CS:4420 Artificial Intelligence

Spring 2017

First-Order Logic

Cesare Tinelli

The University of Iowa

Readings

- Chap. 8 of [Russell and Norvig, 2012]

Pros and cons of Propositional Logic

+ PL is declarative: pieces of syntax correspond to facts

+ PL allows partial/disjunctive/negated information
  (unlike most data structures and databases)

+ Propositional logic is compositional:
  meaning of $B_{1,1} \wedge P_{1,2}$ is derived from meaning of $B_{1,1}$ and of $P_{1,2}$

+ Meaning in propositional logic is context-independent
  (unlike natural language, where meaning depends on context)

- Propositional logic has very limited expressive power (unlike natural language)
  E.g., cannot say "pits cause breezes in adjacent squares" except by writing one sentence for each square

First-order logic

Whereas propositional logic assumes world contains facts,
first-order logic (like natural language) assumes the world contains

- Objects: people, houses, numbers, theories, Ronald McDonald, colors, baseball games, wars, centuries . . .
- Relations: red, round, bogus, prime, multistoried . . .,
brother of, bigger than, inside, part of, has color, occurred after, owns, comes between, . . .
- Functions: father of, best friend, third inning of, one more than, end of . . .

Syntax of FOL: Basic elements

Constants	$KingJohn,\ 2,\ UCB, \ldots$
Predicates	$Brother,\ >, \ldots$
Functions	$Sqrt,\ LeftLegOf, \ldots$
Variables	$x,\ y,\ a,\ b, \ldots$
Connectives	$\land \ \lor \ \neg \ \Rightarrow \ \ \Leftrightarrow$
Equality	$=$
Quantifiers	$\forall \ \exists$

Atomic sentences

Atomic sentence $=$ $predicate(term_1, \ldots, term_n)$
or $term_1 = term_2$

Term $=$ $function(term_1, \ldots, term_n)$
or $constant$ or $variable$

E.g., $Brother(KingJohn, RichardTheLionheart)$
$> (Length(LeftLegOf(Richard)), Length(LeftLegOf(KingJo$

Complex sentences

Complex sentences are made from atomic sentences using connectives

$$\neg S, \quad S_1 \wedge S_2, \quad S_1 \vee S_2, \quad S_1 \Rightarrow S_2, \quad S_1 \Leftrightarrow S_2$$

E.g. $Sibling(KingJohn, Richard) \Rightarrow Sibling(Richard, KingJohn)$
$>(1,2) \vee \leq(1,2)$
$>(1,2) \wedge \neg{>}(1,2)$

Language of FOL: Grammar

Sentence	::=	AtomicS $\mid$ ComplexS
AtomicS	::=	$\mathbf{True} \mid \mathbf{False} \mid$ RelationSymb(Term, ...) $\mid$ Term $=$ Term
ComplexS	::=	(Sentence) $\mid$ Sentence Connective Sentence $\mid$ $\neg$Sentence
	$\mid$	Quantifier Sentence
Term	::=	FunctionSymb(Term, ...) $\mid$ ConstantSymb $\mid$ Variable
Connective	::=	$\wedge \mid \vee \mid \Rightarrow \mid \leftrightarrow$
Quantifier	::=	$\forall$ Variable $\mid$ $\exists$ Variable
Variable	::=	$a \mid b \mid \cdots \mid x \mid y \mid \cdots$
ConstantSymb	::=	$A \mid B \mid \cdots \mid John \mid 0 \mid 1 \mid \cdots \mid \pi \mid \ldots$
FunctionSymb	::=	$F \mid G \mid \cdots \mid Cosine \mid Height \mid FatherOf \mid + \mid \ldots$
RelationSymb	::=	$P \mid Q \mid \cdots \mid Red \mid Brother \mid Apple \mid > \mid \cdots$

Truth in first-order logic

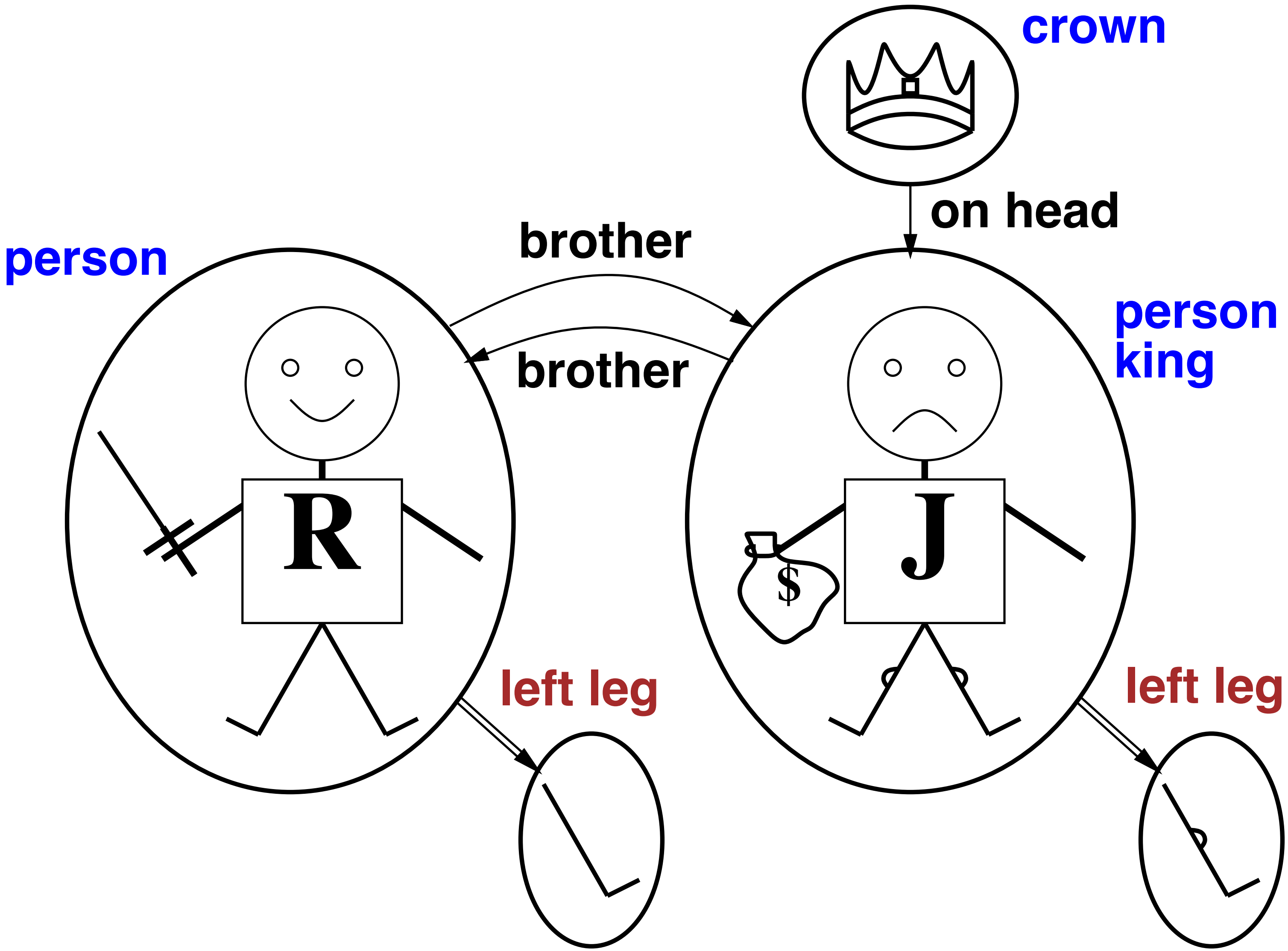
Sentences are true with respect to a model and an interpretation
Model contains ≥ 1 objects (domain elements) and relations among them
Interpretation specifies referents for

constant symbols $\rightarrow$ objects

predicate symbols $\rightarrow$ relations

function symbols $\rightarrow$ functional relations
An atomic sentence $predicate(term_1, \ldots, term_n)$ is true
iff the objects referred to by $term_1, \ldots, term_n$
are in the relation referred to by $predicate$

Models for FOL: Example



Truth example

Consider the interpretation in which

$Richard \rightarrow$ Richard the Lionheart
$John \rightarrow$ the evil King John
$Brother \rightarrow$ the brotherhood relation

Under this interpretation, $Brother(Richard, John)$ is true just in case Richard the Lionheart and the evil King John are in the brotherhood relation in the model

Semantics of First-Order Logic

(A little) more formally:

An **interpretation** is a pair $(\mathcal{D}, \sigma)$ where

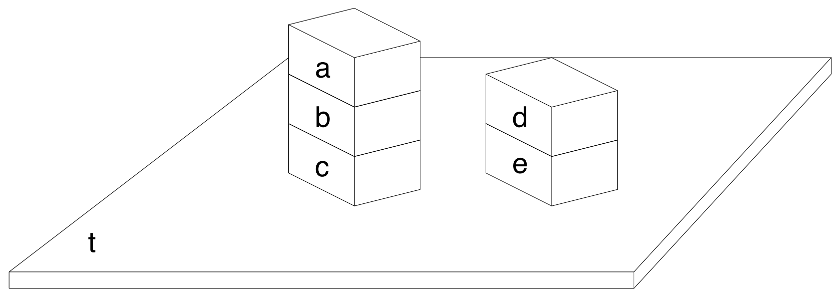
- $\mathcal{D}$ is a set of objects, the universe (or *domain*);
- σ is mapping from variables to objects in $\mathcal{D}$;
- $C^{\mathcal{D}}$ is an object in $\mathcal{D}$ for every constant symbol C;
- $F^{\mathcal{D}}$ is a function from $\mathcal{D}^n$ to $\mathcal{D}$ for every function symbol F of arity n;
- $R^{\mathcal{D}}$ is a relation over $\mathcal{D}^n$ for every relation symbol R of arity n;

An Interpretation I in the Blocks World

Constant Symbols: A, B, C, D, E, T

Function Symbols: $Support$

Relation Symbols: $On, Above, Clear$



$$A^H = A, B^H = B, C^H = C, D^H = D, E^H = E, T^H = T$$

$$
\begin{aligned}
Support^H &= \{\langle A, T\rangle, \langle B, A\rangle, \langle C, B\rangle, \langle D, C\rangle, \langle E, D\rangle\} \\
On^H &= \{\langle A, T\rangle, \langle B, A\rangle, \langle C, B\rangle, \langle D, C\rangle, \langle E, D\rangle\} \\
Above^H &= \{\langle E, D\rangle, \langle D, C\rangle, \ldots\} \\
Clear^H &= \{\langle E\rangle\}
\end{aligned}
$$

Semantics of First-Order Logic

Let $(\mathcal{D}, \sigma)$ be an interpretation and E an expression of FOL. We write $[\![E]\!]^{\mathcal{D}}_{\sigma}$ to denote the *meaning of E in the domain $\mathcal{D}$ under the variable assignment σ*.

The meaning $[\![t]\!]^{\mathcal{D}}_{\sigma}$ of a term t is an object of $\mathcal{D}$. It is inductively defined as follows.

$[\![x]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$\sigma(x)$	for all variables x
$[\![C]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$C^{\mathcal{D}}$	for all constant symbols C
$[\![F(t_1, \ldots, t_n)]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$F^{\mathcal{D}}([\![t_1]\!]^{\mathcal{D}}_{\sigma}, \ldots, [\![t_n]\!]^{\mathcal{D}}_{\sigma})$	for all function symbols F of arity n

Example

Consider the symbols $MotherOf, SchoolOf, Bill$ and the interpretation $(\mathcal{D}, \sigma)$ where

$$
\begin{aligned}
MotherOf^{\mathcal{D}} &\quad \text{is a unary fn mapping people to their mother} \\
FchildOf^{\mathcal{D}} &\quad \text{is a binary fn mapping a couple to their first child} \\
\sigma &\quad := \{x \mapsto \text{ George W Bush}, \ y \mapsto \text{ Barbara Bush}\}
\end{aligned}
$$

What is the meaning of $\boxed{MotherOf(x)}$ according to $(\mathcal{D}, \sigma)$?

$$[\![MotherOf(x)]\!]^{\mathcal{D}}_{\sigma} = [\![MotherOf]\!]^{\mathcal{D}}_{\sigma}([\![x]\!]^{\mathcal{D}}_{\sigma}) = MotherOf^{\mathcal{D}}(\sigma(x)) = \text{ Barbara Bush}$$

Semantics of First-Order Logic

The meaning $[\![\varphi]\!]^{\mathcal{D}}_{\sigma}$ of a formula φ is either $True$ or $False$.

It is inductively defined as follows.

$[\![t_1 = t_2]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$True$	iff	$[\![t_1]\!]^{\mathcal{D}}_{\sigma}$ is the same as $[\![t_2]\!]^{\mathcal{D}}_{\sigma}$
$[\![R(t_1, \ldots, t_n)]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$True$	iff	$\langle [\![t_1]\!]^{\mathcal{D}}_{\sigma}, \ldots, [\![t_n]\!]^{\mathcal{D}}_{\sigma} \rangle \in R^{\mathcal{D}}$
$[\![\neg\varphi]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$True/False$	iff	$[\![\varphi]\!]^{\mathcal{D}}_{\sigma} = False/True$
$[\![\varphi_1 \vee \varphi_2]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$True$	iff	$[\![\varphi_1]\!]^{\mathcal{D}}_{\sigma} = True$ or $[\![\varphi_2]\!]^{\mathcal{D}}_{\sigma} = True$
$[\![\exists x\, \varphi]\!]^{\mathcal{D}}_{\sigma}$	$:=$	$True$	iff	$[\![\varphi]\!]^{\mathcal{D}}_{\sigma'} = True$ for some σ' the same as σ except for x

Semantics of First-Order Logic

The meaning of formulas built with the other logical symbols can be defined by reduction to the previous symbols.

$$
\begin{array}{lcl}
[\![\varphi_1 \wedge \varphi_2]\!]^{\mathcal{D}}_{\sigma} & := & [\![\neg(\neg\varphi_1 \vee \neg\varphi_2)]\!]^{\mathcal{D}}_{\sigma} \\
[\![\varphi_1 \Rightarrow \varphi_2]\!]^{\mathcal{D}}_{\sigma} & := & [\![\neg\varphi_1 \vee \varphi_2]\!]^{\mathcal{D}}_{\sigma} \\
[\![\varphi_1 \leftrightarrow \varphi_2]\!]^{\mathcal{D}}_{\sigma} & := & [\![(\varphi_1 \Rightarrow \varphi_2) \wedge (\varphi_2 \Rightarrow \varphi_1)]\!]^{\mathcal{D}}_{\sigma} \\
[\![\forall x \varphi]\!]^{\mathcal{D}}_{\sigma} & := & [\![\neg\exists x\ \neg\varphi]\!]^{\mathcal{D}}_{\sigma}
\end{array}
$$

If a sentence is closed (no free variables), its meaning *does not depend* on the the variable assignment (although it may depend on the domain):

$$[\![\forall x\ \exists y\ R(x,y)]\!]^{\mathcal{D}}_{\sigma} = [\![\forall x\ \exists y\ R(x,y)]\!]^{\mathcal{D}}_{\sigma'} \quad \text{for any} \quad \sigma, \sigma'$$

Models, Validity, etc. for Sentences

An interpretation $(\mathcal{D}, \sigma)$ **satisfies** a sentence φ, or is a **model** for φ, if $[\![\varphi]\!]^{\mathcal{D}}_{\sigma} = True$.

A sentence is **satisfiable** if it has at least one model.

Examples: $\forall x\, x \geq y, \quad P(x)$

A sentence is **unsatisfiable** *if it has no models.*

Examples: $P(x) \wedge \neg P(x), \quad \neg(x = x)$

A sentence φ is **valid** if every interpretation is a model for φ.

Examples: $P(x) \Rightarrow P(x), \quad x = x$

φ *is valid/unsatisfiable iff* $\neg\varphi$ *is unsatisfiable/valid.*

Models, Validity, etc. for Sets of Sentences

An interpretation $(\mathcal{D}, \sigma)$ **satisfies** a set Γ of sentences, or is a **model** for Γ, if it is a model for *every* sentence in Γ.

A set Γ of sentences is **satisfiable** if it has at least one model.

Ex: $\{\forall x\, x \geq 0,\ \forall x\, x + 1 > x\}$

Γ is **unsatisfiable**, or **inconsistent**, if it has no models.

Ex: $\{P(x),\ \neg P(x)\}$

As in Propositional Logic, Γ **entails** a sentence φ ($\Gamma \models \varphi$), if every model of Γ is also a model of φ.

Ex: $\{\forall x\, P(x) \Rightarrow Q(x),\ P(A_{10})\} \models Q(A_{10})$

Note: Again, $\Gamma \models \varphi$ iff $\Gamma \wedge \neg\varphi$ is unsatisfiable.

Possible Interpretations Semantics

Sentences can be seen as *constraints* on the set S of all possible interpretations.

A sentence *denotes* all the possible interpretations that satisfy it (the models of φ):
If φ_1 denotes a set of interpretations S_1 and φ_2 denotes a set S_2, then

- $\varphi_1 \vee \varphi_2$ denotes $S_1 \cup S_2$,
- $\varphi_1 \wedge \varphi_2$ denotes $S_1 \cap S_2$,
- $\neg\varphi_1$ denotes $S \setminus S_1$,
- $\varphi_1 \models \varphi_2$ iff $S_1 \subseteq S_2$.

A sentence denotes either no interpretations or an infinite number of them!

Valid sentences do not tell us anything about the world. They are satisfied by every possible interpretation!

Models for FOL: Lots!

We *can* enumerate the models for a given FOL sentence:

For each number of universe elements n from 1 to ∞
    For each k-ary predicate P_k in the sentence
        For each possible k-ary relation on n objects
            For each constant symbol C in the sentence
                For each one of n objects mapped to C
                ...

Enumerating models is not going to be easy!

Universal quantification

$\forall \langle variables \rangle \;\; \langle sentence \rangle$

Everyone at Berkeley is smart:
$\forall x \;\; At(x, Berkeley) \Rightarrow Smart(x)$

$\forall x \;\; P$ is true in a model m iff P is true with x being each possible object in the model

Roughly speaking, equivalent to the conjunction of instantiations of P

$$
\begin{array}{ll}
 & (At(KingJohn, Berkeley) \Rightarrow Smart(KingJohn)) \\
\wedge & (At(Richard, Berkeley) \Rightarrow Smart(Richard)) \\
\wedge & (At(Berkeley, Berkeley) \Rightarrow Smart(Berkeley)) \\
\wedge & \ldots
\end{array}
$$

A common mistake to avoid

Typically, $\Rightarrow$ is the main connective with $\forall$
Common mistake: using $\land$ as the main connective with $\forall$:

$$\forall\, x \;\; At(x, Berkeley) \land Smart(x)$$

means "Everyone is at Berkeley and everyone is smart"

Existential quantification

$\exists \langle variables \rangle \ \langle sentence \rangle$

Someone at Stanford is smart:
$\exists x \ \ At(x, Stanford) \wedge Smart(x)$

$\exists x \ \ P$ is true in a model m iff P is true with x being some possible object in the model

Roughly speaking, equivalent to the disjunction of instantiations of P

$$
\begin{array}{ll}
 & (At(KingJohn, Stanford) \wedge Smart(KingJohn)) \\
\vee & (At(Richard, Stanford) \wedge Smart(Richard)) \\
\vee & (At(Stanford, Stanford) \wedge Smart(Stanford)) \\
\vee & \ldots
\end{array}
$$

Another common mistake to avoid

Typically, $\wedge$ is the main connective with $\exists$

Common mistake: using $\Rightarrow$ as the main connective with $\exists$:

$$\exists x \;\; At(x, Stanford) \Rightarrow Smart(x)$$

is true if there is anyone who is not at Stanford!

Properties of quantifiers

$\forall x \; \forall y$ is the same as $\forall y \; \forall x$ (why?)

$\exists x \; \exists y$ is the same as $\exists y \; \exists x$ (why?)

$\exists x \; \forall y$ is not the same as $\forall y \; \exists x$

$\exists x \; \forall y \; Loves(x, y)$
"There is a person who loves everyone in the world"

$\forall y \; \exists x \; Loves(x, y)$
"Everyone in the world is loved by at least one person"

Quantifier duality: each can be expressed using the other

$\forall x \; Likes(x, IceCream) \qquad \neg\exists x \; \neg Likes(x, IceCream)$

$\exists x \; Likes(x, Broccoli) \qquad \neg\forall x \; \neg Likes(x, Broccoli)$

Fun with sentences

Brothers are siblings

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brother(x, y) \Rightarrow Sibling(x, y)$.

"Sibling" is symmetric

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brother(x, y) \Rightarrow Sibling(x, y)$.

"Sibling" is symmetric

$\forall\, x, y \;\; Sibling(x, y) \;\Leftrightarrow\; Sibling(y, x)$.

One's mother is one's female parent

Fun with sentences

Brothers are siblings

$\forall\, x, y\;\; Brother(x, y) \Rightarrow Sibling(x, y)$.

"Sibling" is symmetric

$\forall\, x, y\;\; Sibling(x, y) \;\Leftrightarrow\; Sibling(y, x)$.

One's mother is one's female parent

$\forall\, x, y\;\; Mother(x, y) \;\Leftrightarrow\; (Female(x) \wedge Parent(x, y))$.

A first cousin is a child of a parent's sibling

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brother(x, y) \Rightarrow Sibling(x, y)$.

"Sibling" is symmetric

$\forall\, x, y \;\; Sibling(x, y) \;\Leftrightarrow\; Sibling(y, x)$.

One's mother is one's female parent

$\forall\, x, y \;\; Mother(x, y) \;\Leftrightarrow\; (Female(x) \wedge Parent(x, y))$.

A first cousin is a child of a parent's sibling

$\forall\, x, y \;\; FirstCousin(x, y) \;\Leftrightarrow\; \exists\, p, ps \;\; Parent(p, x) \wedge Sibling(ps, p) \wedge Parent(ps, y)$

Equality

$term_1 = term_2$ is true under a given interpretation
if and only if $term_1$ and $term_2$ refer to the same object

E.g., $1 = 2$ and $\forall x\ \ \times(Sqrt(x), Sqrt(x)) = x$ are satisfiable
$2 = 2$ is valid

E.g., definition of (full) $Sibling$ in terms of $Parent$:

$\forall x, y\ \ Sibling(x, y)\ \Leftrightarrow\ [\neg(x = y) \wedge \exists\, m, f\ \ \neg(m = f) \wedge$
$Parent(m, x) \wedge Parent(f, x) \wedge Parent(m, y) \wedge Parent(f, y)]$