SINTEF

The DISPLIB 2025 train dispatching competition

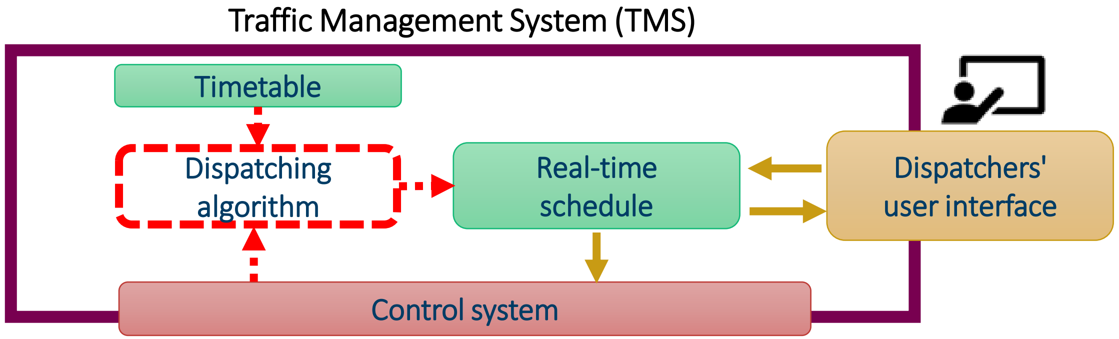
2024-09-10

Giorgio Sartor, Oddvar Kloster,
Bjørnar Luteberget and Carlo Mannino

Teknologi for et bedre samfunn

SINTEF

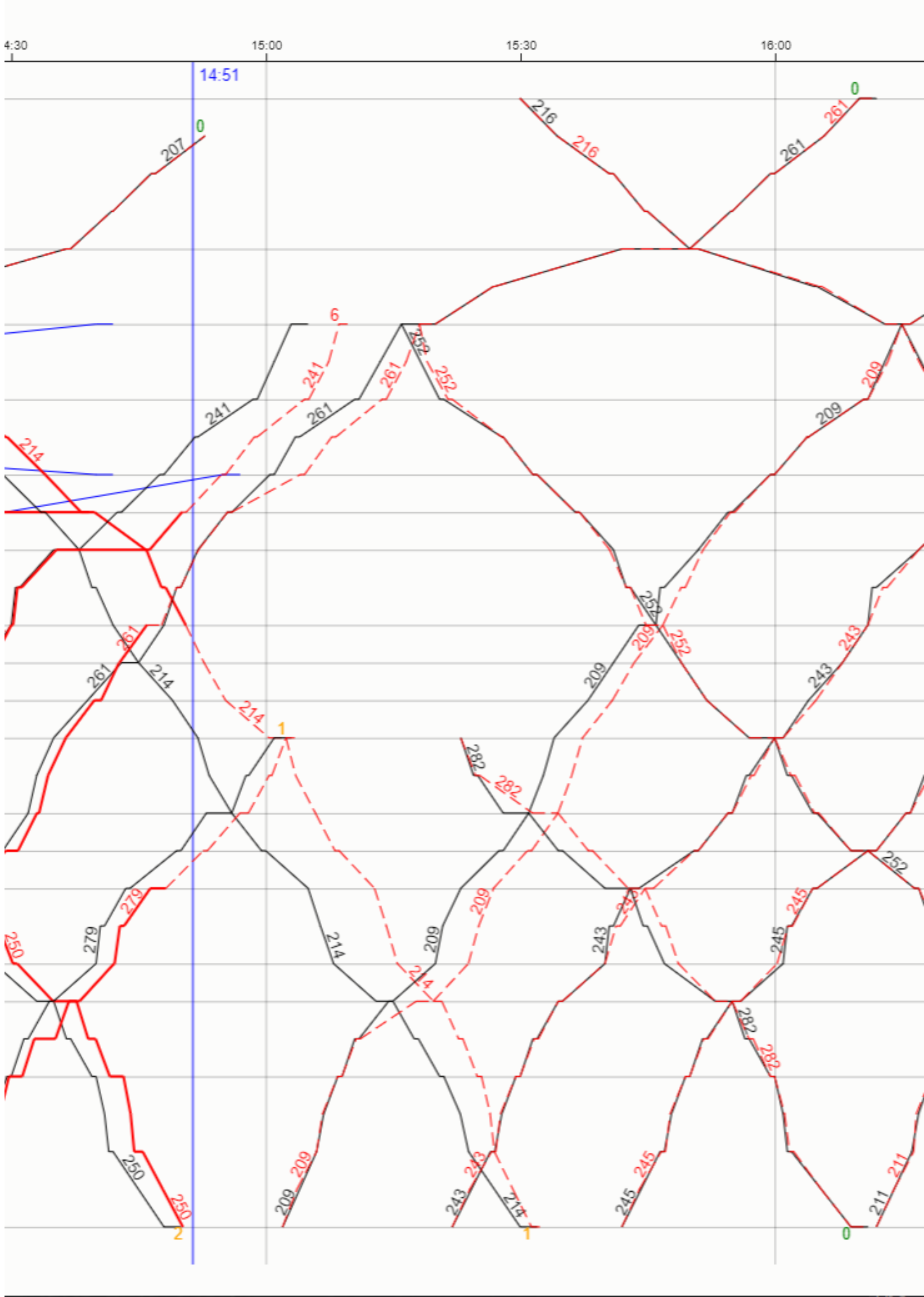
Train dispatching



- **Optimized real-time train dispatching** can significantly increase throughput and reduce delays.

- Yet, it is still the missing piece of many TMS systems

- That is because it requires state-of-the-art optimization methods, advanced software engineering skills, and a tremendous amount of domain-specific expertise
and access to data.



SINTEF

Motivation

- «train dispatching» in Google Scholar returns 17600 results from 2020…
- …and probably each one of them uses a different set of instances!

WHY?? 😭 (crying of frustration)

- Many countries still consider the sharing of railways data as a violation of national security
  - But they publish a public timetable (not in machine-readable format) 🤦
  - And (in Europe) they are required to publish a network statement
- Lack of a standard format
- Existing formats (e.g., RailML) are way too detailed and complex for non-experts

Motivation

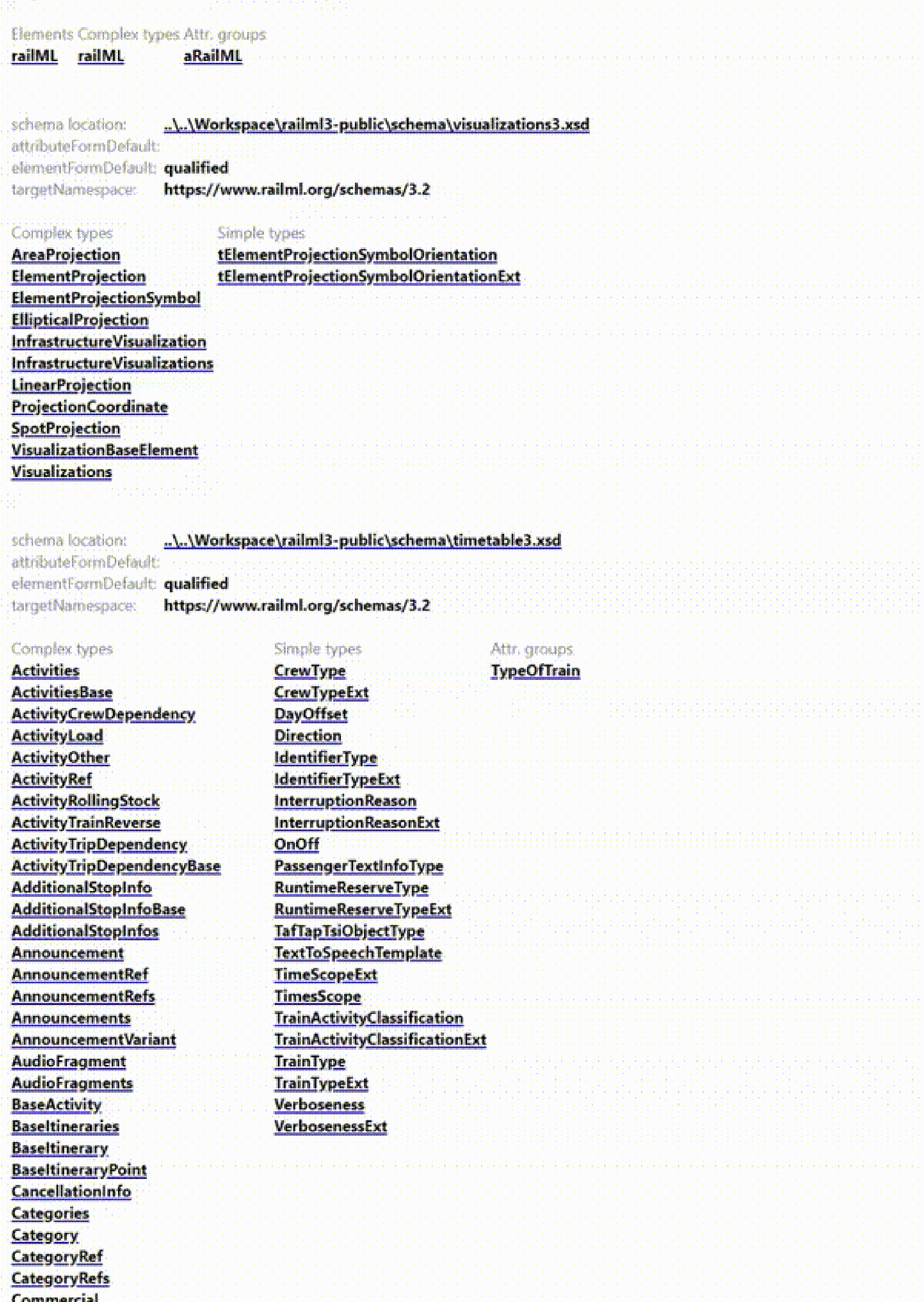
- Other research communities have gained a lot from standardized, comprehensive benchmark libraries:
  - Vehicle Routing (TSP, CVRP, …)
  - SATLIB, MIPLIB, MaxSAT Evalutations, …
  - Optimal power flow (IEEE instances)
  - Scheduling (RCPSP)
  - …

- Examples from the train dispatching community:
  - RAS Problem Solving Competitions (specialized problems, one-off)
  - Flatland (simplified problems)
  - SBB Train Schedule Optimisation Challenge (complex problems)
  - …

SINTEF

Overwhelming data 🤯

- Real-life railway data can be overwhelming:
  - Exact signal positions (a single meter can make a difference)
  - “Weird” train configurations and compositions
  - “Weird” train schedules (e.g., trains going backwards)
  - “You can stop here, but only if the train weighs less than X tons and not in winter”
  - “If a train is more than X meters long and stops in platform Y, then it blocks the pedestrian crossing, which delays connections by 2 minutes”
  - “A train that is more than X meters long occupies station Y and Z at the same time” (quantum trains? 🤔)
  - “If the train is driven by driver X, then it will make this crossing, but driver Y will not be confident doing that”
  - …the list goes on…


Elements Complex types Attr. groups
railML railML aRailML

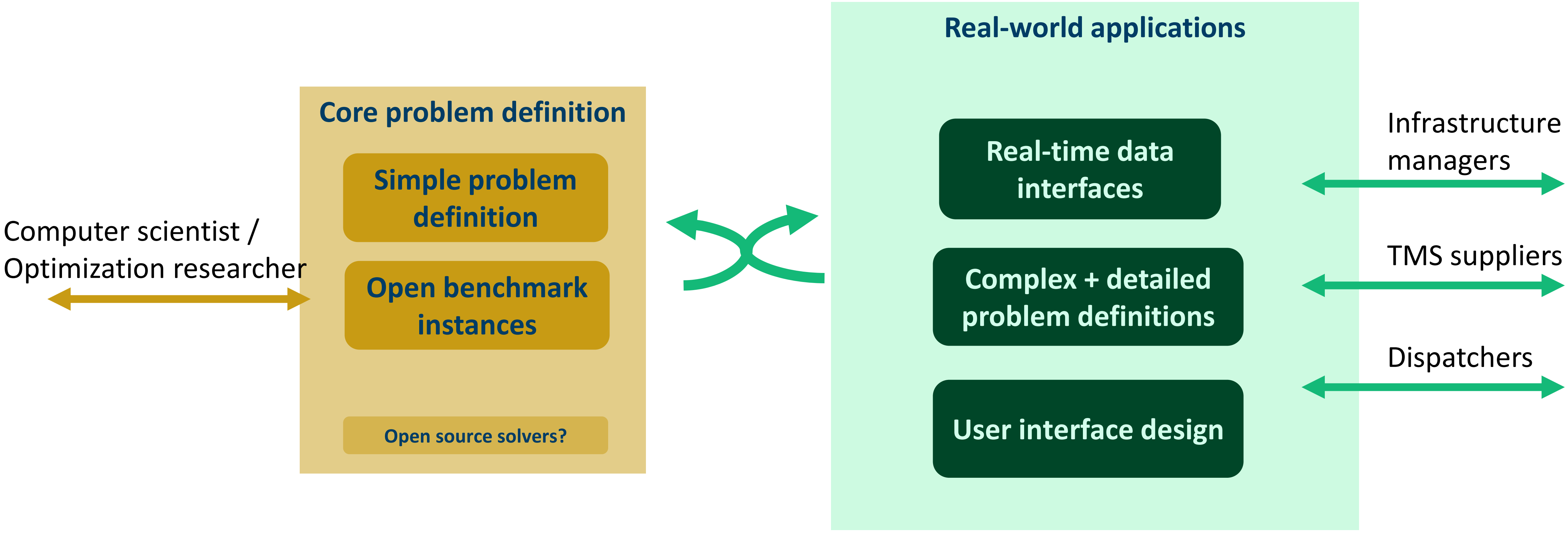
schema location: ..\..\Workspace\railml3-public\schema\visualizations3.xsd
attributeFormDefault:
elementFormDefault: qualified
targetNamespace: https://www.railml.org/schemas/3.2

Complex types	Simple types
AreaProjection	tElementProjectionSymbolOrientation
ElementProjection	tElementProjectionSymbolOrientationExt
ElementProjectionSymbol	
EllipticalProjection	
InfrastructureVisualization	
InfrastructureVisualizations	
LinearProjection	
ProjectionCoordinate	
SpotProjection	
VisualizationBaseElement	
Visualizations	

schema location: ..\..\Workspace\railml3-public\schema\timetable3.xsd
attributeFormDefault:
elementFormDefault: qualified
targetNamespace: https://www.railml.org/schemas/3.2

Complex types	Simple types	Attr. groups
Activities	CrewType	TypeOfTrain
ActivitiesBase	CrewTypeExt	
ActivityCrewDependency	DayOffset	
ActivityLoad	Direction	
ActivityOther	IdentifierType	
ActivityRef	IdentifierTypeExt	
ActivityRollingStock	InterruptionReason	
ActivityTrainReverse	InterruptionReasonExt	
ActivityTripDependency	OnOff	
ActivityTripDependencyBase	PassengerTextInfoType	
AdditionalStopInfo	RuntimeReserveType	
AdditionalStopInfoBase	RuntimeReserveTypeExt	
AdditionalStopInfos	TafTapTsiObjectType	
Announcement	TextToSpeechTemplate	
AnnouncementRef	TimeScopeExt	
AnnouncementRefs	TimesScope	
Announcements	TrainActivityClassification	
AnnouncementVariant	TrainActivityClassificationExt	
AudioFragment	TrainType	
AudioFragments	TrainTypeExt	
BaseActivity	Verboseness	
BaseItineraries	VerbosenessExt	
BaseItinerary		
BaseItineraryPoint		
CancellationInfo		
Categories		
Category		
CategoryRef		
CategoryRefs		
Commercial		

Get to the core!



Two steps towards standardized benchmarks

1. **A suitable problem definition:**
   agreeing on a «core» problem
2. **A competition:**
   creates opportunity get community involvement (instances, solvers, etc.)

Format trade-offs

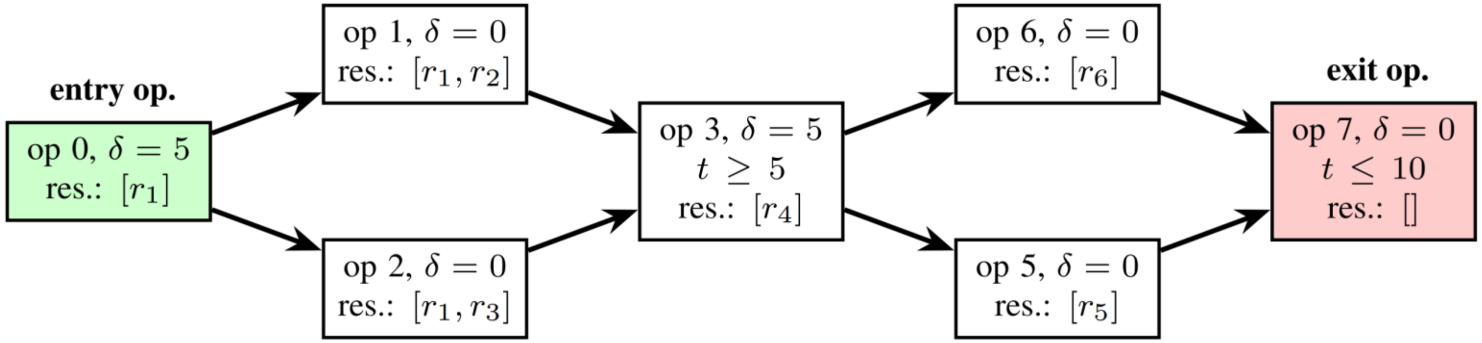
- We need a format that:
  - Simplifies the problem definition, reducing the necessary domain-specific knowledge
  - It is easy to read and interpret
  - Precisely captures all the important aspects of real-life train dispatching

  - Supports instances from all over the world
  - Allows “modelling” of more complex features without significant changes in the algorithm
  - Can be extended to support more complicated details, if necessary

```
{ "trains": [[{ ... operation ... }, ... ], ... ],
  "objective": [{ ... component ... }, ... ] }
```

A train

- Each train is represented as Directed Acyclic Graph (DAG) of train operations
- Exactly one node without incoming edges: Entry Operation
- Exactly one node without outgoing edges: Exit Operation
- A path on this graph represents a possible route (i.e., sequence of resource occupations) for the train



```
{ "start_lb": 7, "start_ub": 8, "min_duration": 5,
  "resources": [{ "resource": "r1" }], "successors": [2] }
```

A train operation

- A train operation is a *train event* to which we need to assign a **start time**
- The **end time** of an operation is equal to the start time of the next operation along the chosen route for that train (except for the Exit Operation)
- It contains the following keys:
  - **`start_lb`**: the earliest start time
  - **`start_ub`**: the latest start time
  - **`min_duration`**: minimum time between the start time and end time
  - **`resources`**: a list of resources used by the train while performing the operation
    - **`resource`**: the id of the resource
    - **`release_time`**: the minimum duration between the end time of this operation and the start time of any subsequent operation (of a different train) using the same resource
  - **`successors`**: a list of alternative successor operations

The objective

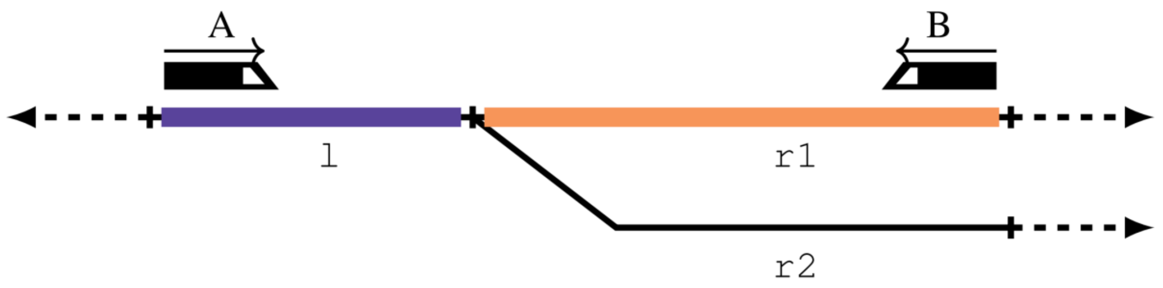
- The objective defines the cost of a solution, which should be minimized.
- The objective consists of a list of objective components, and each component describes an **operation delay**. The total cost is the sum of the operation delays.
- Each component contains the following keys:
  - **`train`**: a reference to a train
  - **`operation`**: a reference to an operation
  - **`threshold`**: a time after which this delay component is activated ($\bar{t}$)
  - **`increment`**: a one time penalty for passing the threshold (d)
  - **`coeff`**: the increase of delay for every time unit above the threshold (c)

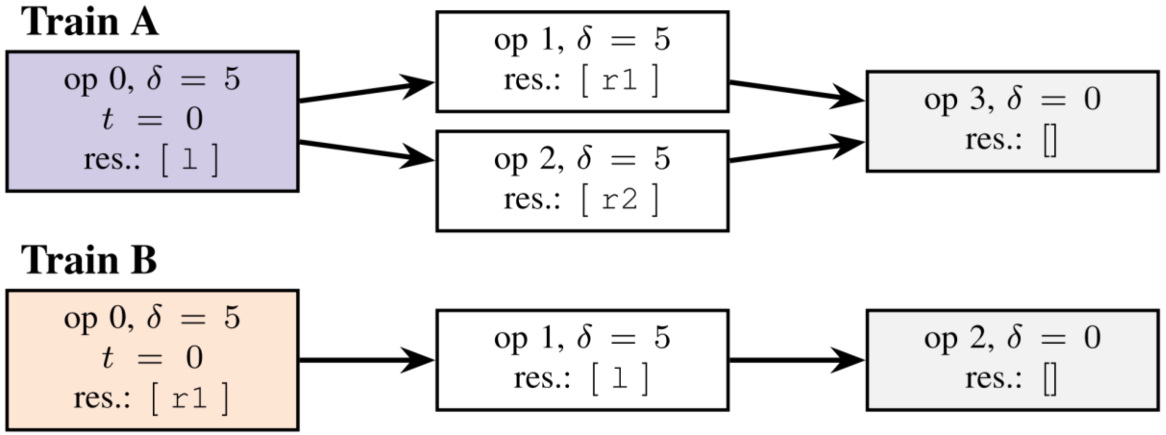
$$v_i = c \cdot \max\{0, t - \bar{t}\} + d \cdot H(t - \bar{t})$$

Heaviside step function

An example

```
{
  "trains": [
    [{ "start_ub": 0,
       "min_duration": 5,
       "resources": [{ "resource": "l" }],
       "successors": [1, 2] },
     { "min_duration": 5,
       "successors": [3],
       "resources": [{ "resource": "r1" }]},
     { "min_duration": 5,
       "successors": [3],
       "resources": [{ "resource": "r2" }]},
     { "min_duration": 5,
       "successors": []}],
    [{ "start_ub": 0,
       "min_duration": 5,
       "resources": [{ "resource": "r1" }],
       "successors": [1]},
     { "min_duration": 5,
       "resources": [{ "resource": "l" }],
       "successors": [2]},
     { "min_duration": 5,
       "successors": []}]],
  "objective": [{ "type": "op_delay",
                  "train": 1,
                  "operation": 2,
                  "coeff": 1}]
}
```





The solution

- A **solution** to the problem consists of an ordered sequence of operation start events across all trains
- The solution is feasible if:
  - For any pair of operations o_1, o_2 that belong to different trains, use a common resource, and $o_2 \succ o_1$:
    - The end event for o_1 precedes the start event of o_2 in the global sequence
    - The duration from the end event of o_1 to the start event of o_2 is greater than or equal to the release time of resource r in o_1.
- A solution contains the following keys:
  - **`objective_value`**: the cost of the solution
  - **`events`**: an ordered list of operation start events
    - **`train`**: a reference to the relevant train
    - **`operation`**: a reference to the relevant operation
    - **`time`**: the time at which the operation starts


```
{ "objective_value": 10, "events": [
  {"time": 0, "train": 0, "operation": 0},
  {"time": 0, "train": 1, "operation": 0},
  {"time": 5, "train": 0, "operation": 2},
  {"time": 5, "train": 1, "operation": 1},
  {"time": 10, "train": 1, "operation": 2},
  {"time": 10, "train": 0, "operation": 3}] }
```


Global ordering is important!

SINTEF

Additional constraints and objectives?

- It is not possible to capture every real-world problem in one formal definition.
- The DISPLIB problem formulation may seem simplistic...
- ... but «re-formulations» can cover a lot of real-world constraints and objectives:
  - Train lengths
  - Partial release (interlocking constraints)
  - Temporary capacity restrictions
  - Passenger exchanges / rolling stock correspondences, mandatory or optional with cost
  - Partial cancellations, short-turning, etc. with cost
- Overall tradeoff: leaning towards simplicity, but covering as much as reasonable

Two steps towards standardized benchmarks

1. **A suitable problem definition:**
   agreeing on a «core» problem
2. **A competition:**
   creates opportunity get community involvement (instances, solvers, etc.)

SINTEF

The DISPLIB 2025 Competition: spirit and rules

- The competition challenges the research community to find innovative and effective algorithms for solving a diverse set of real-life train dispatching instances
- The instances come from different countries and have different characteristics: some have many routing options and few trains while others have few routing options and many trains.
  - (thanks to SINTEF Digital, Siemens Mobility, data.sbb.ch for confirmed sets of instances so far…)
  - (three new data sources under way, pending data release, **more are welcome!**)

- General rules:
  - The usage of commercial MIP solver is allowed
  - The usage of ML pre-training is allowed, and the learning phase does not count against the time limit
  - The time limit to solve each instance is 10 minutes, maximum 8 CPUs and 16GB of RAM.
  - The source code does not need to be submitted, but the winners may be required to show additional proof of compliance to the rules above

The DISPLIB 2025 Competition: timeline

- **Phase 1**: consists of simpler set of instances. Solutions to the phase 1 instances may be submitted early to be part of the Phase 1 scoreboard. This is strongly encouraged!
- **Phase 2**: consists of similar but more challenging instances.

Timeline

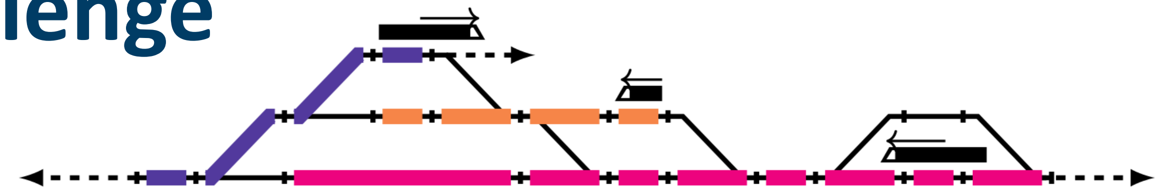
- ✓ **2024-09-10**: Competition announced at ODS conference 2024 and online channels.
- 2024-10-01: First set of problem instances will be published (Phase 1).
- 2025-01-31: Early submission deadline for the Phase 1 scoreboard (optional).
- 2025-02-03: Second set of problem instances will be published (Phase 2).
- 2025-04-30: Final submission deadline for all instances (Phase 1 and Phase 2).

The DISPLIB 2025 Competition: scoring

- The points will be distributed based on a system like Formula One's system
- The competitors are required to submit a short report (max 6 pages) describing the algorithm used
- The winners will be chosen based on a combination of the score, the novelty of the algorithms used, and the quality of the provided report
- A verification program (in Python) is available to check solutions' feasibility

Position	Phase 1	Phase 2
1st	10	15
2nd	7	11
3rd	5	8
4th	3	6
5th	2	4
6th	1	3
7th		2
8th		1

The DISPLIB 2025 Competition: a train dispatching challenge



- DISPLIB: a new train dispatching benchmark library
  - Wide range of real-life instances from all over the world
  - Simple but powerful problem definition

- The DISPLIB 2025 Competition
  - Schedule and route trains from a wide range of real-world use cases
  - No deep knowledge of railways needed to start
  - Winners will be invited to a special session at ODS 2025
  - … and get an expedited review process in the Journal of Rail Transport Planning & Management (JRTPM)
  - Competition details are available now, **instances come 1st October 2024!**

displib.github.io

SINTEF

Teknologi for et bedre samfunn