A Finite Abstraction of Real Valued Functions for Complete Reasoning about Influence

Florian Bruse Martin Lange Sören Möller

[1]Theoretical Computer Science / Formal Methods, University of Kassel, Germany
[2]TUM School of Computation, Information and Technology
Technical University of Munich, Germany

15th Int. Symp. on Frontiers of Combining Systems, FroCoS'25
29/09 - 01/10/2025

Background and Goals

Background: digitalisation in secondary education
▷ natural science experiments

Goals:

- learning tool that allows reasoning about influences
- allows discussing dangerous/time-consuming experiments in class
- tool should provide feedback and run efficiently

Background and Goals

Background: digitalisation in secondary education

▷ natural science experiments

Goals:

- learning tool that allows reasoning about influences
- allows discussing dangerous/time-consuming experiments in class
- tool should provide feedback and run efficiently

[Bruse/Lange/Möller CADE'23]: The Calculus of Influence

- allows abstractly describing sets of variable influence (e.g. time onto growth)
- provides polynomial reasoning via a set of calculus rules
- is not complete in the general case

Solution: develop a more sophisticated approach

Influences and Experiments

What are we trying to achieve?...

- (abstractly) model influences and experiments using simple mathematical terms
- For all described experiments, decide a hypothesis like:

  "When in between 7km and 8km of altitude, does oxygenation decrease in between 60% and 70%?"'

Influences and Experiments

What are we trying to achieve?...

- (abstractly) model influences and experiments using simple mathematical terms
- For all described experiments, decide a hypothesis like:

"When in between 7km and 8km of altitude, does oxygenation decrease in between 60% and 70%?"'

What are experiments?...

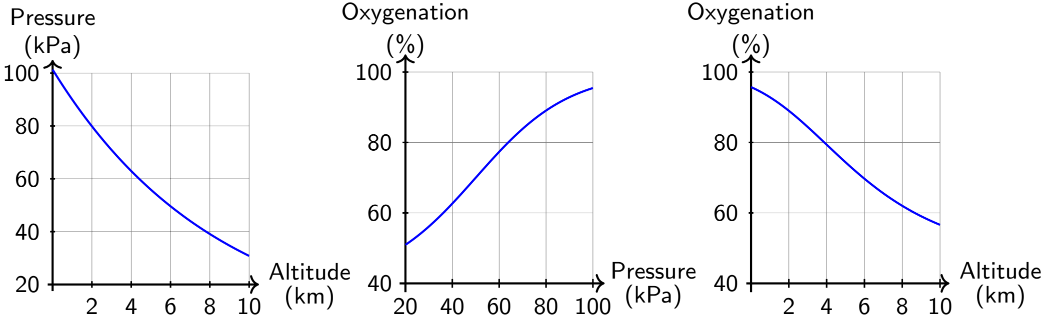
Def.: An influence experiment is a mapping $\mathcal{F} : \mathcal{V} \times \mathcal{V} \to (\mathbb{R} \to \mathbb{R})$ such that:

- $\mathcal{V}$ is a finite signature of ordered variables
- each $\mathcal{F}(a, b) := \mathcal{F}_{a,b}$ is continuous over a closed domain,
- for each $\mathcal{F}_{a,b}$ we have that $a < b$,
- coherence property: for each $a < b < c$ and each $x \in \mathbb{R}$ we have $\mathcal{F}_{a,c}(x) = (\mathcal{F}_{b,c} \circ \mathcal{F}_{a,b})(x)$

Influences and Experiments

A small example of an experiment containing the following influences:

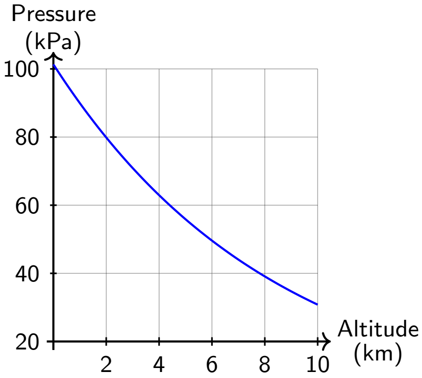
Altitude $\Rightarrow$ Air Pressure $\Rightarrow$ Oxygen Saturation



Influence Statements

Def.: influence statement $S := a \xhookrightarrow{I\, q\, I'} b$, where

- $a, b \in \mathcal{V}$ are variables,
- $I, I' \subseteq \mathbb{R}$ are closed intervals over reals,
- $q \in \{\nearrow, \searrow, \rightsquigarrow, \rightarrow\}$ is the behaviour



We write $\mathcal{F}_{a,b} \models S$, if:

- On the domain I, the (a, b)-influence takes values in I' and behaves like q.

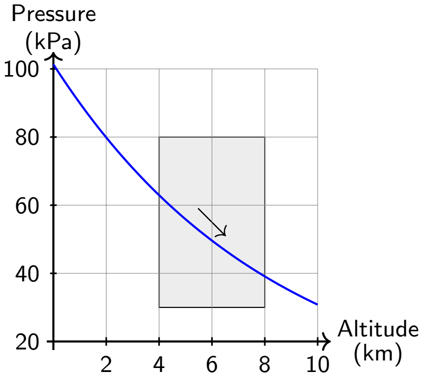
Influence Statements

Def.: influence statement $S := a \xhookrightarrow{I\,q\,I'} b$, where

- $a, b \in \mathcal{V}$ are variables,
- $I, I' \subseteq \mathbb{R}$ are closed intervals over reals,
- $q \in \{\nearrow, \searrow, \rightsquigarrow, \rightarrow\}$ is the behaviour

We write $\mathcal{F}_{a,b} \models S$, if:

- On the domain I, the (a, b)-influence takes values in I' and behaves like q.



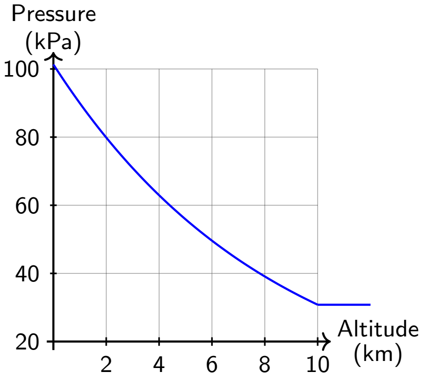
$\mathrm{Alt} \xhookrightarrow{[40,80]\ \searrow\ [30,80]} \mathrm{Press}$

Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$

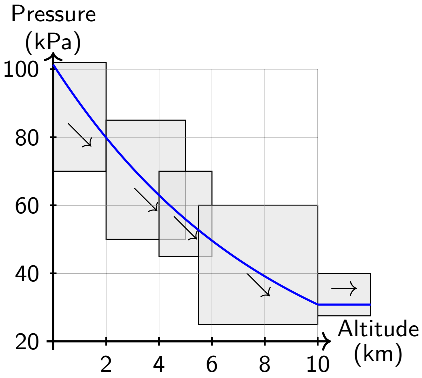


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$

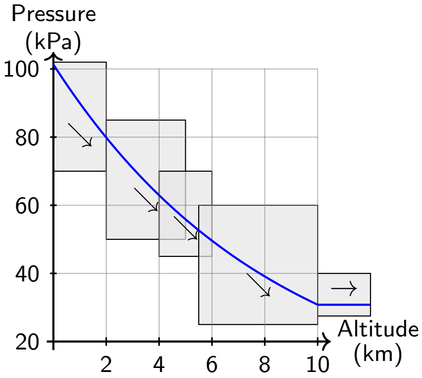


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes

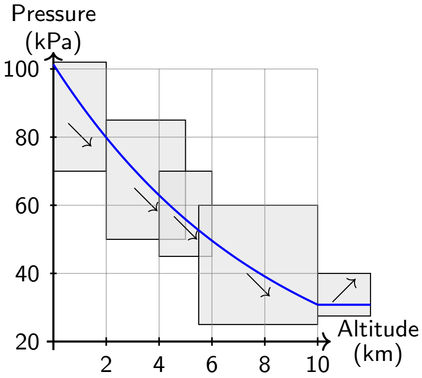


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a,b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes

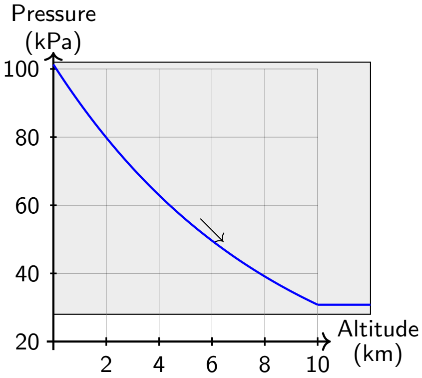


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes

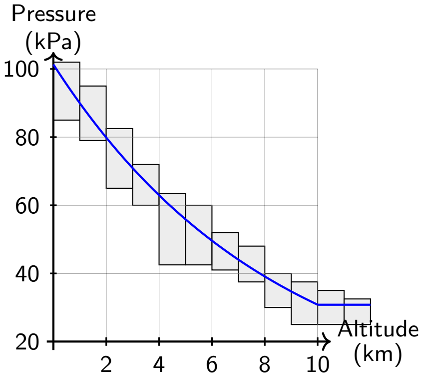


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes

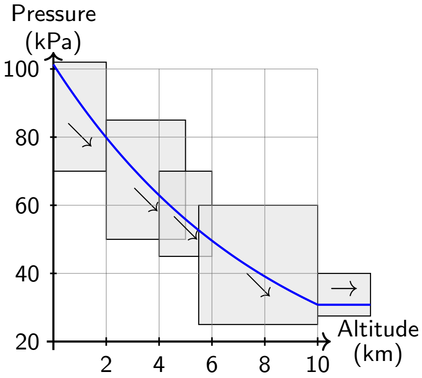


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes
- each scheme has infinitely experiments satisfying it

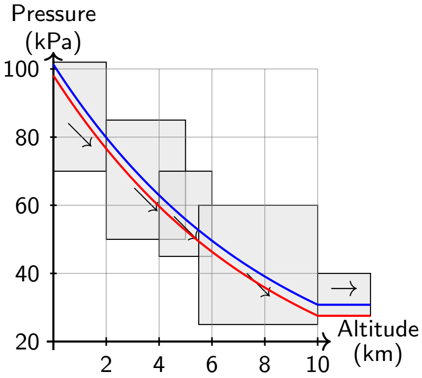


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a, b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes
- each scheme has infinitely experiments satisfying it

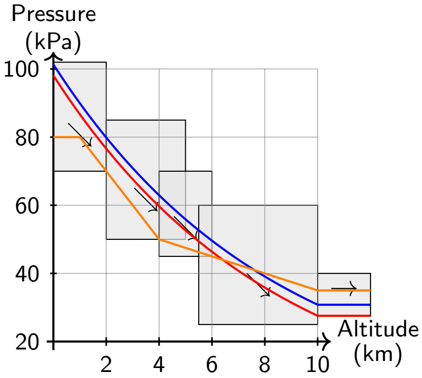


Influence Schemes

Def.: An influence scheme $\mathcal{C}$ is a finite set of influence statements

We write $\mathcal{F}_{a,b} \models \mathcal{C}$ if $\mathcal{F}_{a,b} \models S$ for all $S \in \mathcal{C}_{a,b}$

- $\mathcal{C}_{a,b}$ is the collection of (a,b)-statements in $\mathcal{C}$
- statements in $\mathcal{C}$ are not allowed to include $\rightsquigarrow$
- each experiment can be described by infinitely many schemes
- each scheme has infinitely experiments satisfying it



Hypothesis Validation

Recall the hypothesis from earlier...

"When in between 7km and 8km of altitude, does oxygenation decrease in between 60% and 70%?"'

This directly translates into a (hypothesis)-statement:

$$H := \mathrm{Alt} \xhookrightarrow{[7000{,}8000] \searrow [60{,}70]} \mathrm{Oxy}$$

Hypothesis Validation

Recall the hypothesis from earlier...

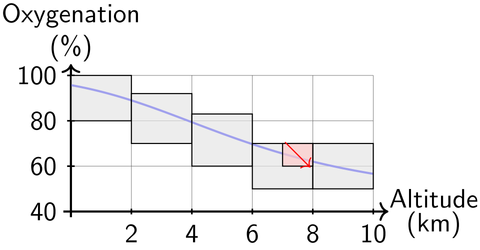
"When in between 7km and 8km of altitude, does oxygenation decrease in between 60% and 70%?"

This directly translates into a (hypothesis)-statement:

$$H := \mathrm{Alt} \xhookrightarrow{[7000,8000]\ \searrow [60,70]} \mathrm{Oxy}$$

Decision Problem:

given:	scheme $\mathcal{C}$ and hypothesis H
problem:	for all experiments $\mathcal{F}$ s.t. $\mathcal{F} \models \mathcal{C}$, does $\mathcal{F} \models H$ hold?



Hypothesis Validation

Recall the hypothesis from earlier...

"When in between 7km and 8km of altitude, does oxygenation decrease in between 60% and 70%?"'

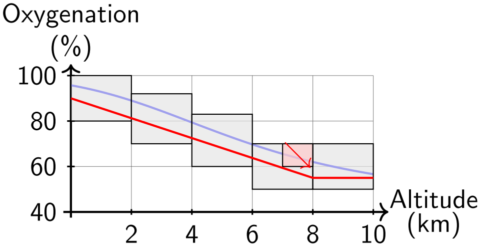
This directly translates into a (hypothesis)-statement:

$$H := \mathrm{Alt} \xhookrightarrow{[7000,8000]\ \searrow [60,70]} \mathrm{Oxy}$$

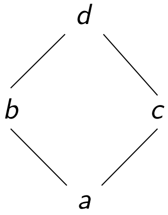
Decision Problem:

given: scheme $\mathcal{C}$ and hypothesis H

problem: for all experiments $\mathcal{F}$ s.t. $\mathcal{F} \models \mathcal{C}$, does $\mathcal{F} \models H$ hold?



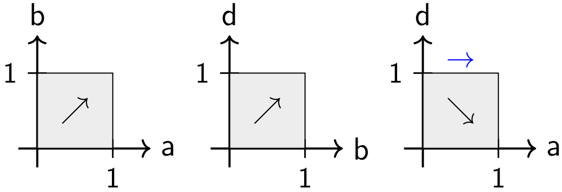
Problems with Completeness



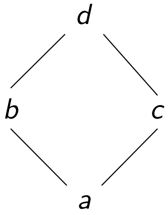
Problems:

- diamonds in the variable order
- statements over non-elementary variable pairs

Consider the composition of statements...



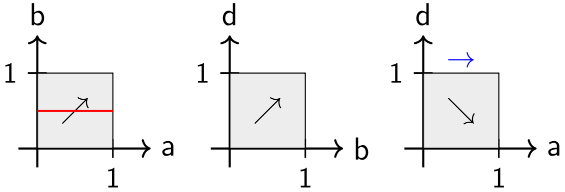
Problems with Completeness



Problems:

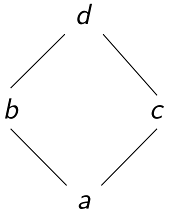
- diamonds in the variable order
- statements over non-elementary variable pairs

Consider the composition of statements...



Introduces non-deteminism and intermediate constraints

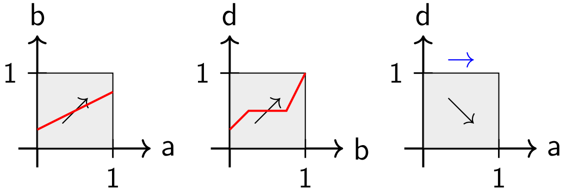
Problems with Completeness



Problems:

- diamonds in the variable order
- statements over non-elementary variable pairs

Consider the composition of statements...



Introduces non-deteminism and intermediate constraints

Alternative Solution

Instead of:

- for all experiments $\mathcal{F}$ s.t. $\mathcal{F} \models \mathcal{C}$ does $\mathcal{F} \models H$ hold?

Do this:

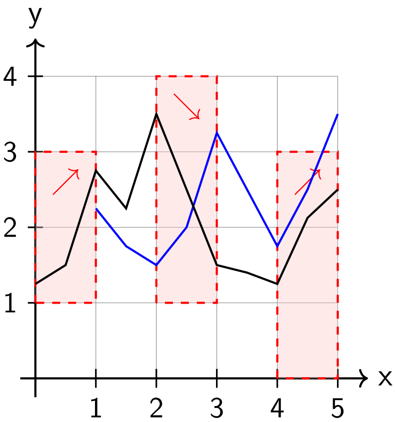
- for all classes of experiments s.t. $\mathcal{F} \models \mathcal{C}$ for all $\mathcal{F}$ in that class, does $\mathcal{F} \models H$ hold for all those $\mathcal{F}$?

This is feasible under the following conditions:

- we have finitely many classes
- for all $S \in \mathcal{C} \cup \{H\}$, we have $\mathcal{F} \models S$ or $\mathcal{F} \not\models S$ for all experiments $\mathcal{F}$ of that class

We will introduce an equivalence relation on influences and naturally extend them to experiments. $\Rightarrow$ $\mathcal{F} \equiv \mathcal{G}$ if $\mathcal{F}_{a,b} \equiv \mathcal{G}_{a,b}$ for all $a, b \in \mathcal{V}$ s.t. $a < b$

An Idea of Categorisation



Initial idea:
Use a grid of boundary points to categorise.

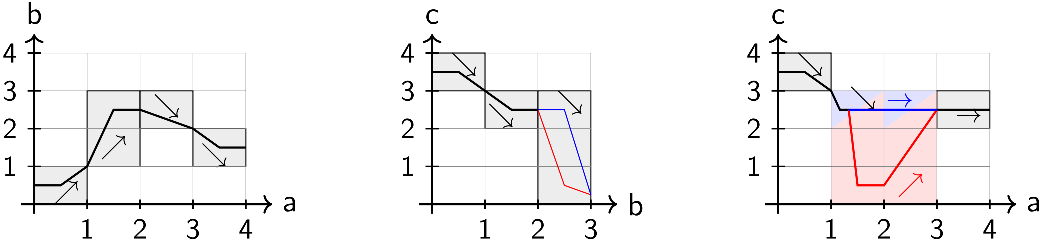
Influences should not be equivalent, if they behave differently on that grid

- different ranges,
- different behaviours,
- different domains

Simplification: Assume all influences are total and assume some integer grid

The Problem with Composition

This does not fix our problem with composition...

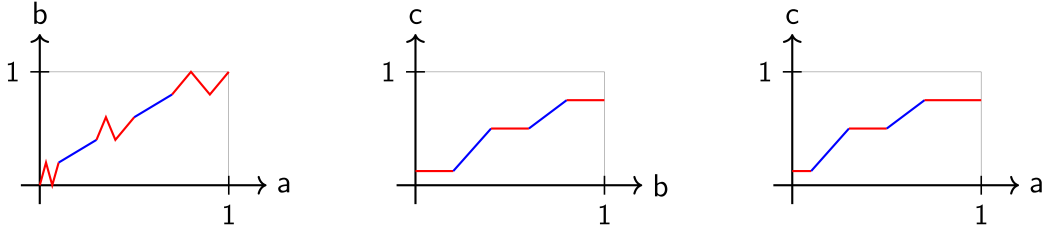


There are two parts to this problem...

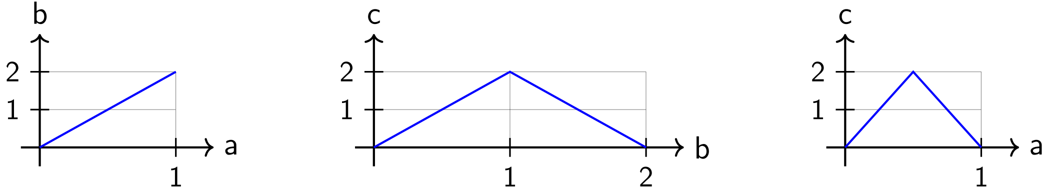
- arbitrary behaviour
- ambiguous ranges

Arbitrary Behaviour

Composing influences over domains exhibiting arbitrary behaviour produces unpredictable behaviour in the composition.

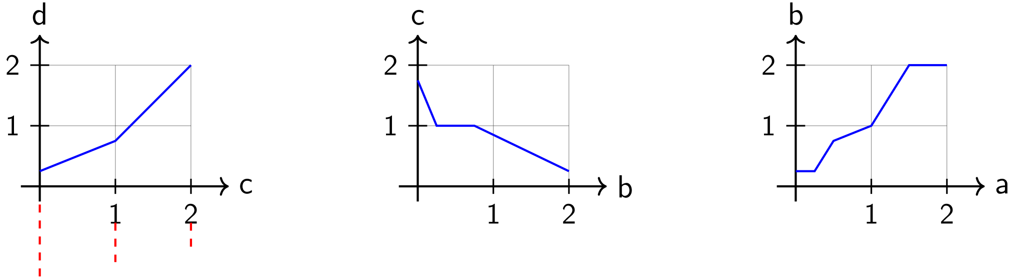


However: Composing non-arbitrary behaviour might introduce arbitrary behavior.



Arbitrary Behaviour: Solution

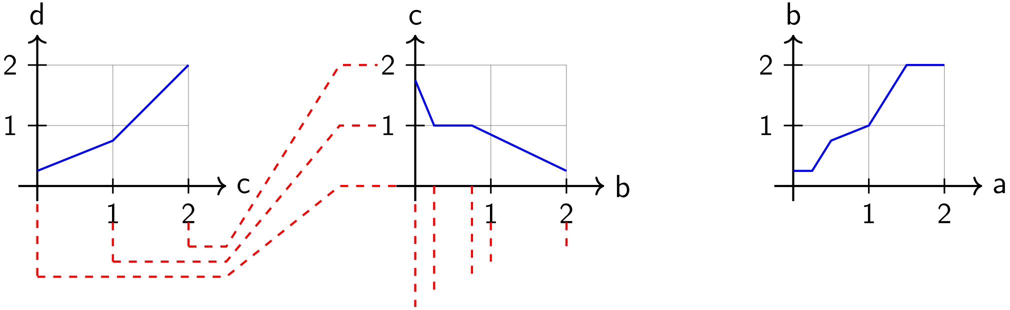
Solution: Decompose the influence into parts where the behaviour might change
$\Rightarrow$ Turning Points



For each $a \in \mathcal{V}$, in a top-down fashion, collect the turning points

Arbitrary Behaviour: Solution

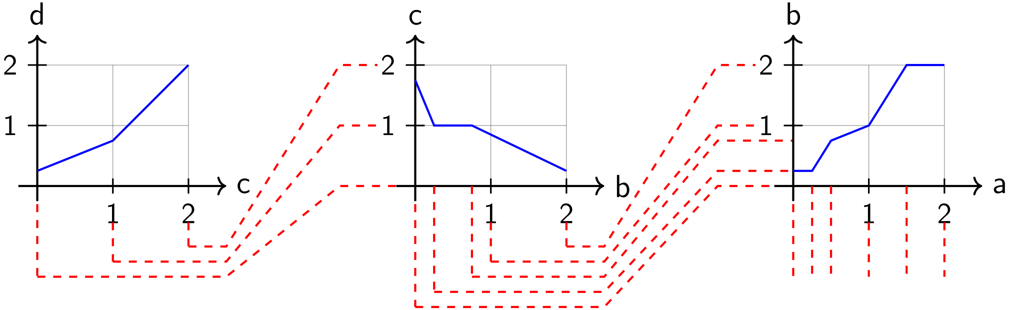
Solution: Decompose the influence into parts where the behaviour might change
$\Rightarrow$ Turning Points



For each $a \in \mathcal{V}$, in a top-down fashion, collect the turning points

Arbitrary Behaviour: Solution

Solution: Decompose the influence into parts where the behaviour might change
$\Rightarrow$ Turning Points

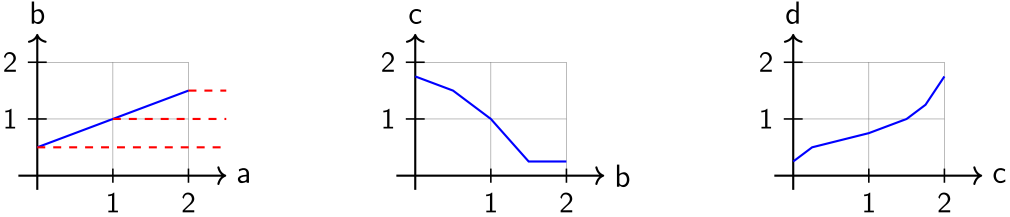


For each $a \in \mathcal{V}$, in a top-down fashion, collect the turning points

Ambigous Ranges: Solution

This does not allow us to compose influence-classes unambigously

Solution: Decompose the influences even further

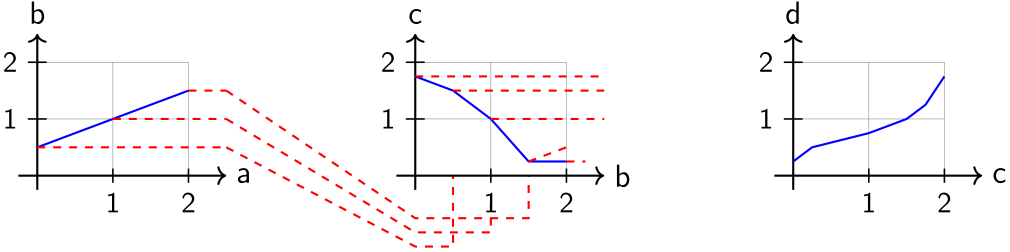


For each $a \in \mathcal{V}$, in a bottom-up fashion, based on the turning points, collect the points of interest

Ambigous Ranges: Solution

This does not allow us to compose influence-classes unambigously

Solution: Decompose the influences even further

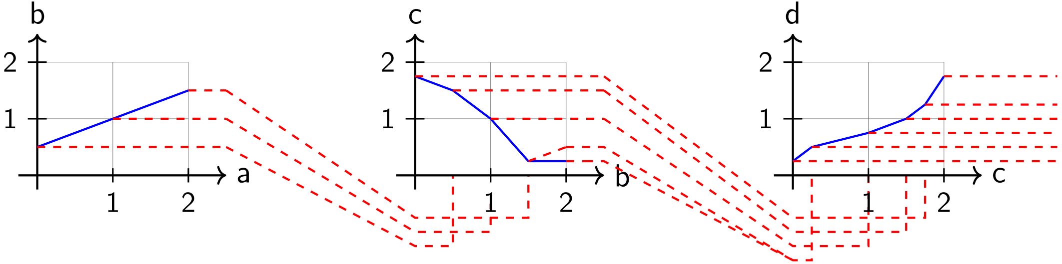


For each $a \in \mathcal{V}$, in a bottom-up fashion, based on the turning points, collect the points of interest

Ambigous Ranges: Solution

This does not allow us to compose influence-classes unambigously

Solution: Decompose the influences even further

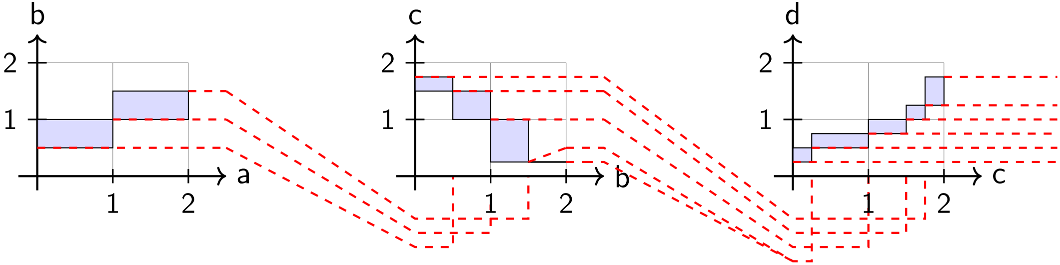


For each $a \in \mathcal{V}$, in a bottom-up fashion, based on the turning points, collect the points of interest

Ambigous Ranges: Solution

This does not allow us to compose influence-classes unambigously

Solution: Decompose the influences even further



For each $a \in \mathcal{V}$, in a bottom-up fashion, based on the turning points, collect the points of interest

Caution: We assumed some experiment and then considered its points of interest
$\Rightarrow$ Consider all possible amounts and orderings of points of interest for all variables

Note: This solves any problems with non-elementary statements and diamonds

Thm 1. given scheme $\mathcal{C}$ and hypothesis H, deciding $\mathcal{C} \models H$ is in coNP

- for all variables, collect all classes of points of interest
- there are only polynomial many points of interest w.r.t. $\mathcal{C}$
  (but exponentially many w.r.t. $\mathcal{V}$)
- filter out classes where the influences do not compose correctly
- for each remaining class, check whether the experiments satisfy $\mathcal{C}$ but not H

Future Work

Biggest Problem: Runtime Efficiency

- Implement a coNP-procedure using SMT-Solvers over different background theories
- Benchmark different solvers
- Fix the overshoot
- Develop a hyposesis-driven (polynomial?) approach

Additionally...

- Extend the results to multi-dimensional experiments
- Enhance the practicality by user-friendly infra-structure