Optimization and synthesis of railway signalling layout from local capacity specifications

Bjørnar Luteberget
Christian Johansen
Martin Steffen

FM, 9 Oct 2019

UiO : University of Oslo

RailCOMPLETE

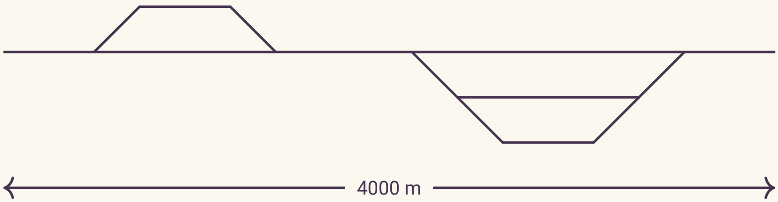
Overview

1. Railway control system design and its challenges.
2. Specifying and verifying capacity within limited scope.

3. Synthesizing control system design from scratch.
4. Optimizing control system design interactively.

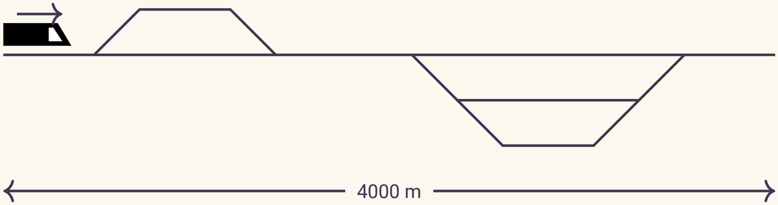
Railway control systems

Constructing a new railway line starts with a track plan:



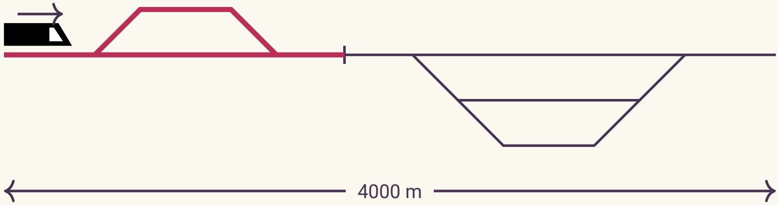
Railway control systems

Constructing a new railway line starts with a track plan:



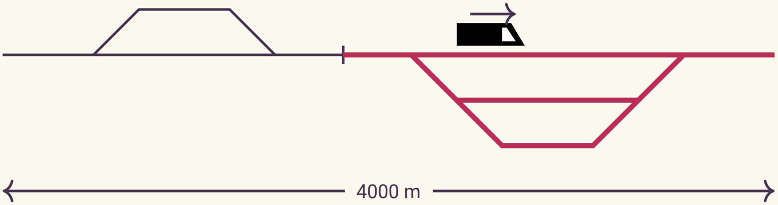
Railway control systems

By adding detectors, we can allocate smaller pieces of tracks to the train:



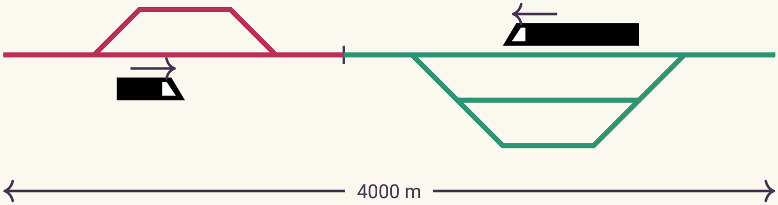
Railway control systems

By adding detectors, we can allocate smaller pieces of tracks to the train:



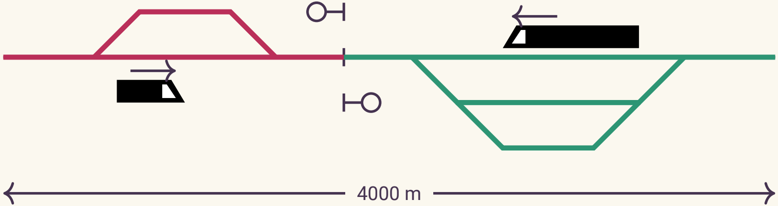
Railway control systems

Now, other trains can occupy different sections.



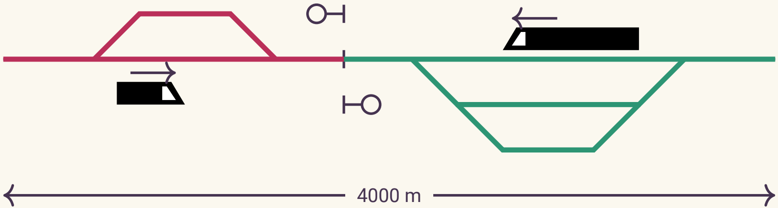
Railway control systems

We add signals to indicate to drivers when they can proceed.



Railway control systems

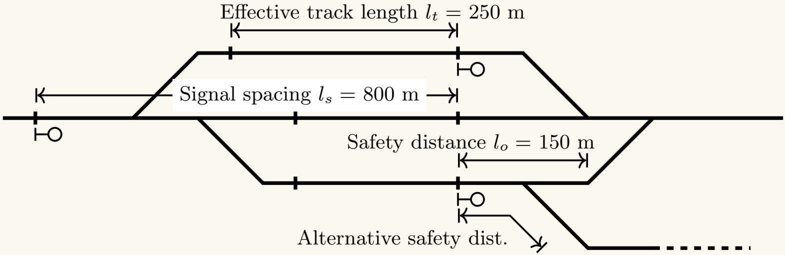
This situation is in principle safe, but is it a good design?



Two views on capacity: schematic track plan

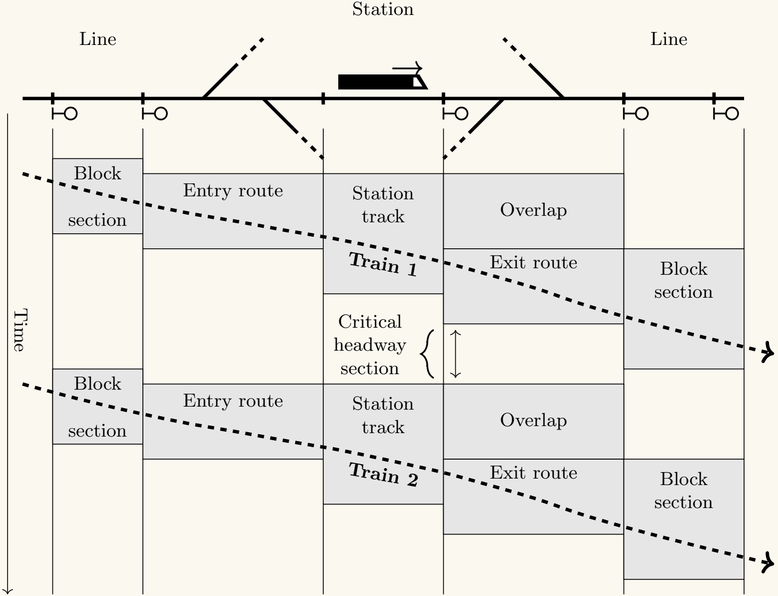
The schematic track plan is a map of tracks and components, such as signals, detectors, etc.

Distance margins determine allowable simultaneous movements.



Two views on capacity: blocking diagram

A single path, or related paths mapped to a linear axis.



Specification capture

Railway engineers gave us examples of performance properties that governed their designs.

Typical categories:

1. Running time (get from A to B)
   - Similar to a simulation test, but smaller specification.
2. Frequency (several consecutive trains)
   - Route trains into alternate tracks.
3. Overtaking
4. Crossing
   - Let one train wait on a side track while another train passes.

Capacity specifications

Local requirements suitable for construction projects.

- Operational scenario $S = (V, M, C)$:
- Vehicle types $V = \{(l_i, v_i^{\text{max}}, a_i, b_i)\}$, defined by length, max velocity, max accel, max braking.
- Movements $M = \{(v_i, \langle q_i \rangle)\}$, defined by vehicle type v and ordered sequence of visits $\langle q_i \rangle$.
  - Each visit $q_i = (\{l_i\}, t_d)$ is a set of alternative locations l_i and an optional dwelling time t_d.
- Timing constraints $C = \{(q_a, q_b, t_c)\}$ which orders two visits and sets a maximum time from the first to the second $t_{q_a} < t_{q_b} < t_{q_a} + t_c$. The maximum time constraint can be omitted ($t_c = \infty$).

Advantages of capacity specification

Can be specified for a single construction project, not dependent on whole-network timetables.

This can give us:

- Improved communication about specifications between contractual parties.
- Automated analysis
  - Early-stage, lower-effort capacity verification
  - Regression testing after changes in design
  - Unifies ad-hoc methods in use today
- Better understanding and communication between construction engineers and timetable planners.

Verification of local capacity specifications

Verification of these specifications would involve finding satisfying train trajectories and control system state:

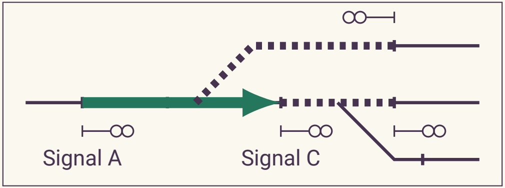
$$\exists p : \ \mathsf{spec}(p)$$

Also, constrained by:

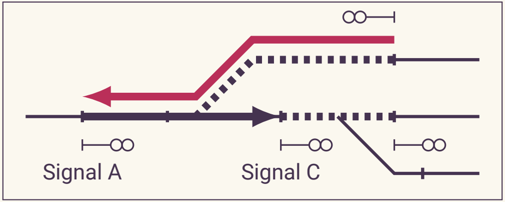
- 1 - Physical infrastructure
- 2 - Allocation of resources (collision safety)
- 3 - Limited communication
- 4 - Laws of motion

Constraints (2) Allocation of resources

An elementary route is a set of resources allocated together.

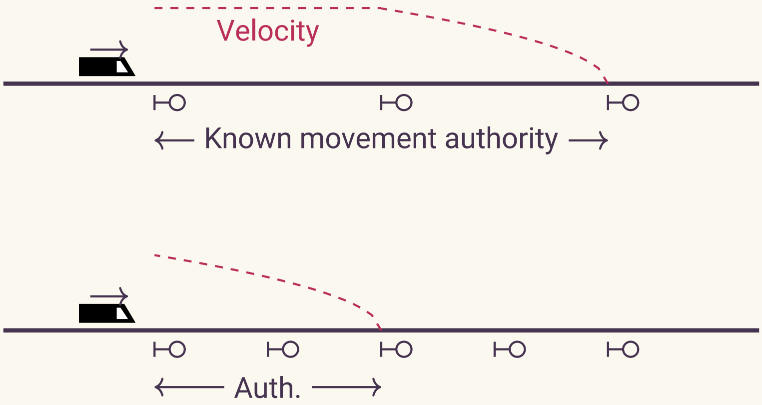


Routes are conflicting if they use any of the same resources.



Constraints (3) Limited communication

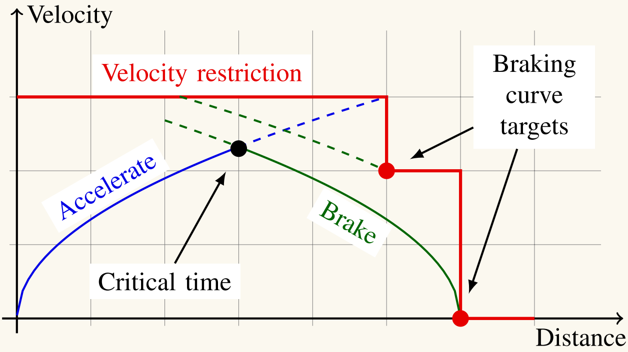
Signal information only carries across two signals ("pre-signalling").



Constraints (4) Laws of motion

Trains move within the limits of given maximum acceleration and braking power. Train drivers need to plan ahead for braking so that the train respects its given movement authority and speed restrictions at all times.

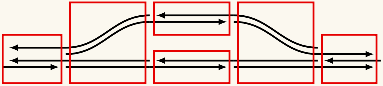
$$v - v_0 \leq a\Delta t, \qquad v^2 - v_i^2 \leq 2bs_i.$$



Dispatch vs. driver

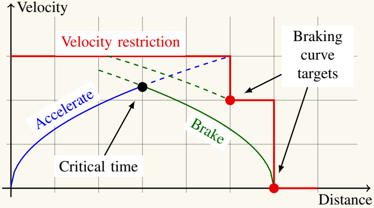
Split the planning work into two separate points of view:

Dispatcher

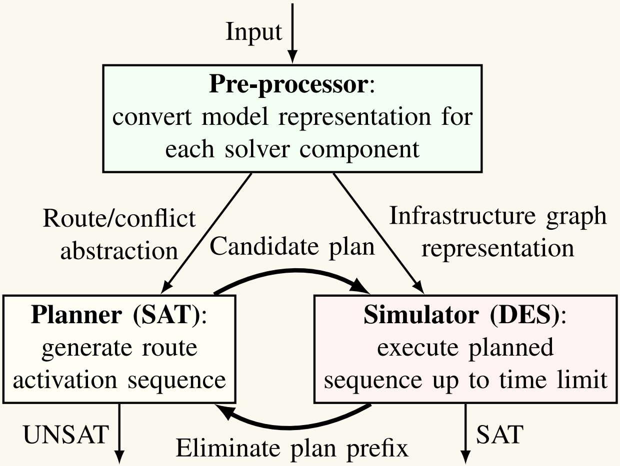


Elementary routes and their conflicts

Train driver

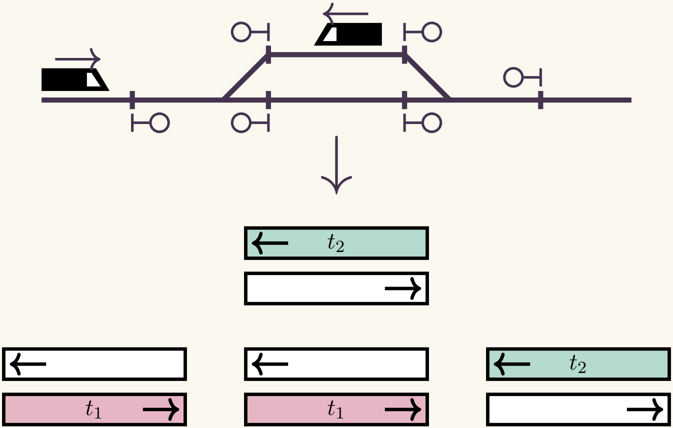


Verification architecture



SAT encoding of dispatch planning

General idea: represent which train occupies which elementary route in each of a sequence of steps.



SAT encoding

Planning as bounded model checking (BMC). Build planning steps as needed using incremental SAT solver interface.

Movement correctness:

- Conflicting routes are not active simultaneously $\text{conflict}(r_1, r_2) \Rightarrow o^i_{r_1} = \text{Free} \vee o^i_{r_2} = \text{Free}$.
- Elementary route allocation is consistent with train movement: $(o^i_r \neq t \wedge o^{i+1}_t = t) \Rightarrow \bigvee \{o^{i+1}_{r_x} = t \mid \text{route}(r_x), \text{entry}(r) = \text{exit}(r_x)\}$

Satisfy specification:

- Visits happen in order (timing requirement is measured on simulation).

From verification to synthesis

Can we use verification techniques
to synthesize signaling designs?

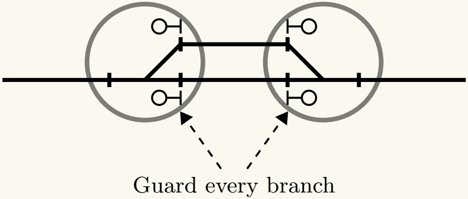
Initial design

- Adding a single component somewhere does not give any good information.

- Let's turn synthesis into optimization by over-approximating required components.

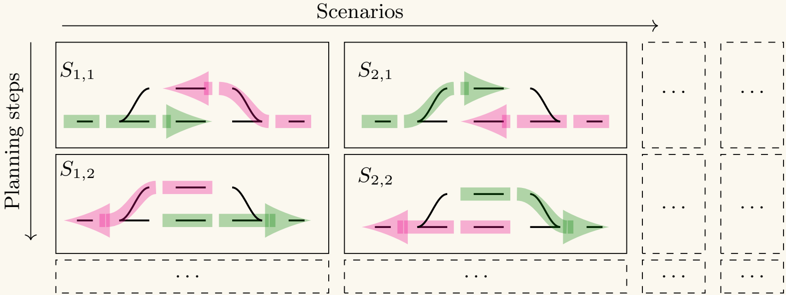
Start with an initial design:

- Include signals at fixed distances from merging paths.
- The distances correspond to choices of overlap distance.



Minimize number of signals

- Instead of verifying each property separately, on a known model ...
- ... we have unknowns in the model, and need to satisfy all properties simultaneously.



Minimize number of signals

- Then, we can add a signal used indicator boolean to the SAT problem, linking the usage of a signal across all planning steps and all scenarions.

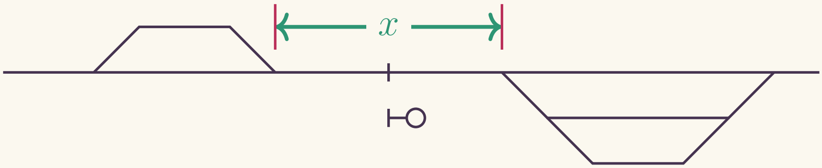
$$\forall i \in \textsf{State} : \forall s \in \textsf{Signal} : \forall t \in \textsf{Train} : \quad \neg u_s \Rightarrow$$
$$\bigvee \left\{ \left(o_r^i \neq t \wedge o_r^{i+1} = t \right) \mid \textsf{exit}(r) = s \right\} \Rightarrow$$
$$\bigvee \left\{ \left(o_r^i \neq t \wedge o_r^{i+1} = t \right) \mid \textsf{entry}(r) = s \right\} \ .$$

- Solve MaxSAT maximising unused signals.

Numerical optimization of component locations

Signal minimization gives a set of signals and a set of corresponding dispatches which fulfil the given specifications.

- Adjusting positions of components may improve timing results in simulator.
- Discontinuous, non-linear, multivariate real-valued optimization problem.



The function to be optimized

The function to be optimized is a weighted sum of dispatch timing measures.

$$f_b(\vec{x}) = \sum_s w_s \left(\frac{1}{n_s} \sum_d t_{b+\vec{x}}(d) \right),$$

where

- $\vec{x}$ represents the location of each signal and detector,
- s indexes capacity specifications,
- w_s is the weight assigned to specification s,
- d indexes dispatch plans for each operational scenario, and
- $t_{b+\vec{x}}(d)$ is the simulation timing result.

(Trading performance and cost is performed by the user)

Powell's method

We fix the set of components, fix the tracks that they belong to, and fix their order within the track.

Powell's method (1964):

- Given domain $D \subset \mathbb{R}^n$, initial point $\vec{x}_0 \in D$, and cost function $f : D \to \mathbb{R}$.
- Iterate through search vectors $\vec{v}_i \in V$ and do a line search for $\alpha \in \mathbb{R}$ minimizing $\vec{x}_{i+1} = f(\vec{x}_i + \alpha\vec{v}_i)$.
- Remove the $\vec{v}_i$ which yielded the highest $|\alpha|$, and replace it with $\vec{x}_{i+1} - \vec{x}_i$ normalized. Repeat until $\|\vec{x}_{i+1} - \vec{x}_i\| < \epsilon$.

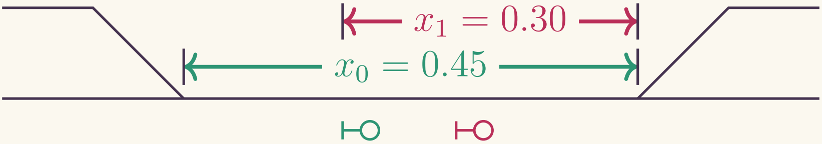
Brent's method (1973):

- A reliable method for root-finding or minimization for non-differentiable functions.
- For well-behaved functions: inverse quadratic interpolation, or linear interpolation.
- For not-so-well-behaved functions: bisection / golden section.

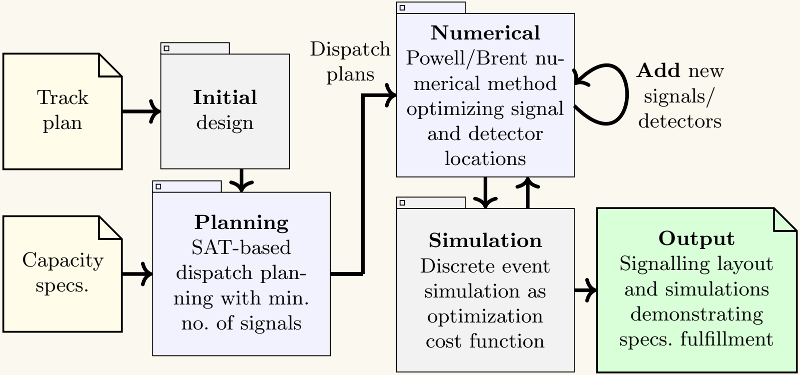
Mapping locations to the unit cube

- Preserve which tracks components are located at, and their order to ensure planned dispatches are still meaningful. Minimum distance d between components.
- Map the component location space to the unit cube $[0,1]^n$ (n-tuples in $[0,1]$) so that the whole of the unit cube is a valid point in the component location space.

Encode: $\text{scan}(0.0, \lambda\, s, x \to \text{linstep}(\text{replace}(s, x) + d, l - d, x))$.
Decode: $\text{scan}(0.0, \lambda\, s, x \to \text{replace}(s, \text{lerp}(s + d, l - d, x)))$.



Synthesis algorithm overview



Local optimization steps

- Synthesis from scratch not always suitable.
- Instead, search for a single step of the synthesis algorithm that gives the most effect on the current design.

1. **Redundant component**: removing a single object while still satisfying specifications.
2. **Local move of component**: moving a single object or a set of nearby objects may improve the overall capacity measure.
3. **Adding component**: adding a single component (and performing local moves) which improves overall capacity measure.

Each of these can be suggested to the user.

Related work

- Formal methods is all about safe implementations of control systems.
- Operations research is all about time tabling on large-scale networks.

- Mao, B. et al.: *Signalling layout for fixed-block railway lines with real-coded genetic algorithms*, Hong Kong Institute of Engineers, Transactions (2006).
- Weits, E. et al.: *Generating optimal signal positions*, Computers in Railways XII (2010).
  - Does not deal with schedulability.
  - Analytical performance models.
- Dillmann, S. and Hähnle, R.: *Automated planning of ETCS tracks*, RSSRAIL 2019.
  - Heuristic algorithm.

Conclusions and future work

- Not a complete method:
  1. initial design does may not have maximum schedulability
  2. simultaneous planning may not be the best starting points.
  3. the cost function may have multiple local optima.
- Scalability concerns:
  1. specification language unsuited for large terminals.
  2. algorithm for adding new signals is naive.

- Assumes fixed block design principles. ERTMS Level 3 with moving block may require different planning algorithm.
- Imperative simulation at the core allows extending timing calculations to be more sophisicated.
- Fast results for small infrastructures.