AI-KI-B

Non-Classical Logics

Ute Schmid & Diedrich Wolter
Practice: Johannes Rabold

Cognitive Systems and Smart Environments
Applied Computer Science, University of Bamberg

last change: 27. Juni 2019, 11:54

Quiz!

1. Is $\exists x.\,(\forall y.\,(x < y))$ over the universe of reals. . .
   1. unsatisfiable, i.e., a contradiction;
   2. satisfiable;
   3. or a contradiction?
2. How about $\exists x.\,(\exists y.\,(P(x, y) \wedge \neg P(x, y))) \rightarrow \forall z.\,(P(z, z))$?
3. How can semantics logic formulae defined?

Agenda

Topics for today

- Motivation of developing new logics
- Fixed relation semantics: qualitative logics
- Example of modal logics: temporal logics, description logics
- Multi-valued validity: probabilistic logics, Bayesian network

Educational objectives: being able to. . .

- Understand and describe trade-offs between representations in different logics
- Describe models of modal and qualitative logics
- Implement reasoning with Bayesian networks

Utility of universal logic applicable to various domains advocated by Gottfried Wilhelm Leibniz.

> *"Quo facto, quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos Computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo ... dicere: calculemus" (Leipniz, 1684)*

FOL has proven to be a very useful logic, yet there are some shortcomings (cp. the five roles of a knowledge representation!):

Utility of universal logic applicable to various domains advocated by Gottfried Wilhelm Leibniz.

> *"Quo facto, quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos Computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo . . . dicere: calculemus" (Leipniz, 1684)*

FOL has proven to be a very useful logic, yet there are some shortcomings (cp. the five roles of a knowledge representation!):

1. Satisfiability problem is undecidable

Utility of universal logic applicable to various domains advocated by Gottfried Wilhelm Leibniz.

> *"Quo facto, quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos Computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo ... dicere: calculemus" (Leipniz, 1684)*

FOL has proven to be a very useful logic, yet there are some shortcomings (cp. the five roles of a knowledge representation!):

1. Satisfiability problem is undecidable
2. Axiomatisation difficult, impossible for equality

Motivation: Avoid challenging axiomatisation

- Modify semantics to have „smaller" (or „northOf", etc.) fixed to their **intended semantics**
- Built-in semantics may be computationally cheaper than expressive logic required for domain axiomatisation
- Key question: How can we reason within a logic that includes predicates with fixed semantics?
  - Example: $\mathrm{northOf}(X, Y) \wedge \mathrm{northof}(Y, X) \vdash \bot$ should be achievable for any X, Y.
  - Observation: Introducing pre-defined predicates requires resolution calculus to be extended
- To avoid undecidability of FOL, mostly restricted logics are considered

Questions of knowledge representation:

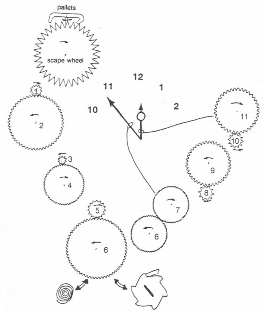
- What primitives are required to represent **common sense** understanding of, e.g., physical mechanisms?
- What are algorithmic properties of reasoning with them?
- To which extent can they represent real world behaviour?

Research area of qualitative representation:

- "First wave" qualitative reasoning (QR) sparked by Pat Hayes' naive physics manifesto (1978) promoting the study of restricted but in-depth theory of common-sense physics
- Since 1990ies specialisations on qualitative spatial and temporal reasoning
- Up to today, more than 40 families of spatial representations explored

> *"[...] the heart of the qualitative reasoning enterprise is to develop computational theories of the core skills underlying engineers, scientists, and just plain folks's ability to hypothesize, test, predict, create, optimize, diagnose and debug physical mechanisms. (Williams and de Kleer, 1991)*

Qualitative reasoning for understanding how a mechanical clock works (Forbus, 1991).



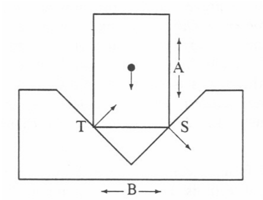
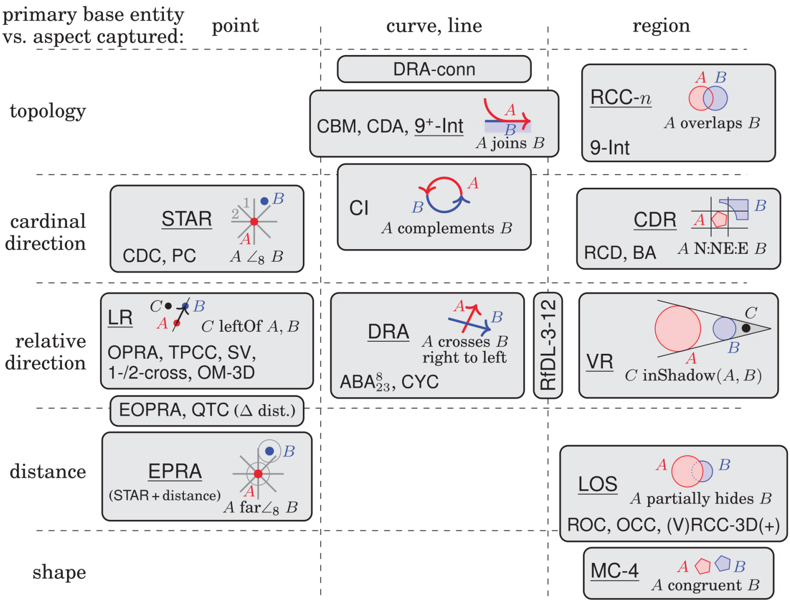


Illustration Qualitative Calculi



Qualitative Spatial Representation: RCC

We consider important family of representations, known as **Region Connection Calculus (RCC)** (Randell, Cui, Cohn 1992), as example.

Representation idea:

- Capture fundamental topological relations of spatial regions
- Build theory on single notion of *connection*

Application areas:

- Query languages for geographic information systems
- Robot navigation
- Language understanding
- Non-spatial concept models, ...

What is the underlying domain theory of *connection*?

Qualitative Spatial Representation: RCC

We consider important family of representations, known as **Region Connection Calculus (RCC)** (Randell, Cui, Cohn 1992), as example.

Representation idea:

- Capture fundamental topological relations of spatial regions
- Build theory on single notion of *connection*

Application areas:

- Query languages for geographic information systems
- Robot navigation
- Language understanding
- Non-spatial concept models, ...

What is the underlying domain theory of *connection*?

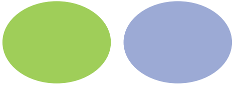
1. $\forall x. C(x, x)$
2. $\forall x. \forall y. (C(x, y) \rightarrow C(y, x))$

Using the primitive connection $C(x, y)$, interesting relations can be defined. We use $\phi := \psi$ as shorthand notation for $\forall x.(\forall y.((\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)))$

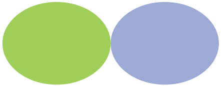
$\mathrm{DC}(x, y)$	$:=$	$\neg C(x, y)$	disconnected
$\mathrm{P}(x, y)$	$:=$	$\forall z.(C(z, x) \rightarrow C(z, y))$	part of
$\mathrm{PP}(x, y)$	$:=$	$P(x, y) \wedge \neg P(y, x)$	proper part of
$\mathrm{EQ}(x, y)$	$:=$	$\mathrm{P}(x, y) \wedge \mathrm{P}(y, x)$	equal
$\mathrm{O}(x, y)$	$:=$	$\exists z.(\mathrm{P}(z, x) \wedge \mathrm{P}(z, y))$	overlaps
$\mathrm{PO}(x, y)$	$:=$	$\mathrm{O}(X, Y) \wedge \neg \mathrm{P}(x, y) \wedge \neg \mathrm{P}(y, x)$	partially overlapping
$\mathrm{DR}(x, y)$	$:=$	$\neg \mathrm{O}(x, y)$	discrete
$\mathrm{EC}(x, y)$	$:=$	$C(x, y) \wedge \neg O(x, y)$	externally connected
$\mathrm{TPP}(x, y)$	$:=$	$\mathrm{PP}(x, y) \wedge \exists z.(\mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y))$	tangential proper part
$\mathrm{NTPP}(x, y)$	$:=$	$\mathrm{PP}(x, y) \wedge \neg \exists z.(\mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y))$	non-tangential proper part
$\mathrm{TPPi}(x, y)$	$:=$	$\mathrm{TPPi}(x, y)$	TPP inverse
$\mathrm{NTPPi}(x, y)$	$:=$	$\mathrm{NTPPi}(x, y)$	NTPP inverse

Illustration

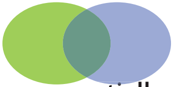
Set of RCC-8 relations defined by the axiomatisation:



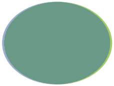
dc disconnected



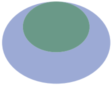
ec externally connected



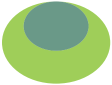
po partially overlapping



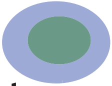
eq equal



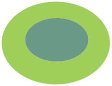
tpp tangential propoer part



tppi tangential proper part of



ntpp non-tangential proper part



ntppi non-tangetial proper part of

Question: What is the core cost of reasoning with RCC relations?

- Idea: consider existentially quantified conjunctive formulae only, e.g.,

$$\exists R_1.\exists R_2.\ldots.\exists R_n.(\mathrm{NTPP}(R_1, R_3) \wedge \mathrm{PO}(R_3, R_2) \wedge \ldots)$$

Question: What is the core cost of reasoning with RCC relations?

- Idea: consider existentially quantified conjunctive formulae only, e.g.,

$$\exists R_1.\exists R_2.\ldots.\exists R_n.(\mathrm{NTPP}(R_1, R_3) \wedge \mathrm{PO}(R_3, R_2) \wedge \ldots)$$

- Observation: These formulae correspond to CSPs!
  - finite set of variables $R_1, \ldots, R_n$
  - constraint relations from RCC, i.e., $\mathrm{NTPP}, \mathrm{EC}, \ldots$
  - domains are **infinite**, though!
- Infinite domains not covered by our definition of CSPs – deciding CSPs over infinite domains may thus be more difficult than NP-complete (one cannot guess!), possibly undecidable

Question: What is the core cost of reasoning with RCC relations?

- Idea: consider existentially quantified conjunctive formulae only, e.g.,

$$\exists R_1.\exists R_2.\ldots.\exists R_n.(\mathrm{NTPP}(R_1, R_3) \wedge \mathrm{PO}(R_3, R_2) \wedge \ldots)$$

- Observation: These formulae correspond to CSPs!
  - finite set of variables $R_1, \ldots, R_n$
  - constraint relations from RCC, i.e., $\mathrm{NTPP}, \mathrm{EC}, \ldots$
  - domains are **infinite**, though!
- Infinite domains not covered by our definition of CSPs – deciding CSPs over infinite domains may thus be more difficult than NP-complete (one cannot guess!), possibly undecidable
- (Surprising) result: CSPs can be solved in just $O(n^3)$ time!
- Reasoning is not based on FOL resolution using axiomatisation, but a dedicated method

It can be shown that for existentially quantified conjunctive formulae it suffices to check that for all variables R_i, R_j, R_k the sub-formulae

$$\alpha(R_i, R_j) \wedge \beta(R_j, R_i)$$

and

$$\gamma(R_i, R_j) \wedge \delta(R_j, R_k) \wedge \sigma(R_i, R_k)$$

are listed in a finite table of possible combinations.

- For example, $\alpha = \mathrm{EQ}$ implies that $\beta = \mathrm{EQ}$
- compare to initial example of dis-allowing $\mathrm{northOf}(x, y)$ and $\mathrm{northOf}(y, x)$ to hold simultaneously!

Up to today, RCC is one of the best-analysed qualitative logics

- Thorough analysis of above method by Renz (2002)
- Several variations of RCC formalism studied (turns out to be NP-complete if regions are restricted to be two-dimensional!)
- RCC formulae characterised as special **modal logic**
- Nevertheless, even up to today improvements for the core reasoning methods are proposed

Wrap-Up Qualitative Logics

- Restrictions of FOL (e.g., CSP) with a set of **qualitative relations** with fixed semantics: avoids explicit axiomatisation
- Develop reasoning algorithms capable of deciding satisfiability in these logics
- Various sets of relations, called **calculi**, studied
- Several calculi motivated by human cognition – bridging formal logics to human-like reasoning
- Several fundamental representation allow for efficient reasoning, yet some spatial relations can very hard to reason about (beyond NP)
- Important for natural language understanding, situated systems, robots, and many other areas

Representation idea:

- Decompose overall representation into distinct parts and specify how the parts are linked
- Extend propositional logic with ‘link’ operators, retain decidability – restricted FOL

Approach:

- Semantics based on interpretations from propositional logics
- **Modal operators** connect distinct interpretations

Application areas

- Representation of temporal knowledge, spatial knowledge, ..., of meta-knowledge (epistemic logic)
- Temporal logics and computational tree logic in software engineering restricted FOL

There exists a wide range of (multi-)modal logics – we only give an overview of the general idea.

Syntax

- Propositional logic formulae are well-formed modal logic formulae
- Additionally, two (or more) unary modal operators are introduced
- In case of temporal logic modal operators $\Box$ (always) and $\diamond$ (eventually)

Examples of well-formed formulae:

$$a \land b \lor c$$
$$\Box a$$
$$(\diamond(b \lor c)) \land (\Box d))$$

Idea:

- Semantics of propositional (sub-formulae) is defined by interpretation functions $I : \mathcal{F} \rightarrow \{\top, \bot\}$ like in standard propositional logic
- Modal operators (e.g., $\Box, \diamond$) switch from one interpretation to another
- Individual interpretations are called **worlds**, the set of all worlds is called **universe**
- Modal logics differ in how switching between worlds is organised – this can lead to a great variety of different logics

w_0:	w_1:	w_2:	w_3:	
$I_0(a) = \top$	$I_1(a) = \bot$	$I_2(a) = \bot$	$I_3(a) = \bot$	…

Linear Temporal Logic (LTL)

In **Linear Temporal Logic**, an infinite set of worlds is connected in a linear fashion:

$\mathcal{M}$

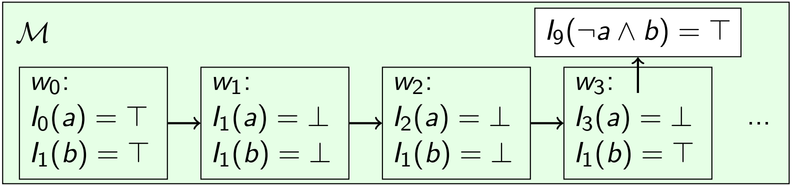
w_0:		w_1:		w_2:		w_3:	
$I_0(a) = \top$	$\longrightarrow$	$I_1(a) = \bot$	$\longrightarrow$	$I_2(a) = \bot$	$\longrightarrow$	$I_3(a) = \bot$	$\dots$
$I_1(b) = \top$		$I_1(b) = \bot$		$I_1(b) = \bot$		$I_1(b) = \top$	

Semantics

- $I(\phi) := I_0(\phi)$
- $I_k(\Box\phi) := \begin{cases} \top & \forall n.(n > k \rightarrow I_n(\phi) = \top) \\ \bot & \text{otherwise} \end{cases}$
- $I_k(\diamond\phi) := \begin{cases} \top & \exists n.(n > k \rightarrow I_n(\phi) = \top) \\ \bot & \text{otherwise} \end{cases}$

Note: Sometime, in particular in software engineering, LTL is used with a richer set of modal operators, e.g., next, until, while.

In **Linear Temporal Logic**, an infinite set of worlds is connected in a linear fashion:



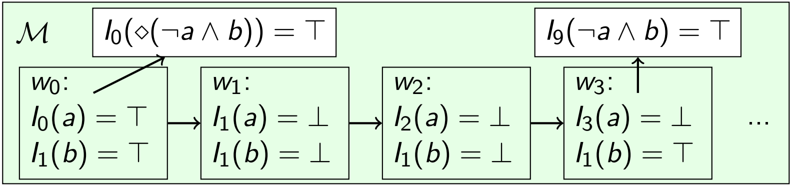
Semantics

- $I(\phi) := I_0(\phi)$
- $I_k(\Box\phi) := \begin{cases} \top & \forall n.(n > k \rightarrow I_n(\phi) = \top) \\ \bot & \text{otherwise} \end{cases}$
- $I_k(\Diamond\phi) := \begin{cases} \top & \exists n.(n > k \rightarrow I_n(\phi) = \top) \\ \bot & \text{otherwise} \end{cases}$

Note: Sometime, in particular in software engineering, LTL is used with a richer set of modal operators, e.g., next, until, while.

Linear Temporal Logic (LTL)

In **Linear Temporal Logic**, an infinite set of worlds is connected in a linear fashion:



Semantics

- $I(\phi) := I_0(\phi)$
- $I_k(\Box\phi) := \begin{cases} \top & \forall n.(n > k \rightarrow I_n(\phi) = \top) \\ \bot & \text{otherwise} \end{cases}$
- $I_k(\diamond\phi) := \begin{cases} \top & \exists n.(n > k \rightarrow I_n(\phi) = \top) \\ \bot & \text{otherwise} \end{cases}$

Note: Sometime, in particular in software engineering, LTL is used with a richer set of modal operators, e.g., next, until, while.

Representation idea:

- tailor logic to specification of concepts by their structure
- logic-based formalisation of semantic networks

Description Logic (DL) is also called a **terminological language** and they are mostly inspired by KL-ONE. An overview is presented at `http://www.dl.kr.org`

Approach:

- Define special **multi-modal** logic which connects concept specification with specifications of their parts

Application areas:

- Semantic web (Wikidata, linked science), databases
- Question answering

There are many distinct DLs that differ in which operators are provided and what kind of information can be represented. Several combinations of operators can lead to undecidable logics. DL research investigates trade-offs between different design decisons. Let three countable set of symbols be given:

- Concept names (convention: starts with a single uppercase letter)
- Role names (convention: starts with a lowercase letter)
- Individual names (convention: in uppercase letters)

DL, depending on the specific variant, allows for three types of definitions addressing **terminology (TBox)**, **roles (RBox)**, and **individuals (ABox)** (assertions).

- Well-formed formulae are defined for three box types individually

Roles are used to represent the relationship between structural elements.

Example: Fathers and daughters may be linked by **role** 'child' and its **inverse role** 'parent'

In the RBox, roles are defined. In the logic $\mathcal{ALC}$ only a single definition is used:

1. Any role name $\mathtt{r}$ represents a role.

In more expressive logics like $\mathcal{SROIQ}$ several more definitions are employed additionally, including

2. For any role name $\mathtt{r}$, $\mathtt{r}^-$ is a role too
3. For any role names $\mathtt{r}_1, \ldots \mathtt{r}_n$, $\mathtt{r}_1 \circ \ldots \circ \mathtt{r}_{n-1} \sqsubseteq \mathtt{r}_n$ is a role too (role chains, role inclusions)

Thus, in $\mathcal{SROIQ}$ one can represent $\mathtt{father} \sqsubseteq \mathtt{child}^-$, whereas in $\mathcal{ALC}$ `father` and `child` cannot be related.

In the TBox, **terminological structure** is described, sometimes referred to as implicit knowledge. Well-formed TBox formulae, i.e., concept definitions, are either of the kind $\mathtt{C} \equiv \mathtt{D}$ or $\mathtt{C} \sqsubseteq \mathtt{D}$ where $\mathtt{C}$ is a concept name and $\mathtt{D}$ is a concept definition defined as follows:

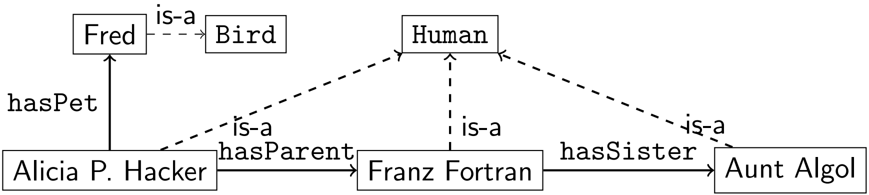
1. $\top, \bot$ are concept definitions
2. Every concept name $\mathtt{C}$ is a concept definition
3. If $\mathtt{C}, \mathtt{D}$ are concept definitions, so are $\neg\mathtt{C}$, $\mathtt{C} \sqcup \mathtt{D}$, and $\mathtt{C} \sqcap \mathtt{D}$
4. If $\mathtt{C}$ is a concept definition and $\mathtt{r}$ is a role definition, then $\forall\mathtt{r}.\mathtt{C}$ and $\exists\mathtt{r}.$Care concepts definitions

In more expressive DLs further constructors may be defined, for example

5. If $\mathtt{C}$ is a concept definition, $\mathtt{r}$ is a role definition, and $n \in \mathbb{N}$, then $(\geq n\mathtt{r}.\mathtt{C})$, $(\leq n\mathtt{r}.\mathtt{C})$ are concepts definitions
6. For individual names $\mathtt{A}_1, \ldots \mathtt{A}_n$, $\{\mathtt{A}_1, \ldots \mathtt{A}_n\}$ is a concept definition – these are called **nominals**

$$\begin{aligned}
\texttt{Parent} &\equiv \exists\texttt{hasChild}.\texttt{Human}\\
\texttt{ProudParent} &\sqsubseteq \exists\texttt{hasChild}.(\texttt{Doctor} \sqcap \exists\texttt{graduatedFrom}.\texttt{UniBamberg})\\
\texttt{Nerd} &\sqsubseteq \texttt{Student} \sqcap (\geq 5\ \texttt{hasObject}.\texttt{Computer})\\
\texttt{Nephew} &\equiv \texttt{Human} \sqcap \exists\texttt{hasParent}.((\exists\texttt{hasBrother}.\top) \sqcup (\exists\texttt{hasSister}.\top))
\end{aligned}$$

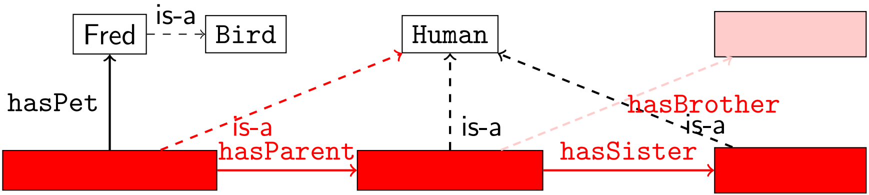
Illustration as semantic network:



Concept definitions are graph patterns in semantic networks!

$$
\begin{aligned}
\texttt{Parent} &\equiv \exists \texttt{hasChild}.\texttt{Human}\\
\texttt{ProudParent} &\sqsubseteq \exists \texttt{hasChild}.(\texttt{Doctor} \sqcap \exists \texttt{graduatedFrom}.\texttt{UniBamberg})\\
\texttt{Nerd} &\sqsubseteq \texttt{Student} \sqcap (\geq 5\ \texttt{hasObject}.\texttt{Computer})\\
\texttt{Nephew} &\equiv \texttt{Human} \sqcap \exists \texttt{hasParent}.((\exists \texttt{hasBrother}.\top) \sqcup (\exists \texttt{hasSister}.\top))
\end{aligned}
$$

Illustration as semantic network:



Concept definitions are graph patterns in semantic networks!

The ABox makes assertions about individuals, it states explicit knowledge.

1. Membership in classes: for a concept definition $\mathtt{C}$ and an individual's name $\mathtt{N}$ one can write $\mathtt{C}(N)$.
2. Role filling: for a role definition $\mathtt{r}$ and individuals' names $\mathtt{N}, \mathtt{M}$ one can write $\mathtt{r}(\mathtt{N}, \mathtt{M})$

Examples according to semantic network from previous slide:

- $\mathtt{Bird}(\mathtt{FRED})$
- $\mathtt{hasParent}(\mathtt{ALICIA\_P\_HACKER}, \mathtt{FRED\_FORTRAN})$
- $(\mathtt{Bird} \sqcup \mathtt{Human})(\mathtt{FRANZ\_FORTRAN})$

Semantics is intuitive:

- concepts represent sets of individuals
- complex DL concepts are defined using set expressions
- semantics given by **set theory**, e.g., membership, inclusion

Semantics is intuitive:

- concepts represent sets of individuals
- complex DL concepts are defined using set expressions
- semantics given by **set theory**, e.g., membership, inclusion

Formal definition is similar to FOL using universe (called **domain** D in DL) and **interpretation** function I

- **Unique Name Assumption** is assumed: distinct individuals are named differently
- every individual's name is mapped to an element $a \in D$
- every role name is mapped to a subset of $D \times D$
- every concept name is mapped to a subset of D

1. $I(\top) := D$
2. $I(\bot) := \emptyset$
3. $I(\neg C) := D \setminus I(C)$
4. $I(C \sqcap D) := I(C) \cap I(D)$
5. $I(C \sqcup D) := I(C) \cup I(D)$
6. $I(\forall r.C) := \{x \in D | \forall y.((x, y) \in I(r) \to y \in I(C))\}$
7. $I(\exists r.C) := \{x \in D | \exists y.((x, y) \in I(r) \wedge y \in I(C))\}$

For TBoxes:

1. I satisfies $C \sqsubseteq D$ iff $I(C) \subseteq I(D)$
2. I satisfies $C \equiv D$ iff $I(C) = I(D)$

For ABoxes:

1. I satisfies $C(a)$ iff $a \in I(C)$
2. I satisfies $r(a, b)$ iff $(I(a), I(b)) \in I(r)$

Typical logic reasoning problems can be applied to DLs such as satisfiability of TBoxes, but some application-specific reasoning tasks are commonly considered:

- Concept satisfiability wrt. a knowledge base K, i.e., does K provide a model for call concepts such that $I(C) \neq \emptyset$?
- Subsumption, i.e., identifying sub-concept relationships
- Instance checking, i.e., identifying whether $C(a)$ is satisfied model wrt. a knowledge base K.
- Answering concept queries, i.e., deciding the concept membership of a given individual wrt. given terminology

Satisfiability can be decided using **tableaux algorithms** (essentially a graph-based variant of last week's resolution), computational complexity for DLs is (mostly) decidable, yet at least PSPACE-complete.

So far, we have only considered **Boolean validity**

- $\top$ representing truth
- $\bot$ representing falsity

There several application of logical reasoning that would benefit from a more fine-grained distinction.

Example Consider **fault analysis**: a machine does not work as expected and one wants to reason what went wrong

$$\Phi_{\text{observation}} \not\models \Psi_{\text{normal behaviour}}, \Phi_{\text{observation}} \models \Psi_{\text{fault}_i}$$

Question: How can we model that some faults are more likely to occur than others?

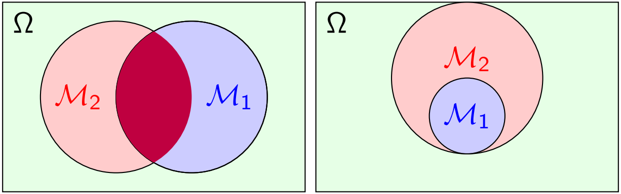
Idea: assign a probability of truth to each atom in a propositional logic, i.e., interpretation $I : A \rightarrow [0, 1]$

- We can then tackle the task of computing the most probable explanation.

Question: How can we define semantics for a propositional probabilistic logic?

Set-Based Semantics

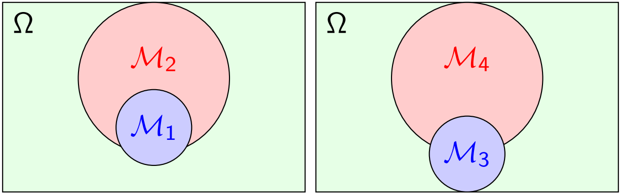
Basic idea is a generalisation of classic propositional logic, considering the set of models $\mathcal{M}_i$ of some formula Φ_i, i.e., $\mathcal{M}_i \models \Phi_i$. Two examples:



$\mathcal{M}_1 \cap \mathcal{M}_2 \models \Phi_1 \wedge \Phi_2$ $\qquad$ $\Phi_1 \models \Phi_2$

We can rewrite logic operations as set-theoretic operations: set algebras $\langle \cup, \cap, \cdot^C \rangle$ are Boolean algebras like $\langle \vee, \wedge, \neg \rangle$!

Consider the following situations, in classical logic we have $\Phi_{2n} \not\models \Phi_{2n+1}, n = \{0, 1\}$



Intuitively, we want $P(\Phi_1 \to \Phi_2) > P(\Phi_3 \to \Phi_4)$.

- Set of all interpretations Ω serves as probability space
- $P(\bot) = P(\emptyset) = 0, P(\top) = P(\Omega) = 1$

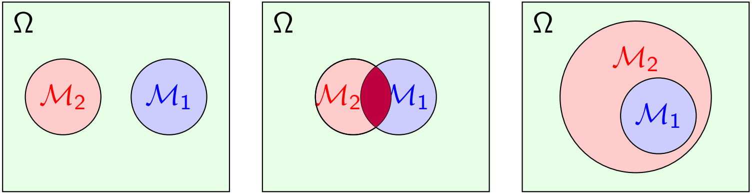
Let Ω be a probability space and $A, B \subseteq \Omega$ subsets, called events (models of formulae will be our events).

1. $P : 2^{\Omega} \to [0, 1]$
2. $P(\Omega) = 1$
3. $A \cap B = \emptyset \to P(A \cup B) = P(A) + P(B)$

We write **conditional probability** of event A, given B as $P(A|B)$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

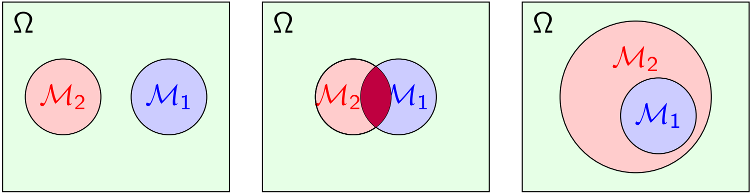
Assume we want to compute the probability of $\Phi_1 \to \Phi_2$ to hold, given models $\mathcal{M}_1$ of premisis Φ_1 and $\mathcal{M}_2$ of conclusion Φ_2.



never true	can hold	always true
0	?	1

Probabilistic Measure of Belief

Assume we want to compute the probability of $\Phi_1 \rightarrow \Phi_2$ to hold, given models $\mathcal{M}_1$ of premisis Φ_1 and $\mathcal{M}_2$ of conclusion Φ_2.



never true	can hold	always true
0	$\frac{P(\mathcal{M}_1 \cap \mathcal{M}_2)}{P(\mathcal{M}_1)}$	1

We can determine probability of logical consequence (entailment) as conditional probabilities!

$$P(\Phi_1 \rightarrow \Phi_2) = P(\mathcal{M}_2 | \ \mathcal{M}_1) \text{ with } \ \mathcal{M}_i \models \Phi_i$$

Problems?

Previous observation suggests we are done now.

- Given a probabilistic evaluation of atoms, i.e., $P(a)$,
- compute probability of composite formulae using set-theoretic semantics and basic probability theory

Problems?

Previous observation suggests we are done now.

- Given a probabilistic evaluation of atoms, i.e., $P(a)$,
- compute probability of composite formulae using set-theoretic semantics and basic probability theory

Unfortunately, this is were the problems just start!
Luckily, researcher *love* problems (discovering new ones, solving old ones)!

Feature of all classical logics is **Monotonicity**:

$$\mathcal{K} \vdash \Phi \Rightarrow \mathcal{K} \cup \{\Psi'\} \vdash \Phi$$

- If we obtain a piece of information Φ, any conclusions not using Φ are still valid
- The set of logical consequences grows monotonously when new knowledge is obtained

Monotonicity does no longer hold with conditional probability. It may easily happen that

$$P(\Phi|\Psi) > P(\Phi|\Psi \underbrace{\wedge \Psi'}_{\text{new fact}})$$

To understand probabilistic reasoning we have to understand **nonmonotonic reasoning**!

Nonmonotonic reasoning confronts us with many challenges

- Defining a semantic that does not give rise to counter-intuitive phenomema
- Developing algorithms that are able to revise previous conclusions, if new knowledge arrives: **truth maintenance systems**

Several AI researchers have worked on overcoming the dilemma, with a variety of approaches

- McCarthy's circumscription
- Reiter's default logic
- Moore's autoepistemic logic
- ...

These logics are beyond the scope of an introduction to AI.

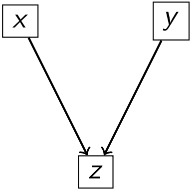
Coming back to the beginning of our KR sessions, first approaches like frames and semantic networks are build on **graphical models** which can be depicted as graphs.
With **Bayesian networks**, we can pursue a similar approach to probabilistic reasoning. Let $X = \{x_1, \ldots, x_n\}$ be a set of variables.

Definition

A Bayesian network over X is a directed acyclic graph (X, E) in which nodes represent variables X and their respective probability distribution can be written as

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P\left(x_i | \mathrm{Pre}(x_i)\right),$$

where $\mathrm{Pre}(x_i)$ denotes the set of predecessors (parents) of variable x_i.



$$P(x) = 0.1$$
$$P(y) = 0.2$$
$$P(z|x, y) = 0.9$$
$$P(z|x, \neg y) = 0.8$$
$$P(z|\neg x, y) = 0.6$$
$$P(z|\neg x, \neg y) = 0.2$$

From the above model we can compute

$$P(z) = 0.9 \cdot 0.1 \cdot 0.2 + 0.8 \cdot 0.1 \cdot 0.8 + 0.6 \cdot 0.9 \cdot 0.2 + 0.2 \cdot 0.9 \cdot 0.8 = 0.334$$

Observation: paths in the graph denote probabilistic dependent variables.

- A variable is conditionally independent of its non-successors (non-descendants), given its predecessors (parents).

Question: How can we compute the joint probability of all variables (efficiently)?

- A naive approach like shown in the example is to iteratively compute $P(A, B)$ from $P(A, B) = P(A|B) \cdot P(B)$ starting with all nodes that have no predecessor (parent).
- More efficient techniques employ a method of **maginalizing** to eliminate some variables from the computation.

Studying the advanced methods is beyond the scope of this introduction course.

Summary

- Besides propositional logic and first-order logic, several other approaches to logic reasoning exist
- Qualitative logics employ a set of predicates with fixed semantics; they define dedicated reasoning procedures
- Modal logics introduce a limited variant of quantification within structures of worlds, where each world is a unique interpretation for propositional sub-formulae. Logic reasoning is for most modal logics decidable, often not easier than PSPACE-complete. Reasoning algorithms are either resolution or tableaux calculi
- Description logics aim at representing structure of concepts. A great variety of DLs has been developed and some specialised reasoning-tasks are commonly considered in DL applications
- Probabilistic logics extend Boolean validity to probabilities, which leads to nonmonotonic reasoning
- Bayesian networks are graphical models to represent conditional (in)dependence among random variables.

Literature

- S. Russel & P. Norvig (2010, 3rd edition). Artificial Intelligence: A Modern Approach, Chapter 14 “Probabilistic Reasoning”
- F. Dylla et al. (2017). A Survey of Qualitative Spatial and Temporal Calculi – Algebraic and Computational Properties, ACM Computing Surveys (CSUR), 50(1), Article 7
- F. Baader et al. (2003). The description logic handbook: theory, implementation, and applications, Cambridge University Press New York, NY, USA

Further Readings

- P. Hayes (1978). The Naive Physics Manifesto, In Expert Systems in the Microelectronic Age, Edinburgh University Press, D. Michie (ed), pp. 242–270
- D. A. Randell and Z. Cui and A. G. Cohn (1992). A Spatial Logic Based on Regions and Connection, In Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference (KR'92)
- J. Renz (2002). Qualitative Spatial Reasoning with Topological Information, LNCS 2293, Springer-Verlag, Berlin