UPPSALA UNIVERSITET

Validation and generation of geographical data using a domain theory

Lars-Henrik Eriksson
Uppsala University

lhe@it.uu.se
http://user.it.uu.se/~lhe

UPPSALA UNIVERSITET

An industrial experience

- This talk will illustrate how domain theories are used in Industrilogik's work in formal specification and verification of railway systems.
- Specification, simulation, modelling and verification (refinement proof) is supported by the GTO toolset.
- The GTO language is predicate logic with finite domains and limited temporalities (previous-moment operator). Independent SAT solvers are used for proving.

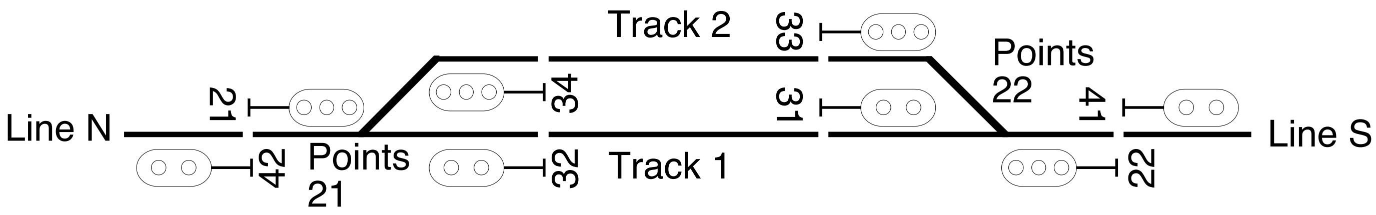
UPPSALA UNIVERSITET

Geographical data

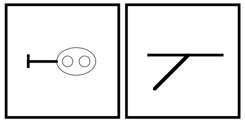
- Railway interlockings are systems governed by general rules. On an abstract level different installations perform the same functions.
- Formal req's specs can be written in a generic manner and instantiated using *geographical data*. Example: Industrilogik's specifications of Swedish and Norwegian signalling.
- Interlocking software can be written as generic modules which are combined and parameterised using geographical data. Examples: Ebilock (Bombardier), Alister (Vossloh/Banverket)

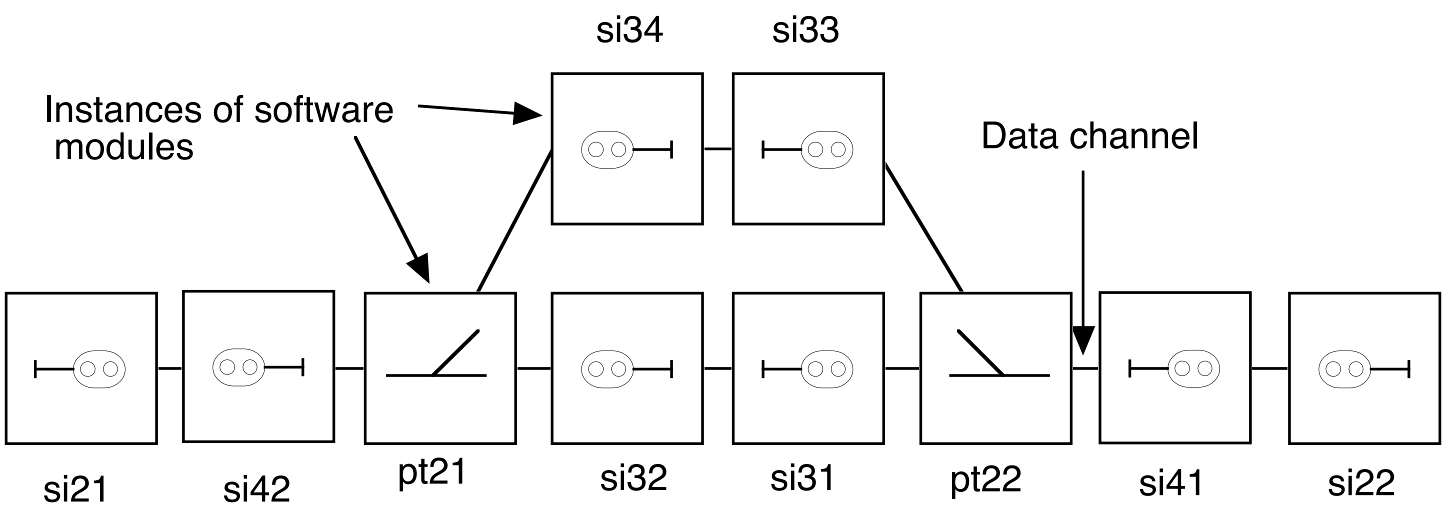
Ebilock software structure

(Bombardier transportation)



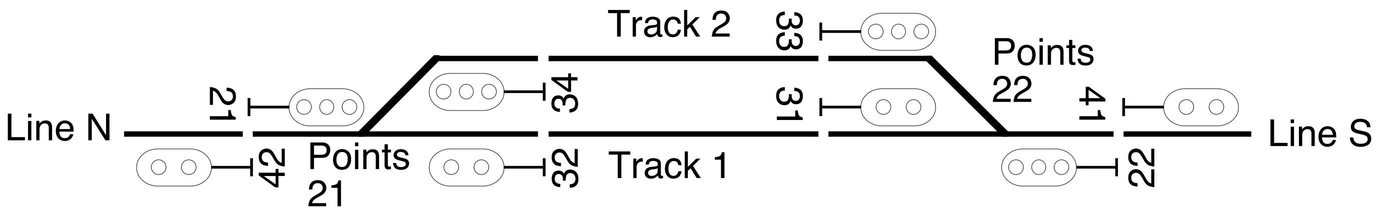
- Software modules:





Primary geographical data

- Primary geographical data relates directly to concrete properties of the installation.



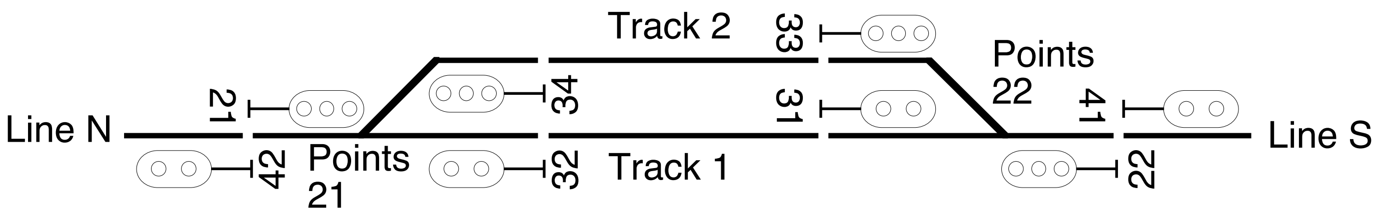
- Units: `ln, pt21, pt22, t1, t2, pt22, ls`
- Signals: `si21,si22,si31,si32,si33,si34,si41,si42`
- Relational data:

```
connectsTo(pt21,ln),    connectsTo(ln, pt21),
connectsTo(pt21,t1),    connectsTo(t1,pt21),
connectsTo(pt21,t2),    connectsTo(t2,pt21),
leftBranch(pt21,t2),    rightBranch(pt21,t1),…

ahead(si21,SL1n),      inRear(si21,pt21),…
```

Secondary geographical data

- Secondary geographical data describe abstract properties wholly or in part determined by primary geographical data.



- Routes: `rt2131, rt2133, rt2232, rt2234, rt3141, rt3242, rt3341, rt3442`

- Relational data: `conflict(rt2131,rt3242),……`

```
entry(si21,rt2131), exit(si31,rt2131),
before(rt2131,ln), first(rt2131,pt21),
partOf(rt2131,pt21), partOf(rt2131,t1),……
```

(Routes are properly sets of units, but the GTO tool is first order...)

Formal req's specification (part)

- $\forall pt \in UNIT\ (\exists rt \in ROUTE\ (routelocked(rt) \wedge partOf(pt,rt)) \wedge points(pt) \rightarrow pointslocked(pt))$
- $\forall rt \in ROUTE\ (ready(rt) \rightarrow \neg \exists rt1 \in ROUTE\ (conflict(rt) \wedge routelocked(rt)))$
- $\forall si \in SIGNAL\ (proceed(si) \rightarrow \exists rt \in ROUTE\ (entry(si,rt) \wedge ready(rt)))$
- `routelocked(rt)`: `rt` is a locked route
- `pointslocked(pt)`: `pt` are locked for reversals
- `ready(rt)`: Trains may enter `rt`
- `proceed(si)`: `si` displays a proceed aspect

UPPSALA UNIVERSITET

Validation of geographical data

- Correct function of specification/interlocking depends on correctness of geographical data.
- Primary geographical data can be checked for internal consistency.
- Secondary geographical data can be verified against the primary geographical data.
- Verification uses a domain theory for railways.
- The geo. data determines a finite interpretation of the logical language so verification amounts to computing truth values.

Domain theory (part)

- $\forall u1,u2 \in UNITS(connectTo(u1,u2) \rightarrow connectsTo(u2,u1))$
- $\forall u \in UNITS\ \neg connectTo(u,u)$
- $\forall w,u \in UNITS(points(w) \wedge rightBranch(u,w) \rightarrow connectsTo(u,w))$
- $\forall w \in UNITS\ (points(w) \rightarrow \exists u1,u2,u3 \in UNITS\ (connectsTo(w,u1) \wedge connectsTo(w,u2) \wedge connectsTo(w,u3) \wedge u1 \neq u2 \wedge u1 \neq u3 \wedge u2 \neq u3 \wedge \forall u4 \in UNITS\ (connectsTo(w,u4) \rightarrow u1=u4 \vee u2=u4 \vee u3=u4)))$
- $\forall s \in SIGNALS\ \exists u \in UNITS\ (ahead(s,u) \wedge \forall u1 \in UNITS\ (ahead(s,u1) \rightarrow u=u1))$

UPPSALA UNIVERSITET

Consistency validation

- Suppose `connectsTo(pt21,t1)` is included in the data but `connectsTo(t1,pt21)` is not.
- Sample GTO session:

```
Welcome to the GTO Formal Tool 0.5.26
> load sampleGeodata1
> listinv domain_1
ALL u1:UNIT ALL u2:UNIT
(connectsTo(u1,u2)->connectsTo(u2,u1))
> evf domain_1
FALSE
> why
Formula is FALSE because
  ~connectsTo(t1,pt21)
```

UPPSALA UNIVERSITET

Domain theory (more)

- $\forall r \in ROUTES\ \exists u \in UNITS\ (before(r,u) \land \forall u1 \in UNITS\ (before(r,u1) \rightarrow u=u1))$
- $\forall r \in ROUTES\ \forall u \in UNITS\ (before(r,u) \rightarrow \neg partOf(r,u) \land \exists u1 \in UNITS\ (partOf(r,u1) \land connectsTo(u,u1)))$
- $first(r,u) \equiv partOf(r,u) \land \forall u1 \in UNITS\ (before(r,u1) \rightarrow connectsTo(u,u1)))$
- $\forall r \in ROUTES\ \exists s \in SIGNALS\ (\forall u \in UNITS(ahead(s,u) \rightarrow before(r,u)) \land \forall u \in UNITS(inRear(s,u) \rightarrow first(r,u)))$
- $\forall r1,r2 \in ROUTES\ (conflict(r1,r2) \leftrightarrow r1 \neq r2 \land \exists u \in UNITS\ (partOf(u,r1) \land partOf(u,r2))$

UPPSALA UNIVERSITET

Generating secondary geo. data

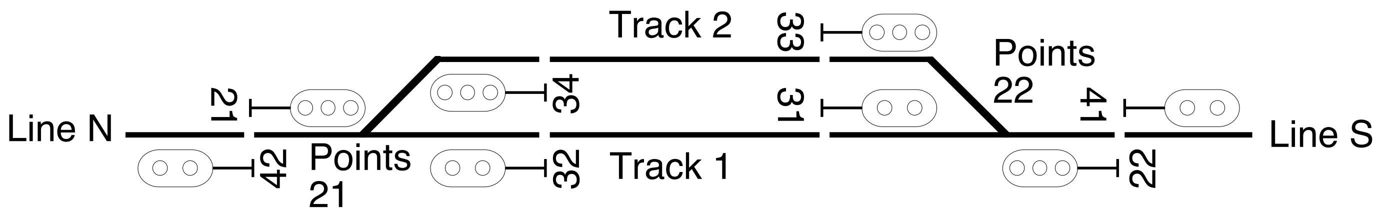
- Some axioms can be made into definitions and computed

```
conflict(r1,r2) ≡
  r1≠r2 ∧ ∃u∈UNITS
          (partOf(u,r1)∧partOf(u,r2))
```

- A SAT solver can be used to find assignments to predicates (given assignment to primary preds).
- In the particular case of train routes where the number of routes is not known in advance, a single route identifier can be used and successive SAT solutions will generate all routes.

UPPSALA UNIVERSITET

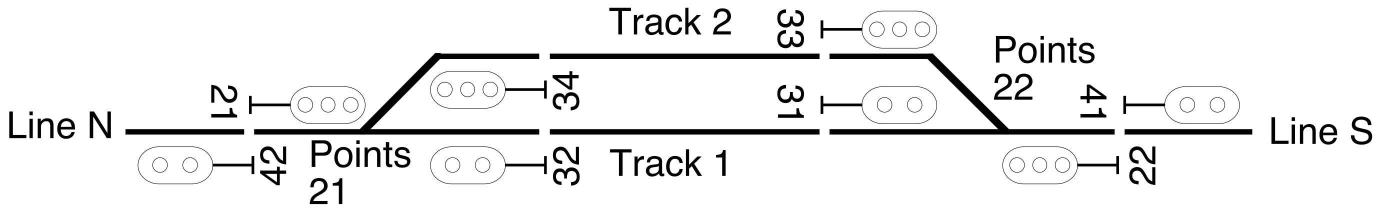
Data generation




```
Welcome to the GTO Formal Tool 0.5.26
> load sampleGeodata2
> listdef conflict
conflict(r1,r2) == r1<>r2&SOME u:UNIT
(partOf(u,r1)&partOf(u,r2))
> list conflict
conflict(rt2131,rt2133)
conflict(rt2131,rt2232)
conflict(rt2131,rt2234)
conflict(rt2131,rt3242)
conflict(rt2131,rt3442)
conflict(rt2232,rt2234)……
```

UPPSALA UNIVERSITET

Train route generation



- A single train route identifier (`route`) is used. No data for secondary predicates.

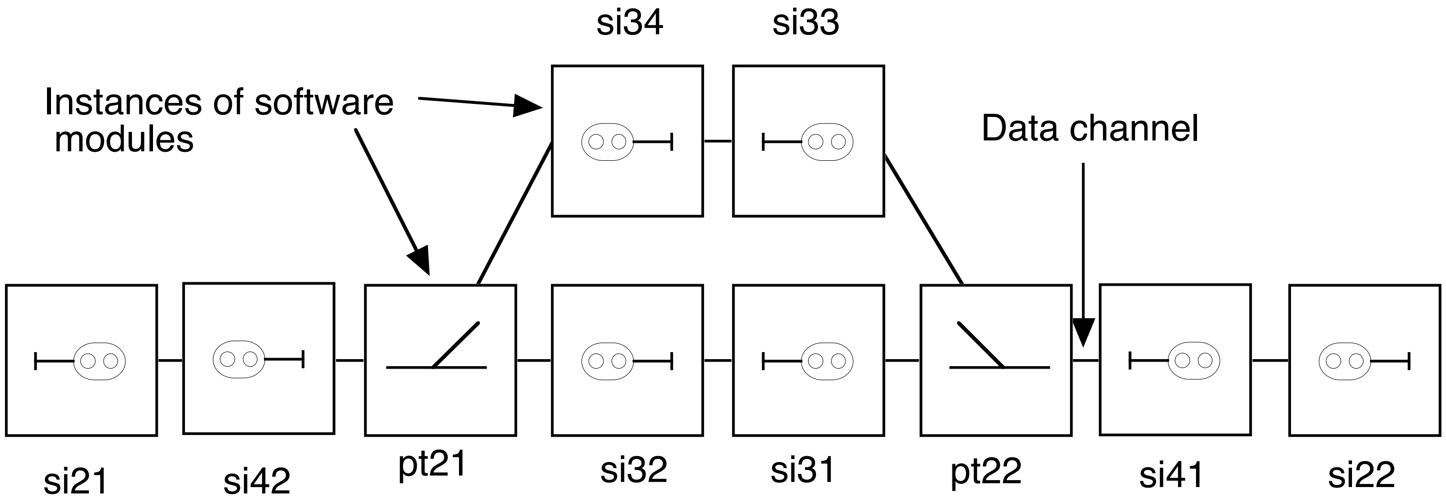
```
Welcome to the GTO Formal Tool 0.5.26
> load sampleGeodata3
> satisfy TRUE
The formula is satisfiable. Time: 0.26
> list partOf entry exit
partOf(route,pt21)
partOf(route,t1)
entry(si21,route)
exit(si31,route)
```

UPPSALA UNIVERSITET

Geographical data at Bombardier

- Formal verification of interlocking software, written in the proprietary language STERNOL.
- Ebilock interlocking systems use data files with geographical data for configuring the software.
- The tools SST (STERNOL specification tool) and SVT (STERNOL verification tool) perform formal verification of STERNOL programs, providing access to geographical data.
- SST+SVT can generate all train routes using domain axioms and exhaustive SAT solving.

Using geographical data with SST



- Specialised specification language includes constructs to access this data:

```
legs(SI42) = 2,
leg(SI42,1) = PT21
leg(SI42,2) = SI21……
```

UPPSALA UNIVERSITET

Dynamic behaviour

(A comment on Dines' talk)

- Railway signalling uses protections to prevent (within reason) accidents when trains do not respect signals. E.g. overlaps, flank protection.
- Protections are typically not defined or determined in a very well-founded manner.
- A domain theory including dynamics of train movements and (probabilistic) behaviour of drivers and equipment could help design protections in a well-founded manner.