routing decisions integral with the parking and matching decisions.

We have seen how the routing and parking decisions restrict each other if via tracks of routes leading to park tracks can be used for parking too. Besides we saw how changing shunt times can imply a different order on a park track, what can lead to better matchings. This is realized in the two-stage solution method for the APT-model.

The freedom in shunt times can also be used by units to route them together with units from another train what also can imply different orders. This option has not been implemented yet. It would be interesting to see what the effect is of inserting this extra options in the APT-model. The same holds for the extra option to park train units on platform tracks. It could be investigated what the effect of these additions is to the calculation times and costs.

During the computational experiments some extra questions came up.

In 16.5.1 we compared the calculation times of three cases with ten units. We determined a significant decreament of the calculation time for cases in which units are of more different types. Perhaps the number of possible matchings is a better indicator for the estimation of the calulation time than the number of units. This would be interesting for further research. We also determined that the calculation time decreases if some units enter and leave using 'last-in-first-out' discipline. It would be interesting to know if this is in general the case.

In 16.5.5 we tried to reduce the calculation time of the one-stage solution method by applying LP-relaxation of some variables. For the cases that we tried, no substantial improvements in calculation time were detected. However, it may be the case that the calculation time for obtaining the first feasible solution is smaller if LP relaxation is applied. More research is needed to be able to make a general statement about this.

In subsection 17.6 we introduced some methods to reduce the calculation time of the two-stage solution method. Some methods reduced the calculating time significantly. In the case that we used, the objective value did not increase by applying these methods. However, the methods imply a reduction of possibile dpt's (no part splits and freedom in time for departing units), so the objective value may become larger. It would be interesting to know how in real life instances the objective value behaves by using these methods.

In subsection 17.8 we tested the algorithm on several cases enlarging the occupancy of the shunt yards. We saw that a growing occupancy raises the calculation time in these cases. More tests are needed to make a general statement about the relationship between occupancy of the shunt yard and the calculation time.

List of symbols

Train units

Symbol	Status	Description/Definition
U^+	Input	Set of arriving units entering the supply
U^-	Input	Set of departing units leaving the supply
U	Construction	Set of all units, $U^+ \cup U^-$
Q	Construction	Set of all possible matchings
ψ_u	Property	Type of unit u
l_u	Property	Length of unit u
t_u	Property	Train service to which unit u belongs
τ_{t_u}	Property	Planned time of train service t_u
$pt(u)$	Assignment	Park track of u
$\tau(u)$	Assignment	Shunt time of u
$\theta(u)$	Assignment	Time on park track of u
$R(u)$	Assignment	Route of u
$T(u)$	Assignment	Departure times of the successive route parts of $R(u)$

Infrastructure

Symbol	Status	Description/Definition
S_p	Input	Set of park tracks
S_{pl}	Input	Set of platform tracks
S	Input	Set of all tracks on the shunt yard
l^s	Property	Length of track s
$\mathcal{R}$	Input	Set of all non sawing routes
o_R	Property	Origin of route R
d_R	Property	Destination of route R
S_R	Property	Set of switches of route R

Time interval Restrictions

See page 27.

Timetables

Symbol	Status	Description/Definition
$\mathcal{T}$	Input	Set of all train services at a station
l_t	Input	Lenght of the composition of train service t
τ_t	Input	Planned arrival time of train service t
τ_t'	Input	Planned departure time of train service t
R_t	Input	Route on the shunt yard until arrival of train service t
R_t'	Input	Route on the shunt yard from departure of train service t
pl_t	Input	Platform track of train service t
s_t	Input	Successor of ending train service t
p_t	Input	Predecessor of starting train service t

Decision variables

Symbol	Status	Description/Definition
x_{apt}	Variable	1 if *apt* is chosen, 0 otherwise
x_{dpt}	Variable	1 if *dpt* is chosen, 0 otherwise
$L_{s,\theta}$	Variable	length of the units on track s at time θ
$B_{s,\theta}$	Variable	1 if $L_{s,\theta} > 0$, 0 otherwise
$\theta(u)$	Variable	arrival time of $u \in U^+$ at $pt(u)$
$\theta(v)$	Variable	departure time of $v \in U^-$ from $pt(u)$
$T_{u,v}$	Variable	1 if $\theta(v) > \theta(u)$, 0 if $\theta(v) < \theta(u)$
$M_{u,v}$	Variable	1 if u is matched to v, 0 otherwise
$a(u, u')$	Variable	1 if u' stands more close to the A-side than u, 0 otherwise

Apt/Dpt

Symbol	Status	Description/Definition
APT	Construction	Set of all apt's
DPT	Construction	Set of all dpt's
c_{apt}	Main	Composition of *apt*
R_{apt}	Main	Route of *apt*
T_{apt}	Main	Departure times of the route parts of *apt*
ds_{apt}	Main	Dwellside on the platform track of *apt*
pl_{apt}	Derivative	Platform track of *apt*, o_R
pt_{apt}	Derivative	Park track of *apt*, d_R
τ_{apt}	Derivative	Shunt time of *apt*, T^1
θ_{apt}	Derivative	Arrival time on park track of *apt*, $T^{last} + tir(DrivingTime)$
mo_{apt}	Derivative	Maximal overstep
y_{apt}	Variable	Overstep in arrival time on park track
$plsi_{apt}$	Derivative	The leaving side of the platform track of *apt*
si_{apt}	Derivative	The entering side of the park track of *apt*
l_{apt}	Derivative	Length of the composition of *apt*, $\sum_{u \in c_{apt}} l_u$
C_{apt}	Derivative	Cost of *apt*
c_{dpt}	Main	Composition of *dpt*
R_{dpt}	Main	Route of *dpt*
T_{dpt}	Main	Departure times of the route parts of *dpt*
ds_{dpt}	Main	Dwellside on the platform track of *dpt*
pl_{dpt}	Derivative	Platform track of *dpt*, d_R
pt_{dpt}	Derivative	Park track of *dpt*, o_R
mo_{dpt}	Derivative	Maximal overstep
y_{dpt}	Variable	Overstep in departure time from park track
τ_{dpt}	Derivative	Shunt time of *dpt*, $T^{last} + tir(DrivingTime)$
θ_{dpt}	Derivative	Time on park track of *dpt*, T^1
$plsi_{dpt}$	Derivative	The **entering** side of the platform track of *dpt*
si_{dpt}	Derivative	The **leaving** side of the park track of *dpt*
l_{dpt}	Derivative	Length of the composition of *dpt*, $\sum_{u \in c_{dpt}} l_u$
C_{dpt}	Derivative	Cost of *dpt*
APT_u	Construction	Set of apt's for unit u, $\{apt \in APT \mid u \in c_{apt}\}$
DPT_v	Construction	Set of dpt's for unit v, $\{dpt \in DPT \mid u \in c_{dpt}\}$

Literature

[Len06] Ramon Martijn Lentink. *Algorithmic decisions support for shunt planning*, PhD thesis, Erasmus Research Institute of Management (ERIM), Rotterdam, 2006. ISBN 978-90-5892-104-8, chapter 7.

[vdBro09] John Johannes Jacobus van den Broek. *MIP-based approaches for complex planning problems*, PhD thesis, Beta, Research school for operations management and logistics, Eindhoven, 2009. ISBN 978-90-386-2060-2, Chapter 4.