Extracting Confident General Concept Inclusions from Finite Interpretations

Daniel Borchmann

16. Oct. 2012

Overview

1. Motivation and Introduction
2. Formal Concept Analysis
3. Description Logics
4. Confident General Concept Inclusions of Finite Interpretations

Automation is a Good Idea

Goal (of AI)

Let computers do what they can do (and leave all the rest to the humans.)

Automation is a Good Idea

Goal (of AI)

Let computers do what they can do (and leave all the rest to the humans.)

Requirement

Computers must know about the real world.

Automation is a Good Idea

Goal (of AI)

Let computers do what they can do (and leave all the rest to the humans.)

Requirement

Computers must know about the real world.

Goal (of Knowledge Representation)

Represent knowledge in a way suitable for computers,

Automation is a Good Idea

Goal (of AI)

Let computers do what they can do (and leave all the rest to the humans.)

Requirement

Computers must know about the real world.

Goal (of Knowledge Representation)

Represent knowledge in a way suitable for computers, i. e. as a *description logics ontology*.

Description Logics Ontologies

Ontologies contain assertional knowledge and terminological knowledge.

Description Logics Ontologies

Ontologies contain assertional knowledge and terminological knowledge.

Example ($\mathcal{EL}^{\bot}$-Ontology)

$(\mathcal{T}, \mathcal{A})$ is an ontology, where

$$\begin{aligned}\mathcal{T} &= \{\, \mathsf{Cat} \sqsubseteq \mathsf{Animal} \sqcap \exists \mathsf{hunts}.\mathsf{Mouse}, \\ &\qquad \mathsf{Cat} \sqcap \mathsf{Mouse} \sqsubseteq \bot \,\} \\ \mathcal{A} &= \{\, \mathsf{Cat}(\mathsf{Tom}), \mathsf{Mouse}(\mathsf{Jerry}), \mathsf{hunts}(\mathsf{Tom}, \mathsf{Jerry}) \,\}\end{aligned}$$

Description Logics Ontologies

Ontologies contain assertional knowledge and terminological knowledge.

Example ($\mathcal{EL}^{\bot}$-Ontology)

$(\mathcal{T}, \mathcal{A})$ is an ontology, where

$$\begin{aligned}\mathcal{T} = \{\, & \mathsf{Cat} \sqsubseteq \mathsf{Animal} \sqcap \exists \mathsf{hunts}.\mathsf{Mouse}, \\ & \mathsf{Cat} \sqcap \mathsf{Mouse} \sqsubseteq \bot \,\} \\ \mathcal{A} = \{\, & \mathsf{Cat}(\mathsf{Tom}), \mathsf{Mouse}(\mathsf{Jerry}), \mathsf{hunts}(\mathsf{Tom}, \mathsf{Jerry}) \,\}\end{aligned}$$

Definition

Terminological axioms of the form $C \sqsubseteq D$ are called *general concept inclusions* (GCIs.)

Description Logics Ontologies

Problem

Construction of real world ontologies is a difficult task

Description Logics Ontologies

Problem

Construction of real world ontologies is a difficult task

But unstructured information is often already available (i. e. as textual publication)

Description Logics Ontologies

Problem

Construction of real world ontologies is a difficult task

But unstructured information is often already available (i. e. as textual publication)

Goal

Automatically construct ontologies from unstructured data

Description Logics Ontologies

Problem

Construction of real world ontologies is a difficult task

But unstructured information is often already available (i. e. as textual publication)

Goal

Semi-Automatically construct the terminological part of ontologies from unstructured data

Unstructured Data

Question

What is unstructured data?

Unstructured Data

Question

What is unstructured data?

Example (RDF Triples)

Unstructured Data

Question

What is unstructured data?

Example (RDF Triples)

```
<http://dbpedia.org/resource/Autism>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://dbpedia.org/ontology/Disease> .

<http://dbpedia.org/resource/Aristotle>
<http://dbpedia.org/ontology/influenced>
<http://dbpedia.org/resource/Western_philosophy> .
```

Unstructured Data

Question

What is unstructured data?

Example (RDF Triples)

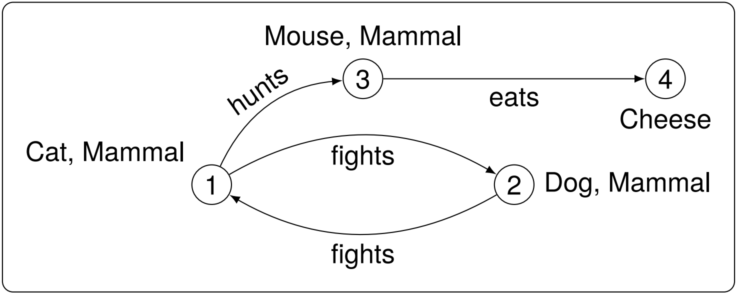
```
<http://dbpedia.org/resource/Autism>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://dbpedia.org/ontology/Disease> .

<http://dbpedia.org/resource/Aristotle>
<http://dbpedia.org/ontology/influenced>
<http://dbpedia.org/resource/Western_philosophy> .
```

Approach

Unstructured data is given as a *finite interpretation* (finite vertex- and edge-labeled graphs)

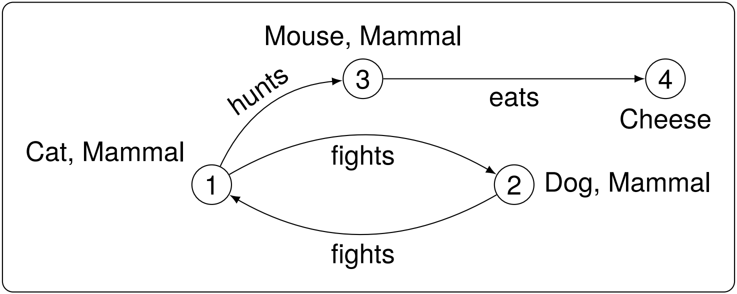
A Simple Example

Example (Interpretation $\mathcal{I}_{\text{pets}}$)



A Simple Example

Example (Interpretation $\mathcal{I}_{\text{pets}}$)



The *elements* (vertices) satisfying $C = \text{Mammal} \sqcap \exists \text{hunts.Mouse}$ are

$$C^{\mathcal{I}} = \{ 1 \}.$$

$C^{\mathcal{I}}$ is called the *extension* of C.

Terminological Knowledge of Interpretations

Goal

Extract all *terminological knowledge*, i. e. all valid GCIs, from $\mathcal{I}$.

Terminological Knowledge of Interpretations

Definition

Let C, D be $\mathcal{EL}^{\perp}$-concept descriptions. Then the GCI $C \sqsubseteq D$ *holds* in $\mathcal{I}$ if and only if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

Terminological Knowledge of Interpretations

Definition

Let C, D be $\mathcal{EL}^{\bot}$-concept descriptions. Then the GCI $C \sqsubseteq D$ *holds* in $\mathcal{I}$ if and only if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

Problem

The number of valid GCIs of $\mathcal{I}$ is (normally) infinite.

Terminological Knowledge of Interpretations

Definition

Let C, D be $\mathcal{EL}^{\perp}$-concept descriptions. Then the GCI $C \sqsubseteq D$ *holds* in $\mathcal{I}$ if and only if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

Problem

The number of valid GCIs of $\mathcal{I}$ is (normally) infinite.

Example

Cat $\sqsubseteq$ Mammal holds in $\mathcal{I}_{\text{pets}}$,

Terminological Knowledge of Interpretations

Definition

Let C, D be $\mathcal{EL}^{\bot}$-concept descriptions. Then the GCI $C \sqsubseteq D$ *holds* in $\mathcal{I}$ if and only if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

Problem

The number of valid GCIs of $\mathcal{I}$ is (normally) infinite.

Example

$\mathsf{Cat} \sqsubseteq \mathsf{Mammal}$ holds in $\mathcal{I}_{\mathsf{pets}}$, and so do

$$\exists\mathsf{hunts}.\mathsf{Cat} \sqsubseteq \exists\mathsf{hunts}.\mathsf{Mammal},$$
$$\exists\mathsf{hunts}.\exists\mathsf{hunts}.\mathsf{Cat} \sqsubseteq \exists\mathsf{hunts}.\exists\mathsf{hunts}.\mathsf{Mammal}, \ldots$$

Bases of Valid GCIs

Approach

Consider *bases* of valid GCIs of $\mathcal{I}$, i. e. sets $\mathcal{B}$ of GCIs such that

Bases of Valid GCIs

Approach

Consider *bases* of valid GCIs of $\mathcal{I}$, i. e. sets $\mathcal{B}$ of GCIs such that

- $\mathcal{B}$ contains only valid GCIs of $\mathcal{I}$ ($\mathcal{B}$ is *sound*)

Bases of Valid GCIs

Approach

Consider *bases* of valid GCIs of $\mathcal{I}$, i. e. sets $\mathcal{B}$ of GCIs such that

- $\mathcal{B}$ contains only valid GCIs of $\mathcal{I}$ ($\mathcal{B}$ is *sound*)
- every valid GCI of $\mathcal{I}$ already *follows* from $\mathcal{B}$ ($\mathcal{B}$ is *complete*.)

Bases of Valid GCIs

Approach

Consider *bases* of valid GCIs of $\mathcal{I}$, i. e. sets $\mathcal{B}$ of GCIs such that

- $\mathcal{B}$ contains only valid GCIs of $\mathcal{I}$ ($\mathcal{B}$ is *sound*)
- every valid GCI of $\mathcal{I}$ already *follows* from $\mathcal{B}$ ($\mathcal{B}$ is *complete*.)

Goal

Find a *finite* base of all valid GCIs of $\mathcal{I}$.

Bases of Valid GCIs

Approach

Consider *bases* of valid GCIs of $\mathcal{I}$, i. e. sets $\mathcal{B}$ of GCIs such that

- $\mathcal{B}$ contains only valid GCIs of $\mathcal{I}$ ($\mathcal{B}$ is *sound*)
- every valid GCI of $\mathcal{I}$ already *follows* from $\mathcal{B}$ ($\mathcal{B}$ is *complete*.)

Goal

Find a *finite* base of all valid GCIs of $\mathcal{I}$.

Theorem (Baader, Distel 2008)

Finite bases of all valid $\mathcal{EL}^{\bot}$-GCIs of $\mathcal{I}$ always exists. One can be constructed effectively.

Problem: Errors in DBpedia

Problem

Approach assumes data set $\mathcal{I}$ to be complete and free of errors.

Problem: Errors in DBpedia

Problem

Approach assumes data set $\mathcal{I}$ to be complete and free of errors.

Example

The GCI

$$\exists \mathsf{child}.\top \sqsubseteq \mathsf{Person}$$

does not hold in $\mathcal{I}_{\mathsf{DBpedia}}$,

Problem: Errors in DBpedia

Problem

Approach assumes data set $\mathcal{I}$ to be complete and free of errors.

Example

The GCI

$$\exists \text{child}.\top \sqsubseteq \text{Person}$$

does not hold in $\mathcal{I}_{\text{DBpedia}}$, but there are only four *erroneous counterexamples* (in 5262 individuals.)

Problem: Errors in DBpedia

Problem

Approach assumes data set $\mathcal{I}$ to be complete and free of errors.

Example

The GCI

$$\exists \text{child}.\top \sqsubseteq \text{Person}$$

does not hold in $\mathcal{I}_{\text{DBpedia}}$, but there are only four *erroneous counterexamples* (in 5262 individuals.)

Idea

Consider *confident GCIs*, i. e. GCIs that allow some few “exceptions.”

Formal Concept Analysis

From Mathematical Order Theory to a «Theory of Data»

Formal Concept Analysis

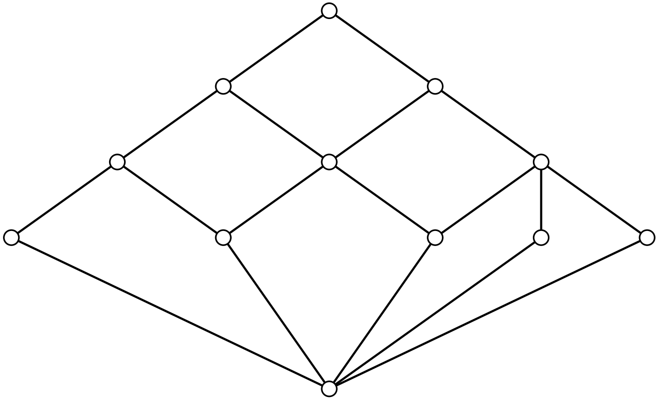
What is FCA?

Formal Concept Analysis is a restructuring attempt to modern lattice theory.

Formal Concept Analysis

What is FCA?

Formal Concept Analysis is a restructuring attempt to modern lattice theory.



Formal Concept Analysis

Motivation for FCA (back in the 1980s)

Formal Concept Analysis

Motivation for FCA (back in the 1980s)

- Claim: lattice theory has turned into a meaningless manipulation of symbols

Formal Concept Analysis

Motivation for FCA (back in the 1980s)

- Claim: lattice theory has turned into a meaningless manipulation of symbols
- Goal: (re)introduce *meaning* into this theory

Formal Concept Analysis

Motivation for FCA (back in the 1980s)

- Claim: lattice theory has turned into a meaningless manipulation of symbols
- Goal: (re)introduce *meaning* into this theory
- Use a theory of *concepts* for this

Formal Concept Analysis

Motivation for FCA (back in the 1980s)

- Claim: lattice theory has turned into a meaningless manipulation of symbols
- Goal: (re)introduce *meaning* into this theory
- Use a theory of *concepts* for this

Literature

- Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts; R. Wille 1982
- Formal Concept Analysis Mathematical Foundations; R. Wille and B. Ganter; 1999

Formal Contexts

The fundamental notion of FCA is the one of a *formal context*.

Formal Contexts

The fundamental notion of FCA is the one of a *formal context*.

Definition

Let G, M be sets and let $I \subseteq G \times M$. Then the triple $\mathbb{K} = (G, M, I)$ is called a *formal context*.

Formal Contexts

The fundamental notion of FCA is the one of a *formal context*.

Definition

Let G, M be sets and let $I \subseteq G \times M$. Then the triple $\mathbb{K} = (G, M, I)$ is called a *formal context*.

Example

$$(\{1, \dots, 5\}, \{1, \dots, 5\}, \{(x, y) \mid x \leqslant y\})$$

Formal Contexts

The fundamental notion of FCA is the one of a *formal context*.

Definition

Let G, M be sets and let $I \subseteq G \times M$. Then the triple $\mathbb{K} = (G, M, I)$ is called a *formal context*.

Example

$$(\{1, \ldots, 5\}, \{1, \ldots, 5\}, \{(x, y) \mid x \leqslant y\})$$

Uhm …

Meaning?

Formal Contexts – Basic Interpretation

Let $\mathbb{K} = (G, M, I)$ be a formal context.
We then introduce the following interpretation:

Formal Contexts – Basic Interpretation

Let $\mathbb{K} = (G, M, I)$ be a formal context.
We then introduce the following interpretation:

- Elements of G are called *objects* (Gegenstände)

Formal Contexts – Basic Interpretation

Let $\mathbb{K} = (G, M, I)$ be a formal context.
We then introduce the following interpretation:

- Elements of G are called *objects* (Gegenstände)
- Elements of M are called *attributes* (Merkmale)

Formal Contexts – Basic Interpretation

Let $\mathbb{K} = (G, M, I)$ be a formal context.
We then introduce the following interpretation:

- Elements of G are called *objects* (Gegenstände)
- Elements of M are called *attributes* (Merkmale)
- We say that the object g *has* the attribute m if and only if $(g, m) \in I$

Formal Contexts – Graphical Representation

	1	2	3	4	5
1	×	×	×	×	×
2		×	×	×	×
3			×	×	×
4				×	×
5					×

Formal Contexts – Graphical Representation

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Formal Concepts

Definition (Derivation Operators)

Let $A \subseteq G, B \subseteq M$. Then we define

$$A' := \{\, m \in M \mid \forall g \in A \colon g\ I\ m \,\},$$
$$B' := \{\, g \in G \mid \forall m \in B \colon g\ I\ m \,\}.$$

Formal Concepts

Definition (Derivation Operators)

Let $A \subseteq G, B \subseteq M$. Then we define

$$A' := \{\, m \in M \mid \forall g \in A \colon g\ I\ m \,\},$$
$$B' := \{\, g \in G \mid \forall m \in B \colon g\ I\ m \,\}.$$

Definition (Formal Concepts)

The pair (A, B) is called a *formal concept* of $\mathbb{K}$ if and only if $A \subseteq G, B \subseteq M$ and

$$A' = B \quad \text{and} \quad B' = A.$$

Formal Concepts

Definition (Derivation Operators)

Let $A \subseteq G, B \subseteq M$. Then we define

$$A' := \{\, m \in M \mid \forall g \in A \colon g \mathrel{I} m \,\},$$
$$B' := \{\, g \in G \mid \forall m \in B \colon g \mathrel{I} m \,\}.$$

Definition (Formal Concepts)

The pair (A, B) is called a *formal concept* of $\mathbb{K}$ if and only if $A \subseteq G, B \subseteq M$ and

$$A' = B \quad \text{and} \quad B' = A.$$

The set of all formal contexts of $\mathbb{K}$ is denoted by $\mathfrak{B}(\mathbb{K})$.

Formal Concepts – Example

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Formal Concepts – Example

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Example (Formal Concepts)

Formal Concepts – Example

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Example (Formal Concepts)

- $(\{\text{Mercury, Venus, Earth, Mars, Pluto}\}, \{\text{small}\})$

Formal Concepts – Example

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Example (Formal Concepts)

- $(\{\text{ Mercury, Venus, Earth, Mars, Pluto }\}, \{\text{ small }\}) \mathrel{\hat{=}}$ *small planets*

Formal Concepts – Example

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Example (Formal Concepts)

- $(\{\text{Mercury, Venus, Earth, Mars, Pluto}\}, \{\text{small}\}) \mathrel{\hat{=}}$ *small planets*
- $(\{\text{Pluto}\}, \{\text{small, far, moon}\})$

Formal Concepts – Example

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Example (Formal Concepts)

- ({ Mercury, Venus, Earth, Mars, Pluto }, { small }) $\hat{=}$ *small planets*
- ({ Pluto }, { small, far, moon }) $\hat{=}$ *small planets far away from sun*

Concept Lattices

Observation

Concepts can be ordered by *generality*.

Concept Lattices

Observation

Concepts can be ordered by *generality*.

Example (Formal Concepts)

- $(\{\,$Mercury, Venus, Earth, Mars, Pluto$\,\}, \{\,$small$\,\}) \mathrel{\hat{=}}$ *small planets*
- $(\{\,$Pluto$\,\}, \{\,$small, far, moon$\,\}) \mathrel{\hat{=}}$ *small planets far away from sun*

Concept Lattices

Observation

Concepts can be ordered by *generality*.

Example (Formal Concepts)

- $(\{\,\text{Mercury, Venus, Earth, Mars, Pluto}\,\}, \{\,\text{small}\,\}) \mathrel{\hat{=}}$ *small planets*
- $(\{\,\text{Pluto}\,\}, \{\,\text{small, far, moon}\,\}) \mathrel{\hat{=}}$ *small planets far away from sun*

Definition

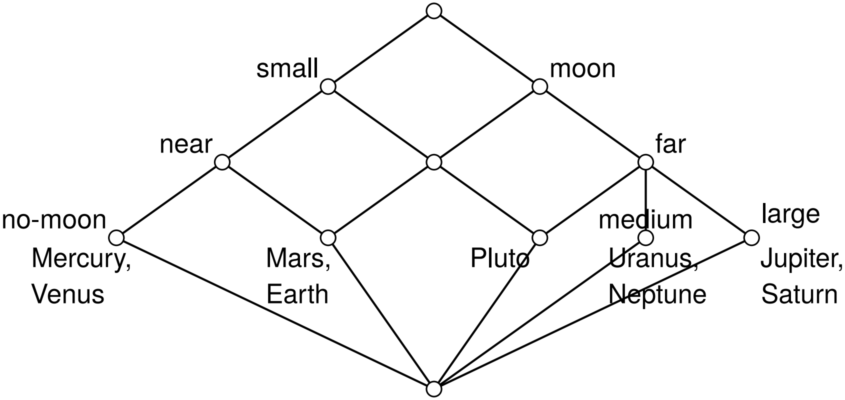
Let $(A_1, B_1), (A_2, B_2) \in \mathfrak{B}(\mathbb{K})$. Then define

$$(A_1, B_1) \leqslant (A_2, B_2) \iff A_1 \subseteq A_2.$$

Concept Lattices

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Concept Lattices



Implications

FCA can also be used to examine *dependencies* between attributes of $\mathbb{K}$.

Implications

FCA can also be used to examine *dependencies* between attributes of $\mathbb{K}$.

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Implications

FCA can also be used to examine *dependencies* between attributes of $\mathbb{K}$.

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Observation

Every planet, that is far away from sun has a moon.

Implications

FCA can also be used to examine *dependencies* between attributes of $\mathbb{K}$.

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Observation

Every planet, that is far away from sun has a moon.

far planet $\triangleq$

Implications

FCA can also be used to examine *dependencies* between attributes of $\mathbb{K}$.

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Observation

Every planet, that is far away from sun has a moon.

far planet $\hat{=}$ ({ Jupiter, Saturn, Uranus, Neptune, Pluto },

Implications

FCA can also be used to examine *dependencies* between attributes of $\mathbb{K}$.

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Observation

Every planet, that is far away from sun has a moon.

$$\text{far planet} \mathrel{\hat{=}} (\{\,\text{Jupiter, Saturn, Uranus, Neptune, Pluto}\,\}, \{\,\text{far, moon}\,\}).$$

Implications

Definition (Implication (Syntax))

Let M be a set, $A, B \subseteq M$. Then the pair (A, B) may be called an *implication* on M and is written as $A \longrightarrow B$.

Implications

Definition (Implication (Syntax))

Let M be a set, $A, B \subseteq M$. Then the pair (A, B) may be called an *implication* on M and is written as $A \longrightarrow B$.

Definition (Implication (Semantics))

Let $\mathbb{K} = (G, M, I)$ be a formal context and let $A \longrightarrow B$ be an implication M.

Implications

Definition (Implication (Syntax))

Let M be a set, $A, B \subseteq M$. Then the pair (A, B) may be called an *implication* on M and is written as $A \longrightarrow B$.

Definition (Implication (Semantics))

Let $\mathbb{K} = (G, M, I)$ be a formal context and let $A \longrightarrow B$ be an implication M. Then $A \longrightarrow b$ *holds* in $\mathbb{K}$ if and only if

$$A' \subseteq B'.$$

Implications

Definition (Implication (Syntax))

Let M be a set, $A, B \subseteq M$. Then the pair (A, B) may be called an *implication* on M and is written as $A \longrightarrow B$.

Definition (Implication (Semantics))

Let $\mathbb{K} = (G, M, I)$ be a formal context and let $A \longrightarrow B$ be an implication M. Then $A \longrightarrow b$ *holds* in $\mathbb{K}$ if and only if

$$A' \subseteq B'.$$

Remark

$A \longrightarrow B$ holds in $\mathbb{K}$ if and only if all objects that have all attributes from A also have all attributes from B.

Implications

Definition (Implication (Syntax))

Let M be a set, $A, B \subseteq M$. Then the pair (A, B) may be called an *implication* on M and is written as $A \longrightarrow B$.

Definition (Implication (Semantics))

Let $\mathbb{K} = (G, M, I)$ be a formal context and let $A \longrightarrow B$ be an implication M. Then $A \longrightarrow b$ *holds* in $\mathbb{K}$ if and only if

$$A' \subseteq B'.$$

Remark

$A \longrightarrow B$ holds in $\mathbb{K}$ if and only if all objects that have all attributes from A also have all attributes from B.

This is a *model-based* semantics!

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find all valid implications of $\mathbb{K}$

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find *a good representation* of all valid implications of $\mathbb{K}$

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find a *good representation* of all valid implications of $\mathbb{K}$

Definition

Let $\mathcal{B}$ be a set of implications of $\mathbb{K}$.

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find a *good representation* of all valid implications of $\mathbb{K}$

Definition

Let $\mathcal{B}$ be a set of implications of $\mathbb{K}$.

- $\mathcal{B}$ is called *sound*, if all implications in $\mathcal{B}$ hold in $\mathbb{K}$;

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find a *good representation* of all valid implications of $\mathbb{K}$

Definition

Let $\mathcal{B}$ be a set of implications of $\mathbb{K}$.

- $\mathcal{B}$ is called *sound*, if all implications in $\mathcal{B}$ hold in $\mathbb{K}$;
- $\mathcal{B}$ is called *complete*, if all implications valid in $\mathbb{K}$ follow from $\mathcal{B}$.

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find a *good representation* of all valid implications of $\mathbb{K}$

Definition

Let $\mathcal{B}$ be a set of implications of $\mathbb{K}$.

- $\mathcal{B}$ is called *sound*, if all implications in $\mathcal{B}$ hold in $\mathbb{K}$;
- $\mathcal{B}$ is called *complete*, if all implications valid in $\mathbb{K}$ follow from $\mathcal{B}$.

$\mathcal{B}$ is called a *base* if it is sound and complete.

Bases of Implications

Recall

Want to find a finite base of all GCIs of a finite interpretation

In terms of FCA

Find a *good representation* of all valid implications of $\mathbb{K}$

Definition

Let $\mathcal{B}$ be a set of implications of $\mathbb{K}$.

- $\mathcal{B}$ is called *sound*, if all implications in $\mathcal{B}$ hold in $\mathbb{K}$;
- $\mathcal{B}$ is called *complete*, if all implications valid in $\mathbb{K}$ follow from $\mathcal{B}$.

$\mathcal{B}$ is called a *base* if it is sound and complete. $\mathcal{B}$ is called an *irredundant* base if $\mathcal{B}$ is a base and every proper subset $\mathcal{B}' \subsetneq \mathcal{B}$ is not a base.

Canonical Base

One can explicitly describe some bases of $\mathbb{K}$

Canonical Base

One can explicitly describe some bases of $\mathbb{K}$

Theorem

The set

$$\{ A \longrightarrow A'' \mid A \subseteq M \}$$

is a base of $\mathbb{K}$*.*

Canonical Base

One can explicitly describe some bases of $\mathbb{K}$

Theorem

The set

$$\{ A \longrightarrow A'' \mid A \subseteq M \}$$

is a base of $\mathbb{K}$.

This base is in general not irredundant.

Canonical Base

One can explicitly describe some bases of $\mathbb{K}$

Theorem

The set

$$\{ A \longrightarrow A'' \mid A \subseteq M \}$$

is a base of $\mathbb{K}$*.*

This base is in general not irredundant.

Remark

One can explicitly describe a base of $\mathbb{K}$ *with minimal cardinality*, the so-called *canonical base* of $\mathbb{K}$.

Description Logics

Formalizing Knowledge The Right Way

What are Description Logics about?

In a Nutshell

Description Logics are formal languages to represent knowledge that provide methods to reason about this knowledge.

What are Description Logics about?

In a Nutshell

Description Logics are formal languages to represent knowledge that provide methods to reason about this knowledge.

What are Description Logics about?

In a Nutshell

Description Logics are formal languages to represent knowledge that provide methods to reason about this knowledge.

What are Description Logics about?

In a Nutshell

Description Logics are formal languages to represent knowledge that provide methods to reason about this knowledge.

The Plan

- Syntax of $\mathcal{ALC}$
- Semantics of $\mathcal{ALC}$
- TBoxes, ABoxes and Ontologies
- Standard Reasoning Tasks

Literature

Baader et. al. ; *The Description Logic Handbook, Theory, Implementation and Applications*; Cambridge University Press; second edition; 2007

Syntax of $\mathcal{ALC}$

Fix the following sets:

- N_C of *concept names*
- N_R of *role names*

Syntax of $\mathcal{ALC}$

Fix the following sets:

- N_C of *concept names*
- N_R of *role names*

Example

$$N_C = \{ \text{Person}, \text{Male}, \text{Female} \}$$
$$N_R = \{ \text{hasChild} \}$$

Syntax of $\mathcal{AL}$

Definition (Syntax of $\mathcal{ALC}$)

The following terms form the set $\mathcal{C}$ of all $\mathcal{ALC}$-concept descriptions

Syntax of $\mathcal{AL}$

Definition (Syntax of $\mathcal{ALC}$)

The following terms form the set $\mathcal{C}$ of all $\mathcal{ALC}$-concept descriptions

- $\top, \bot$ (universal and bottom concept)
- A for $A \in N_C$ (atomic concepts)
- $\neg C$ for $C \in \mathcal{C}$ (negation)
- $C \sqcap D$ for $C, D \in \mathcal{C}$ (conjunction)
- $C \sqcup D$ for $C, D \in \mathcal{C}$ (disjunction)
- $\forall r.C$ for $r \in N_R, C \in \mathcal{C}$ (value restriction)
- $\exists r.C$ for $r \in N_R, C \in \mathcal{C}$ (existential restriction)

Syntax of $\mathcal{AL}$

Definition (Syntax of $\mathcal{ALC}$)

The following terms form the set $\mathcal{C}$ of all $\mathcal{ALC}$-concept descriptions

- $\top, \bot$ (universal and bottom concept)
- A for $A \in N_C$ (atomic concepts)
- $\neg C$ for $C \in \mathcal{C}$ (negation)
- $C \sqcap D$ for $C, D \in \mathcal{C}$ (conjunction)
- $C \sqcup D$ for $C, D \in \mathcal{C}$ (disjunction)
- $\forall r.C$ for $r \in N_R, C \in \mathcal{C}$ (value restriction)
- $\exists r.C$ for $r \in N_R, C \in \mathcal{C}$ (existential restriction)

Example

$$\text{Person} \sqcap \text{Female} \sqcap \exists \text{hasChild}.\top \sqcap \forall \text{hasChild}.\text{Male}$$

Syntax of $\mathcal{AL}$

Definition (Syntax of $\mathcal{ALC}$)

The following terms form the set $\mathcal{C}$ of all $\mathcal{ALC}$-concept descriptions

- $\top, \bot$ (universal and bottom concept)
- A for $A \in N_C$ (atomic concepts)
- $\neg C$ for $C \in \mathcal{C}$ (negation)
- $C \sqcap D$ for $C, D \in \mathcal{C}$ (conjunction)
- $C \sqcup D$ for $C, D \in \mathcal{C}$ (disjunction)
- $\forall r.C$ for $r \in N_R, C \in \mathcal{C}$ (value restriction)
- $\exists r.C$ for $r \in N_R, C \in \mathcal{C}$ (existential restriction)

Example

$$\text{Person} \sqcap \text{Female} \sqcap \exists \text{hasChild}.\top \sqcap \forall \text{hasChild}.\text{Male}$$

A mother which has only sons.

Interpretations

Semantics of description logics are defined using *interpretations*.

Interpretations

Semantics of description logics are defined using *interpretations*.

Definition

An *interpretation* $\mathcal{I}$ is a pair $(\Delta_{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta_{\mathcal{I}}$ is a set and $\cdot^{\mathcal{I}}$ is mapping such that

Interpretations

Semantics of description logics are defined using *interpretations*.

Definition

An *interpretation* $\mathcal{I}$ is a pair $(\Delta_\mathcal{I}, \cdot^\mathcal{I})$ where $\Delta_\mathcal{I}$ is a set and $\cdot^\mathcal{I}$ is mapping such that

- $A^\mathcal{I} \subseteq \Delta_\mathcal{I}$ for each $A \in N_C$

Interpretations

Semantics of description logics are defined using *interpretations*.

Definition

An *interpretation* $\mathcal{I}$ is a pair $(\Delta_\mathcal{I}, \cdot^\mathcal{I})$ where $\Delta_\mathcal{I}$ is a set and $\cdot^\mathcal{I}$ is mapping such that

- $A^\mathcal{I} \subseteq \Delta_\mathcal{I}$ for each $A \in N_C$
- $r^\mathcal{I} \subseteq \Delta_\mathcal{I} \times \Delta_\mathcal{I}$ for each $r \in N_R$

Interpretations

Example

Consider $\Delta_{\mathcal{I}} = \{ 1, 2, 3, 4 \}$ and

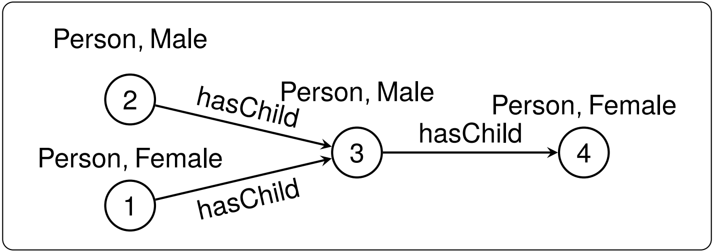
$$\begin{aligned} \mathsf{Person}^{\mathcal{I}} &= \{ 1, 2, 3, 4 \} \\ \mathsf{Male}^{\mathcal{I}} &= \{ 2, 3 \} \\ \mathsf{Female}^{\mathcal{I}} &= \{ 1, 4 \} \\ \mathsf{hasChild}^{\mathcal{I}} &= \{ (1, 3), (2, 3), (3, 4) \}. \end{aligned}$$

Interpretations

Example

Consider $\Delta_{\mathcal{I}} = \{1, 2, 3, 4\}$ and

$$\begin{aligned}
\mathsf{Person}^{\mathcal{I}} &= \{1, 2, 3, 4\} \\
\mathsf{Male}^{\mathcal{I}} &= \{2, 3\} \\
\mathsf{Female}^{\mathcal{I}} &= \{1, 4\} \\
\mathsf{hasChild}^{\mathcal{I}} &= \{(1, 3), (2, 3), (3, 4)\}.
\end{aligned}$$



Semantics of $\mathcal{ALC}$

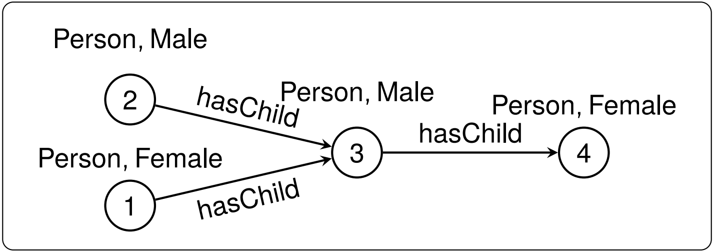
Definition

Let C, D be $\mathcal{ALC}$-concept descriptions, $r \in N_R$.

$$\begin{aligned}
\top^{\mathcal{I}} &= \Delta_{\mathcal{I}} \\
\bot^{\mathcal{I}} &= \emptyset \\
(C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
(C \sqcup D)^{\mathcal{I}} &= C^{\mathcal{I}} \cup D^{\mathcal{I}} \\
(\forall r.C)^{\mathcal{I}} &= \{\, x \in \Delta_{\mathcal{I}} \mid \forall y \in \Delta_{\mathcal{I}} \colon (x, y) \in r^{\mathcal{I}} \implies y \in C^{\mathcal{I}} \,\} \\
(\exists r.C)^{\mathcal{I}} &= \{\, x \in \Delta_{\mathcal{I}} \mid \exists y \in \Delta_{\mathcal{I}} \colon (x, y) \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}} \,\}
\end{aligned}$$

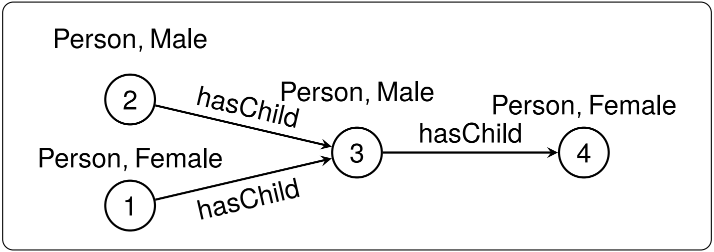
Semantics of $\mathcal{ALC}$

Example



Semantics of $\mathcal{ALC}$

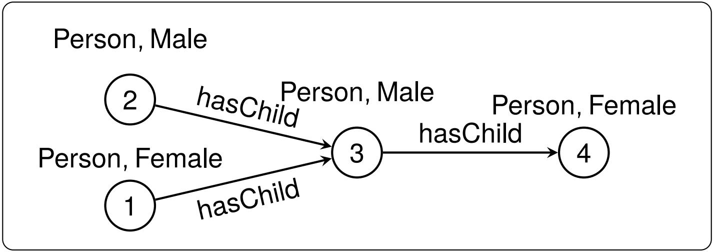
Example



$$(\text{Person} \sqcap \text{Female} \sqcap \exists \text{hasChild}.\top \sqcap \forall \text{hasChild}.\text{Male})^{\mathcal{I}} =$$

Semantics of $\mathcal{ALC}$

Example



$$(\mathsf{Person} \sqcap \mathsf{Female} \sqcap \exists \mathsf{hasChild}.\top \sqcap \forall \mathsf{hasChild}.\mathsf{Male})^{\mathcal{I}} = \{\, 1 \,\}$$

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Different forms of knowledge:

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Different forms of knowledge:

- *terminological knowledge*, i. e. “a cat is a mammal which hunts mice”

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Different forms of knowledge:

- *terminological knowledge*, i. e. “a cat is a mammal which hunts mice”
  $\rightsquigarrow$ TBoxes $\mathcal{T}$

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Different forms of knowledge:

- *terminological knowledge*, i. e. “a cat is a mammal which hunts mice”
  $\rightsquigarrow$ TBoxes $\mathcal{T}$
- *assertional knowledge*, i. e. “Tom is a cat”

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Different forms of knowledge:

- *terminological knowledge*, i. e. “a cat is a mammal which hunts mice”
  $\rightsquigarrow$ TBoxes $\mathcal{T}$
- *assertional knowledge*, i. e. “Tom is a cat”
  $\rightsquigarrow$ ABoxes $\mathcal{A}$

Description Logic Ontologies

Goal

Use Description Logics to represent knowledge

Different forms of knowledge:

- *terminological knowledge*, i. e. “a cat is a mammal which hunts mice”
  $\rightsquigarrow$ TBoxes $\mathcal{T}$
- *assertional knowledge*, i. e. “Tom is a cat”
  $\rightsquigarrow$ ABoxes $\mathcal{A}$

Definition (Ontology)

An *ontology* is a pair $(\mathcal{T}, \mathcal{A})$, where $\mathcal{T}$ is a TBox and $\mathcal{A}$ is an ABox.

Description Logic Ontologies – An Example

Example ($\mathcal{EL}^{\bot}$-Ontology)

$(\mathcal{T}, \mathcal{A})$ is an ontology, where

$$\begin{aligned}\mathcal{T} &= \{\, \mathsf{Cat} \sqsubseteq \mathsf{Animal} \sqcap \exists \mathsf{hunts}.\mathsf{Mouse}, \\ &\qquad \mathsf{Cat} \sqcap \mathsf{Mouse} \sqsubseteq \bot \,\} \\ \mathcal{A} &= \{\, \mathsf{Cat}(\mathsf{Tom}), \mathsf{Mouse}(\mathsf{Jerry}), \mathsf{hunts}(\mathsf{Tom}, \mathsf{Jerry}) \,\}\end{aligned}$$

Terminological Knowledge and TBoxes

Definition (Terminological Axioms)

Terminological Axioms are of the form

Terminological Knowledge and TBoxes

Definition (Terminological Axioms)

Terminological Axioms are of the form

- $A \equiv C$, where C is a concept description and $A \notin N_C$ is a *defined concept name* (*concept definition*)

Terminological Knowledge and TBoxes

Definition (Terminological Axioms)

Terminological Axioms are of the form

- $A \equiv C$, where C is a concept description and $A \notin N_C$ is a *defined concept name* (*concept definition*)
- $C \sqsubseteq D$, where C, D are concept descriptions (*general concept inclusion*)

Terminological Knowledge and TBoxes

Definition (Terminological Axioms)

Terminological Axioms are of the form

- $A \equiv C$, where C is a concept description and $A \notin N_C$ is a *defined concept name* (*concept definition*)
- $C \sqsubseteq D$, where C, D are concept descriptions (*general concept inclusion*)

A TBox $\mathcal{T}$ is a finite set of terminological axioms, where each defined concept name appears at most once.

Terminological Knowledge and TBoxes

Definition (Terminological Axioms)

Terminological Axioms are of the form

- $A \equiv C$, where C is a concept description and $A \notin N_C$ is a *defined concept name* (*concept definition*)
- $C \sqsubseteq D$, where C, D are concept descriptions (*general concept inclusion*)

A TBox $\mathcal{T}$ is a finite set of terminological axioms, where each defined concept name appears at most once.

Example

$$\mathcal{T} = \{\, \text{Cat} \sqsubseteq \text{Animal} \sqcap \exists \text{hunts.Mouse}, \\ \text{Cat} \sqcap \text{Mouse} \sqsubseteq \bot \,\}$$

TBox Semantics

Definition (Descriptive Semantics)

An interpretation $\mathcal{I} = (\Delta_{\mathcal{I}}, \cdot^{\mathcal{I}})$ is a *model* of a TBox $\mathcal{T}$ if and only if

$$A^{\mathcal{I}} = C^{\mathcal{I}} \quad \text{and} \quad C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$$

for all $(A \equiv C), (C \sqsubseteq D) \in \mathcal{T}$.

TBox Semantics

Definition (Descriptive Semantics)

An interpretation $\mathcal{I} = (\Delta_{\mathcal{I}}, \cdot^{\mathcal{I}})$ is a *model* of a TBox $\mathcal{T}$ if and only if

$$A^{\mathcal{I}} = C^{\mathcal{I}} \quad \text{and} \quad C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$$

for all $(A \equiv C), (C \sqsubseteq D) \in \mathcal{T}$.

Extend the interpretation function $\cdot^{\mathcal{I}}$ to all defined concept names such that

$$A^{\mathcal{I}} \subseteq \Delta_{\mathcal{I}}.$$

TBox Semantics

Definition (Descriptive Semantics)

An interpretation $\mathcal{I} = (\Delta_{\mathcal{I}}, \cdot^{\mathcal{I}})$ is a *model* of a TBox $\mathcal{T}$ if and only if

$$A^{\mathcal{I}} = C^{\mathcal{I}} \quad \text{and} \quad C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$$

for all $(A \equiv C), (C \sqsubseteq D) \in \mathcal{T}$.

Extend the interpretation function $\cdot^{\mathcal{I}}$ to all defined concept names such that

$$A^{\mathcal{I}} \subseteq \Delta_{\mathcal{I}}.$$

Other semantics:

- greatest fixpoint semantics
- least fixpoint semantics

Confident GCIs of Finite Interpretations

Handling Errors in Knowledge

Work by Baader and Distel

Theorem (Baader, Distel 2008)

Finite bases of all valid $\mathcal{EL}^{\bot}$-GCIs of $\mathcal{I}$ always exists. One can be constructed effectively.

Work by Baader and Distel

Theorem (Baader, Distel 2008)

Finite bases of all valid $\mathcal{EL}^{\bot}$-GCIs of $\mathcal{I}$ always exists. One can be constructed effectively.

Goal

Extend approach to also handle errors.

Work by Baader and Distel

Theorem (Baader, Distel 2008)

Finite bases of all valid $\mathcal{EL}^{\perp}$-GCIs of $\mathcal{I}$ always exists. One can be constructed effectively.

Goal

Extend approach to also handle errors.

Plan

- Introduce necessary terminology
- Define *confident GCIs* as an approach to handle errors
- Discuss some relevant ideas from FCA
- Present first results

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$.

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$.

Questions:

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$*.*

Questions:

- What is $M_{\mathcal{I}}$?

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$*.*

Questions:

- What is $M_{\mathcal{I}}$? $\rightsquigarrow$ set of concept descriptions (no more details here)

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$.

Questions:

- What is $M_{\mathcal{I}}$? $\rightsquigarrow$ set of concept descriptions (no more details here)
- What is $\bigsqcap U$?

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$*.*

Questions:

- What is $M_{\mathcal{I}}$? $\rightsquigarrow$ set of concept descriptions (no more details here)
- What is $\bigsqcap U$?
- What is $((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}}$?

In More Detail

Theorem

The set

$$\mathcal{B}_2 := \{ \bigsqcap U \sqsubseteq ((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}} \mid U \subseteq M_{\mathcal{I}} \}$$

is a finite base of $\mathcal{I}$.

Questions:

- What is $M_{\mathcal{I}}$? $\rightsquigarrow$ set of concept descriptions (no more details here)
- What is $\bigsqcap U$?
- What is $((\bigsqcap U)^{\mathcal{I}})^{\mathcal{I}}$?

Definition

$$\bigsqcap U := \begin{cases} \top & U = \emptyset \\ \bigsqcap_{V \in U} V & \text{otherwise.} \end{cases}$$

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Definition

A concept description C is a *model-based most-specific concept description* of X in $\mathcal{I}$ iff

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Definition

A concept description C is a *model-based most-specific concept description* of X in $\mathcal{I}$ iff

- $C^{\mathcal{I}} \supseteq X$,

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Definition

A concept description C is a *model-based most-specific concept description* of X in $\mathcal{I}$ iff

- $C^{\mathcal{I}} \supseteq X$,
- if D is a concept description such that $D^{\mathcal{I}} \supseteq X$, then $C \sqsubseteq D$.

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Definition

A concept description C is a *model-based most-specific concept description* of X in $\mathcal{I}$ iff

- $C^{\mathcal{I}} \supseteq X$,
- if D is a concept description such that $D^{\mathcal{I}} \supseteq X$, then $C \sqsubseteq D$.

Observation

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Definition

A concept description C is a *model-based most-specific concept description* of X in $\mathcal{I}$ iff

- $C^{\mathcal{I}} \supseteq X$,
- if D is a concept description such that $D^{\mathcal{I}} \supseteq X$, then $C \sqsubseteq D$.

Observation

- C (as above) is a *most specific concept description that describes* X.

Model-Based Most-Specific Concept Descriptions

Let $X \subseteq \Delta_{\mathcal{I}}$. Then $X^{\mathcal{I}}$ denotes the *model-based most-specific concept description* of X in $\mathcal{I}$.

Definition

A concept description C is a *model-based most-specific concept description* of X in $\mathcal{I}$ iff

- $C^{\mathcal{I}} \supseteq X$,
- if D is a concept description such that $D^{\mathcal{I}} \supseteq X$, then $C \sqsubseteq D$.

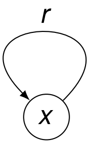
Observation

- C (as above) is a *most specific concept description that describes* X.
- C is unique up to equivalence, denoted by $X^{\mathcal{I}}$.

Model-Based Most-Specific Concept Descriptions

Problem

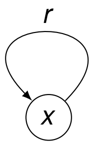
Model-based most-specific concept descriptions do not need to exist in $\mathcal{EL}^{\bot}$.



Model-Based Most-Specific Concept Descriptions

Problem

Model-based most-specific concept descriptions do not need to exist in $\mathcal{EL}^\bot$.

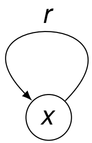


Solution: Consider $\mathcal{EL}^\bot_{\mathrm{gfp}}$ concept descriptions.

Model-Based Most-Specific Concept Descriptions

Problem

Model-based most-specific concept descriptions do not need to exist in $\mathcal{EL}^{\bot}$.



Solution: Consider $\mathcal{EL}^{\bot}_{\mathrm{gfp}}$ concept descriptions.

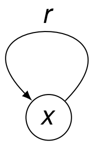
Lemma

In $\mathcal{EL}^{\bot}_{\mathrm{gfp}}$ model-based most-specific concept descriptions always exist.

Model-Based Most-Specific Concept Descriptions

Problem

Model-based most-specific concept descriptions do not need to exist in $\mathcal{EL}^{\bot}$.



Solution: Consider $\mathcal{EL}^{\bot}_{\mathrm{gfp}}$ concept descriptions.

Lemma

In $\mathcal{EL}^{\bot}_{gfp}$ model-based most-specific concept descriptions always exist.

Lemma

If $\mathcal{B}$ is an $\mathcal{EL}^{\bot}_{gfp}$-base of $\mathcal{I}$, then one can effectively compute an $\mathcal{EL}^{\bot}$ base $\mathcal{B}'$ from $\mathcal{B}$.

Ontologies from Data: an Example

Experiment (B. 2010)

DBpedia: automatically extracted RDF triples from Wikipedia Infoboxes

Ontologies from Data: an Example

Experiment (B. 2010)

DBpedia: automatically extracted RDF triples from Wikipedia Infoboxes

Take relation hasChild $\leadsto$ interpretation $\mathcal{I}_{\text{DBpedia}}$

Ontologies from Data: an Example

Experiment (B. 2010)

DBpedia: automatically extracted RDF triples from Wikipedia Infoboxes
Take relation hasChild $\rightsquigarrow$ interpretation $\mathcal{I}_{\text{DBpedia}}$
$|\Delta_{\mathcal{I}_{\text{DBpedia}}}| = 5626$, Base of GCIs of size 1252.

Ontologies from Data: an Example

Experiment (B. 2010)

DBpedia: automatically extracted RDF triples from Wikipedia Infoboxes
Take relation hasChild $\rightsquigarrow$ interpretation $\mathcal{I}_{\text{DBpedia}}$
$|\Delta_{\mathcal{I}_{\text{DBpedia}}}| = 5626$, Base of GCIs of size 1252.

Observation

$$\exists \text{hasChild}.\top \sqsubseteq \text{Person}$$

does not hold in $\mathcal{I}_{\text{DBpedia}}$

Ontologies from Data: an Example

Experiment (B. 2010)

DBpedia: automatically extracted RDF triples from Wikipedia Infoboxes

Take relation hasChild $\leadsto$ interpretation $\mathcal{I}_{\text{DBpedia}}$

$|\Delta_{\mathcal{I}_{\text{DBpedia}}}| = 5626$, Base of GCIs of size 1252.

Observation

$$\exists \text{hasChild}.\top \sqsubseteq \text{Person}$$

does not hold in $\mathcal{I}_{\text{DBpedia}}$, but there are only 4 *erroneous* counterexamples.

Ontologies from Data: an Example

Experiment (B. 2010)

DBpedia: automatically extracted RDF triples from Wikipedia Infoboxes

Take relation hasChild $\rightsquigarrow$ interpretation $\mathcal{I}_{\text{DBpedia}}$

$|\Delta_{\mathcal{I}_{\text{DBpedia}}}| = 5626$, Base of GCIs of size 1252.

Observation

$$\exists \text{hasChild}.\top \sqsubseteq \text{Person}$$

does not hold in $\mathcal{I}_{\text{DBpedia}}$, but there are only 4 *erroneous* counterexamples.

Idea

Also consider GCIs that “almost” hold in $\mathcal{I}_{\text{DBpedia}}$.

Confidence of GCIs

Definition

The *confidence* of $C \sqsubseteq D$ in $\mathcal{I}$ is defined as

$$\mathrm{conf}_{\mathcal{I}}(C \sqsubseteq D) := \begin{cases} 1 & \text{if } C^{\mathcal{I}} = \emptyset, \\ \frac{|(C \sqcap D)^{\mathcal{I}}|}{|C^{\mathcal{I}}|} & \text{otherwise.} \end{cases}$$

Confidence of GCIs

Definition

The *confidence* of $C \sqsubseteq D$ in $\mathcal{I}$ is defined as

$$\mathrm{conf}_{\mathcal{I}}(C \sqsubseteq D) := \begin{cases} 1 & \text{if } C^{\mathcal{I}} = \emptyset, \\ \frac{|(C \sqcap D)^{\mathcal{I}}|}{|C^{\mathcal{I}}|} & \text{otherwise.} \end{cases}$$

Let $c \in [0, 1]$. Define $\mathrm{Th}_c(\mathcal{I})$ as the set of all GCIs having confidence of at least c in $\mathcal{I}$.

Confidence of GCIs

Definition

The *confidence* of $C \sqsubseteq D$ in $\mathcal{I}$ is defined as

$$\mathrm{conf}_{\mathcal{I}}(C \sqsubseteq D) := \begin{cases} 1 & \text{if } C^{\mathcal{I}} = \emptyset, \\ \frac{|(C \sqcap D)^{\mathcal{I}}|}{|C^{\mathcal{I}}|} & \text{otherwise.} \end{cases}$$

Let $c \in [0, 1]$. Define $\mathrm{Th}_c(\mathcal{I})$ as the set of all GCIs having confidence of at least c in $\mathcal{I}$.

Approach

Consider $\mathrm{Th}_c(\mathcal{I})$ as set of “almost” valid GCIs of $\mathcal{I}$.

Confidence of GCIs

Definition

The *confidence* of $C \sqsubseteq D$ in $\mathcal{I}$ is defined as

$$\mathrm{conf}_{\mathcal{I}}(C \sqsubseteq D) := \begin{cases} 1 & \text{if } C^{\mathcal{I}} = \emptyset, \\ \frac{|(C \sqcap D)^{\mathcal{I}}|}{|C^{\mathcal{I}}|} & \text{otherwise.} \end{cases}$$

Let $c \in [0, 1]$. Define $\mathrm{Th}_c(\mathcal{I})$ as the set of all GCIs having confidence of at least c in $\mathcal{I}$.

Approach

Consider $\mathrm{Th}_c(\mathcal{I})$ as set of “almost” valid GCIs of $\mathcal{I}$.

Question

Can we find a *finite base* for $\mathrm{Th}_c(\mathcal{I})$?

A Base for Confident GCIs

Answer

There exist finite bases of $\mathsf{Th}_c(\mathcal{I})$.

A Base for Confident GCIs

Answer

There exist finite bases of $\mathsf{Th}_c(\mathcal{I})$.

Use ideas from Formal Concept Analysis for this!

Implications with Confidence

Definition

For an implication $A \longrightarrow B$ of a formal context $\mathbb{K}$ define its *confidence* to be

$$\operatorname{conf}_{\mathbb{K}}(A \longrightarrow B) := \begin{cases} 1 & A' = \emptyset \\ \frac{|(A \cup B)'|}{|A'|} & \text{otherwise.} \end{cases}$$

Implications with Confidence

Definition

For an implication $A \longrightarrow B$ of a formal context $\mathbb{K}$ define its *confidence* to be

$$\operatorname{conf}_{\mathbb{K}}(A \longrightarrow B) := \begin{cases} 1 & A' = \emptyset \\ \frac{|(A \cup B)'|}{|A'|} & \text{otherwise.} \end{cases}$$

Goal

Find "small" representation of all implications with confidence at least $c \in [0, 1]$.

Implications with Confidence

Definition

For an implication $A \longrightarrow B$ of a formal context $\mathbb{K}$ define its *confidence* to be

$$\operatorname{conf}_{\mathbb{K}}(A \longrightarrow B) := \begin{cases} 1 & A' = \emptyset \\ \frac{|(A \cup B)'|}{|A'|} & \text{otherwise.} \end{cases}$$

Goal

Find "small" representation of all implications with confidence at least $c \in [0, 1]$. More precisely, let

$$\operatorname{Th}_c(\mathbb{K}) := \{ A \longrightarrow B \mid \operatorname{conf}_{\mathbb{K}}(A \longrightarrow B) \geqslant c \},$$

Implications with Confidence

Definition

For an implication $A \longrightarrow B$ of a formal context $\mathbb{K}$ define its *confidence* to be

$$\mathrm{conf}_{\mathbb{K}}(A \longrightarrow B) := \begin{cases} 1 & A' = \varnothing \\ \frac{|(A \cup B)'|}{|A'|} & \text{otherwise.} \end{cases}$$

Goal

Find "small" representation of all implications with confidence at least $c \in [0, 1]$. More precisely, let

$$\mathrm{Th}_c(\mathbb{K}) := \{ A \longrightarrow B \mid \mathrm{conf}_{\mathbb{K}}(A \longrightarrow B) \geqslant c \},$$

Then: find a set $\mathcal{B} \subseteq \mathrm{Th}_c(\mathbb{K})$ that is *complete* for $\mathrm{Th}_c(\mathbb{K})$, i. e. that entails all implications from $\mathrm{Th}_c(\mathbb{K})$.

Implications with Confidence

Observation

Plan (Luxenburger)

Implications with Confidence

Observation

Plan (Luxenburger)

- Restrict attention to implications with confidence < 1

Implications with Confidence

Observation

Plan (Luxenburger)

- Restrict attention to implications with confidence < 1
- Consider only implications of the form $A'' \longrightarrow B''$, where $B'' \supseteq A''$

Implications with Confidence

Observation

Plan (Luxenburger)

- Restrict attention to implications with confidence < 1
- Consider only implications of the form $A'' \longrightarrow B''$, where $B'' \supseteq A''$
- Consider only implications $A'' \longrightarrow B''$ where A'' and B'' are *directly neighbored*

Implications with Confidence

Observation

Plan (Luxenburger)

- Restrict attention to implications with confidence < 1
- Consider only implications of the form $A'' \longrightarrow B''$, where $B'' \supseteq A''$
- Consider only implications $A'' \longrightarrow B''$ where A'' and B'' are *directly neighbored*

Lemma

For $A \subseteq B \subseteq C \subseteq M$ it is true that

$$\mathrm{conf}_{\mathbb{K}}(A \longrightarrow C) = \mathrm{conf}_{\mathbb{K}}(A \longrightarrow B) \cdot \mathrm{conf}_{\mathbb{K}}(B \longrightarrow C).$$

Implications with Confidence

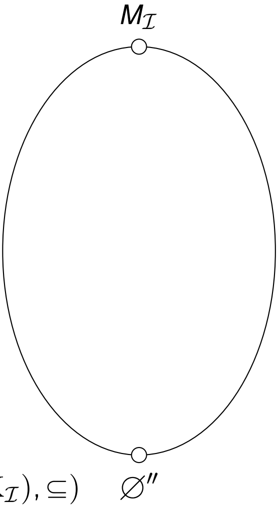
Theorem

Let $\mathbb{K} = (G, M, I)$ *be a finite non-empty formal context and* $c \in [0, 1]$. *Let* $\mathcal{B}$ *be a base of* $\mathbb{K}$ *and define*

$$\mathcal{C} := \{\, A'' \longrightarrow C'' \mid A \subseteq C \subseteq M, \operatorname{conf}_{\mathbb{K}}(A'' \longrightarrow C'') \in [c, 1), \\ \nexists B'' : A'' \subsetneq B'' \subsetneq C'' \,\}.$$

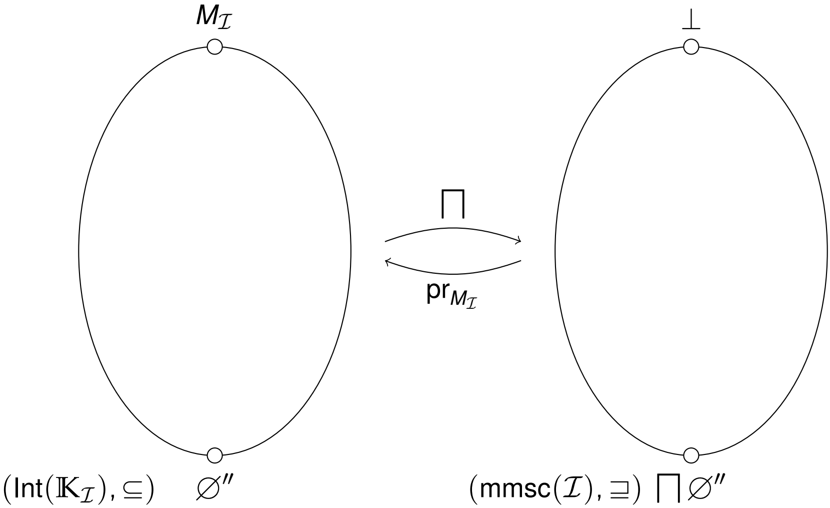
Then $\mathcal{B} \cup \mathcal{C}$ *is a base of* $\operatorname{Th}_c(\mathbb{K})$.

An Order Isomorphism

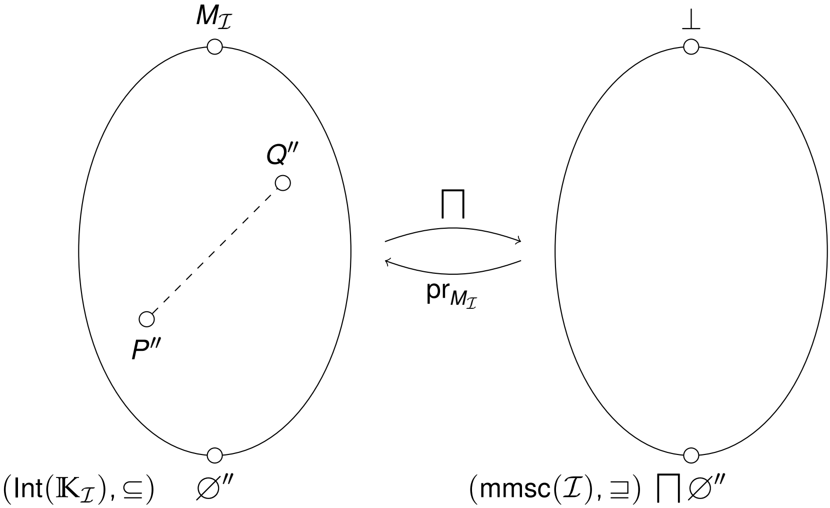


$(\mathrm{Int}(\mathbb{K}_{\mathcal{I}}), \subseteq)$

An Order Isomorphism

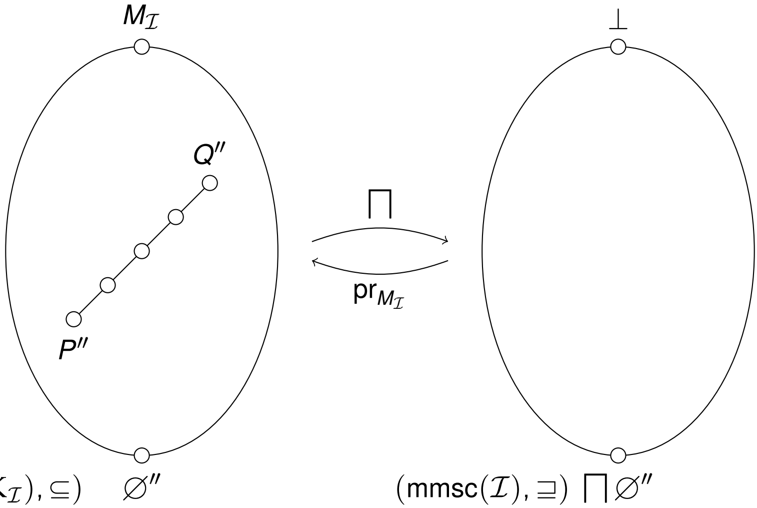


An Order Isomorphism



$\mathrm{conf}_{\mathbb{K}_{\mathcal{I}}}(P'' \longrightarrow Q'') \in [c, 1)$

An Order Isomorphism

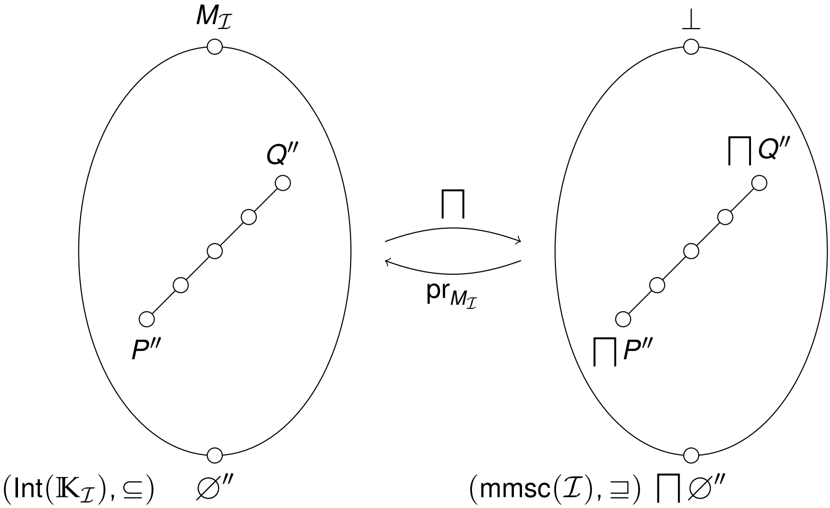


$(\mathrm{Int}(\mathbb{K}_{\mathcal{I}}), \subseteq)$

$(\mathrm{mmsc}(\mathcal{I}), \sqsupseteq)$

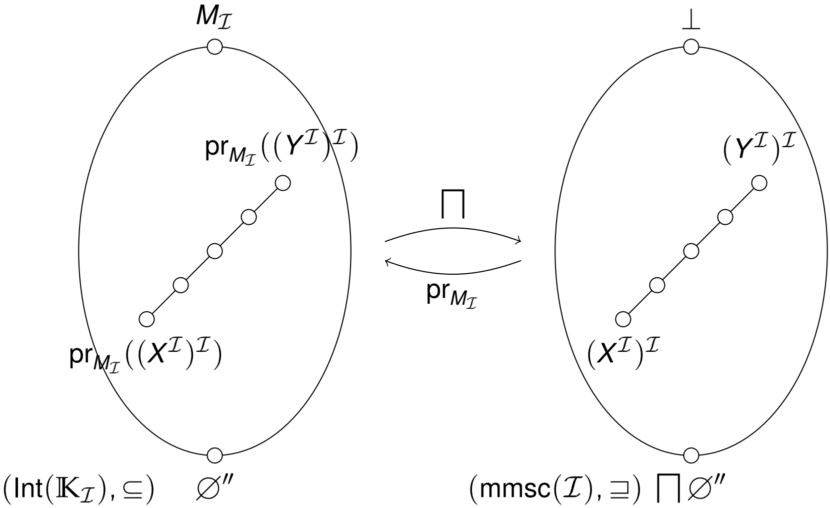
$\mathrm{conf}_{\mathbb{K}_{\mathcal{I}}}(P'' \longrightarrow Q'') \in [c, 1)$

An Order Isomorphism



$\mathrm{conf}_{\mathbb{K}_{\mathcal{I}}}(P'' \longrightarrow Q'') \in [c, 1)$

An Order Isomorphism



$\operatorname{conf}_{\mathcal{I}}((X^{\mathcal{I}})^{\mathcal{I}} \sqsubseteq (Y^{\mathcal{I}})^{\mathcal{I}}) \in [c, 1)$

A Base for Confident GCIs

Theorem (B. 2012)

Let $\mathcal{B}$ be a finite base of $\mathcal{I}$, $c \in [0, 1]$ and

$$\mathrm{Conf}(\mathcal{I}, c) := \{ X^{\mathcal{I}} \sqsubseteq Y^{\mathcal{I}} \mid Y \subseteq X \subseteq \Delta_{\mathcal{I}}, 1 > \mathrm{conf}_{\mathcal{I}}(X^{\mathcal{I}} \sqsubseteq Y^{\mathcal{I}}) \geqslant c \}.$$

Then $\mathcal{B} \cup \mathcal{C}$ is a finite base of $\mathrm{Th}_c(\mathcal{I})$.

A Base for Confident GCIs

Theorem (B. 2012)

Let $\mathcal{B}$ be a finite base of $\mathcal{I}$, $c \in [0, 1]$ and

$$\operatorname{Conf}(\mathcal{I}, c) := \{\, X^{\mathcal{I}} \sqsubseteq Y^{\mathcal{I}} \mid Y \subseteq X \subseteq \Delta_{\mathcal{I}}, 1 > \operatorname{conf}_{\mathcal{I}}(X^{\mathcal{I}} \sqsubseteq Y^{\mathcal{I}}) \geqslant c \,\}.$$

Then $\mathcal{B} \cup \mathcal{C}$ is a finite base of $\operatorname{Th}_c(\mathcal{I})$.

Theorem (B. 2012)

The set

$$\mathcal{D} := \{\, (X^{\mathcal{I}} \sqsubseteq Y^{\mathcal{I}}) \in \operatorname{Conf}(\mathcal{I}, c) \mid \nexists Z \subseteq \Delta_{\mathcal{I}} \colon Y^{\mathcal{I}} \subsetneq Z^{\mathcal{I}} \subsetneq X^{\mathcal{I}} \,\}$$

is complete for $\mathcal{C}$. In particular, $\mathcal{B} \cup \mathcal{D}$ is a finite base for $\operatorname{Th}_c(\mathcal{I})$.

Thank You for Your Attention!