CS:4420 Artificial Intelligence

Spring 2019

First-Order Logic

Cesare Tinelli

The University of Iowa

Readings

- Chap. 8 of [Russell and Norvig, 3rd Edition]

Knowledge Representation and Logic

Recall:

The field of Mathematical Logic provides powerful, formal knowledge representation languages and inference systems to build reasoning agents

We will consider two languages, and associated inference systems, from mathematical logic:

- Propositional Logic
- First-order Logic

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, 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

Constant symbols	$KingJohn,\ 2,\ Potus,\ [],\ \ldots$
Relation symbols	$Brothers(\_,\_),\ \_ > \_,\ Red(\_),\ \ldots$
Function symbols	$Sqrt(\_),\ LeftLegOf(\_),\ \_ + \_,\ \ldots$
Variables	$x,\ y,\ a,\ b, \ldots$
Connectives	$\wedge\ \vee\ \neg\ \Rightarrow\ \ \Leftrightarrow$
Equality	$=$
Quantifiers	$\forall\ \exists$

Atomic sentences

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

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

E.g., $Brother(KingJohn, RichardTheLionheart)$,

$Length(LeftLegOf(RobinHood)) > Length(LeftLegOf(KingJohn)))$

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. $Siblings(KingJohn, Richard) \Rightarrow Siblings(Richard, KingJohn)$

$x > 2 \ \vee \ 1 < x$

$1 > 2 \ \wedge \ \neg y > 2$

Language of FOL: Grammar

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

Truth in FOL

Sentences are true with respect to a model and an interpretation

A model contains ≥ 1 objects (domain elements) and relations and functions over them them

An interpretation specifies referents for

variables $\rightarrow$ objects

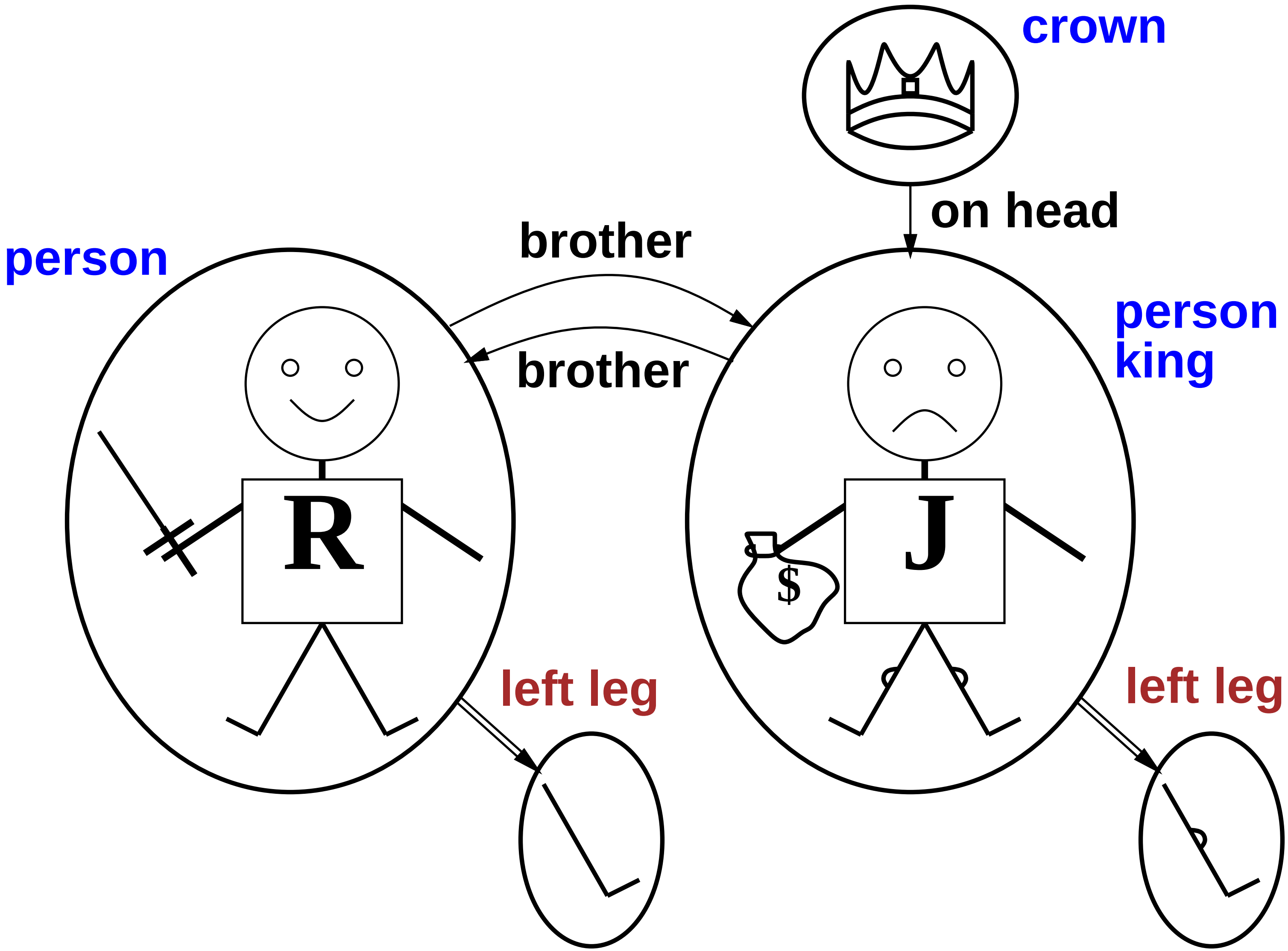
constant symbols $\rightarrow$ objects

predicate symbols $\rightarrow$ relations

function symbols $\rightarrow$ functional relations

An atomic sentence $P(t_1, \ldots, t_n)$ is true in an interpretation iff the objects referred to by $t_1, \ldots, t_n$ are in the relation referred to by P

Models for FOL: Example



Truth example

Consider the interpretation in which

$Richard$ → Richard the Lionheart
$John$ → the evil King John
$Brother$ → 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* $\mathcal{I}$ is a pair $(\mathcal{D}, \sigma)$ where

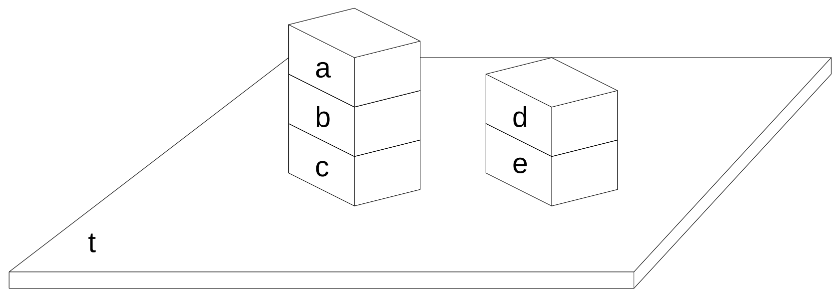
- $\mathcal{D}$ is a set of objects, the universe (or *domain*)
- σ is mapping from variables to objects in $\mathcal{D}$
- $c^{\mathcal{I}}$ is an object in $\mathcal{D}$ for every constant symbol c
- $f^{\mathcal{I}}$ is a function from $\mathcal{D}^n$ to $\mathcal{D}$ for every function symbol f of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

An Interpretation $\mathcal{I}$ in the Blocks World

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

Function Symbols: $Support$

Relation Symbols: $On, Above, Clear$



$A^{\mathcal{I}} = \mathsf{a},\ B^{\mathcal{I}} = \mathsf{b},\ C^{\mathcal{I}} = \mathsf{c},\ D^{\mathcal{I}} = \mathsf{d},\ E^{\mathcal{I}} = \mathsf{e},\ T^{\mathcal{I}} = \mathsf{t}$

$$
\begin{aligned}
Support^{\mathcal{I}} &= \{\langle \mathsf{a}, \mathsf{b}\rangle, \langle \mathsf{b}, \mathsf{c}\rangle, \langle \mathsf{c}, \mathsf{t}\rangle, \langle \mathsf{d}, \mathsf{e}\rangle, \langle \mathsf{e}, \mathsf{t}\rangle, \langle \mathsf{t}, \mathsf{t}\rangle\} \\
On^{\mathcal{I}} &= \{\langle \mathsf{a}, \mathsf{b}\rangle, \langle \mathsf{b}, \mathsf{c}\rangle, \langle \mathsf{c}, \mathsf{t}\rangle, \langle \mathsf{d}, \mathsf{e}\rangle, \langle \mathsf{e}, \mathsf{t}\rangle\} \\
Above^{\mathcal{I}} &= \{\langle \mathsf{a}, \mathsf{b}\rangle, \langle \mathsf{a}, \mathsf{c}\rangle, \langle \mathsf{a}, \mathsf{t}\rangle, \ldots\} \\
Clear^{\mathcal{I}} &= \{\langle \mathsf{a}\rangle, \langle \mathsf{d}\rangle\}
\end{aligned}
$$

Semantics of First-Order Logic

Let $\mathcal{I} = (\mathcal{D}, \sigma)$ be an interpretation and E an expression of FOL

We write $[\![e]\!]^{\mathcal{I}}$ to denote the *meaning of e in $\mathcal{I}$*

The meaning $[\![t]\!]^{\mathcal{I}}$ of a term t is an object of $\mathcal{D}$, inductively defined as follows:

$[\![x]\!]^{\mathcal{I}}$	$:=$	$\sigma(x)$	for all variables x
$[\![c]\!]^{\mathcal{I}}$	$:=$	$c^{\mathcal{I}}$	for all constant symbols c
$[\![f(t_1, \ldots, t_n)]\!]^{\mathcal{I}}$	$:=$	$f^{\mathcal{I}}([\![t_1]\!]^{\mathcal{I}}, \ldots, [\![t_n]\!]^{\mathcal{I}})$	for all n-ary function symbols f

Example

Consider the symbols $\mathit{MotherOf}$, $\mathit{SpouseOf}$ and the interpretation $\mathcal{I} = (\mathcal{D}, \sigma)$ where

$$\begin{array}{rl} \mathit{MotherOf}^{\mathcal{I}} & \text{is a unary fn mapping people to their mother} \\ \mathit{SpouseOf}^{\mathcal{I}} & \text{is a unary fn mapping people to their spouse} \\ \sigma & := \{x \mapsto \text{Bart},\ y \mapsto \text{Homer}, \ldots\} \end{array}$$

What is the meaning of $\mathit{SpouseOf}(\mathit{MotherOf}(x))$ in $\mathcal{I}$?

$$\begin{array}{rcl} [\![\mathit{SpouseOf}(\mathit{MotherOf}(x))]\!]^{\mathcal{I}} & = & \mathit{SpouseOf}^{\mathcal{I}}([\![\mathit{MotherOf}(x)]\!]^{\mathcal{I}}) \\ & = & \mathit{SpouseOf}^{\mathcal{I}}(\mathit{MotherOf}^{\mathcal{I}}([\![x]\!]^{\mathcal{I}})) \\ & = & \mathit{SpouseOf}^{\mathcal{I}}(\mathit{MotherOf}^{\mathcal{I}}(\sigma(x))) \\ & = & \mathit{SpouseOf}^{\mathcal{I}}(\mathit{MotherOf}^{\mathcal{I}}(\text{Bart})) \\ & = & \mathit{SpouseOf}^{\mathcal{I}}(\text{Marge}) \\ & = & \text{Homer} \end{array}$$

Semantics of First-Order Logic

Let $\mathcal{I} = (\mathcal{D}, \sigma)$ be an interpretation

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

It is inductively defined as follows:

$[\![t_1 = t_2]\!]^{\mathcal{I}}$	$:=$	$True$	iff	$[\![t_1]\!]^{\mathcal{I}}$ is the same as $[\![t_2]\!]^{\mathcal{I}}$
$[\![r(t_1, \ldots, t_n)]\!]^{\mathcal{I}}$	$:=$	$True$	iff	$\langle [\![t_1]\!]^{\mathcal{I}}, \ldots, [\![t_n]\!]^{\mathcal{I}} \rangle \in r^{\mathcal{I}}$
$[\![\neg\varphi]\!]^{\mathcal{I}}$	$:=$	$True/False$	iff	$[\![\varphi]\!]^{\mathcal{I}} = False/True$
$[\![\varphi_1 \vee \varphi_2]\!]^{\mathcal{I}}$	$:=$	$True$	iff	$[\![\varphi_1]\!]^{\mathcal{I}} = True$ or $[\![\varphi_2]\!]^{\mathcal{I}} = True$
$[\![\exists x\, \varphi]\!]^{\mathcal{I}}$	$:=$	$True$	iff	$[\![\varphi]\!]^{\mathcal{I}}_{\sigma'} = True$ for some σ' that disagrees with σ at most on x

Semantics of First-Order Logic

Let $\mathcal{I} = (\mathcal{D}, \sigma)$ be an interpretation

The meaning of formulas built with the other logical symbols:

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

If a sentence is *closed*, i.e., it has no *free* variables, its meaning does not depend on the the variable assignment—although it may depend on the domain:

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

Models, Validity, etc. for Sentences

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

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

Ex: $\forall x\, x \geq y$, $P(x)$

A sentence is unsatisfiable if it has no models

Ex: $P(x) \wedge \neg P(x)$, $\neg(x = x)$, $(\forall x\, Q(x, y)) \Rightarrow \neg Q(a, b)$

A sentence φ is *valid* if every interpretation is a model of it

Ex: $P(x) \Rightarrow P(x)$, $x = x$, $(\forall x\, P(x)) \Rightarrow \exists x\, P(x)$

Note: *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 of* 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: $\{P(x),\ \neg P(x)\}$

Γ *entails* a sentence φ ($\Gamma \models \varphi$), if every model for Γ is also a model for φ

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

Note: As in propositional logic, $\Gamma \models \varphi$ 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 $S_1 \subseteq S_2$.

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 $S_1 \subseteq S_2$.

Note 1: A sentence denotes either no interpretations or an infinite number of them!

Note 2: 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 an interpretation $\mathcal{I}$ iff P is true with x being each possible object in $\mathcal{I}$'s domain

Roughly speaking, equivalent to the conjunction of instantiations of P

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

Existential quantification

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

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

$\exists x \;\; P$ is true in an interpretation $\mathcal{I}$ iff P is true with x being some possible object in $\mathcal{I}$'s domain

Roughly speaking, equivalent to the disjunction of instantiations of P

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

Properties of quantifiers

$\forall x \ \forall y \ \varphi$ is equivalent to $\forall y \ \forall x \ \varphi$ (why?)

$\exists x \ \exists y \ \varphi$ is equivalent to $\exists y \ \exists x \ \varphi$ (why?)

$\exists x \ \forall y \ \varphi$ is not equivalent to $\forall y \ \exists x \ \varphi$

Ex.

$\exists x \ \forall y \ \ Loves(x, y)$
"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

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

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

From English prepositions to FOL connectives

English	Logic
A and B $\mid$ A but B	$A \wedge B$
A if B $\mid$ A when B $\mid$ A whenever B	$B \Rightarrow A$
if A, then B $\mid$ A implies B $\mid$ A forces B	$A \Rightarrow B$
only if A, B $\mid$ B only if A $\mid$	$B \Rightarrow A$
A precisely when B $\mid$ A if and only if B	$B \Leftrightarrow A \mid A \Leftrightarrow B$
A or B (or both) $\mid$ A unless B	$A \vee B$ (logical or)
either A or B (but not both)	$A \oplus B$ (exclusive or)

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"

A common mistake to avoid

Typically, $\Rightarrow$ is the main connective with $\forall$

Common mistake: using $\wedge$ as the main connective with $\forall$:

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

means "Everyone is at Berkeley and everyone is smart"

Compare with

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

"Everyone at Berkeley is smart"

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!

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!

Compare with

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

"Someone at Stanford is smart"

Fun with sentences

Brothers are siblings

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall\, x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall\, x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

One's mother is one's female parent

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall\, x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

One's mother is one's female parent

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

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall\, x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

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

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall\, x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

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

$$\forall\, x_1, x_2 \;\; FirstCousin(x_1, x_2) \;\Leftrightarrow\; \exists\, p_1, p_2 \;\; Siblings(p_1, p_2) \wedge Parent(p_1, x_1) \wedge Parent(p_2, x_2)$$

Fun with sentences

Brothers are siblings

$\forall\, x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall\, x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

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

$$\forall\, x_1, x_2 \;\; FirstCousin(x_1, x_2) \;\Leftrightarrow\; \exists\, p_1, p_2 \;\; Siblings(p_1, p_2) \wedge Parent(p_1, x_1) \wedge Parent(p_2, x_2)$$

Dogs are mammals

Fun with sentences

Brothers are siblings

$\forall x, y \;\; Brothers(x, y) \Rightarrow Siblings(x, y)$

"Siblings" is symmetric

$\forall x, y \;\; Siblings(x, y) \;\Leftrightarrow\; Siblings(y, x)$

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

$$\forall x_1, x_2 \;\; FirstCousin(x_1, x_2) \;\Leftrightarrow\; \exists p_1, p_2 \;\; Siblings(p_1, p_2) \wedge Parent(p_1, x_1) \wedge Parent(p_2, x_2)$$

Dogs are mammals

$\forall x \;\; Dog(x) \Rightarrow Mammal(x)$

Equality

Recall that $t_1 = t_2$ is true under a given interpretation if and only if t_1 and t_2 refer to the same object

E.g., $1 = 2$ and $x * x = x$ are satisfiable
$2 = 2$ is valid

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

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

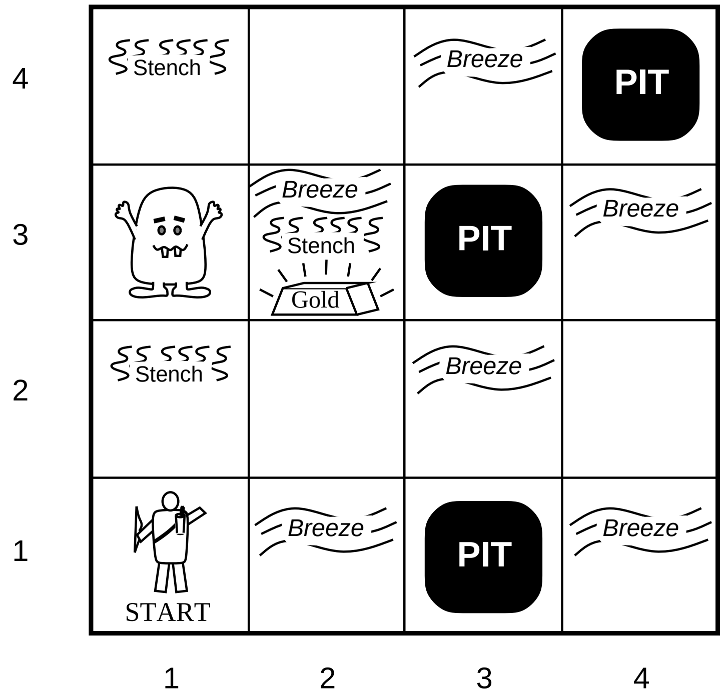
More fun with sentences

1. No one is his/her own sibling
2. Sisters are female, brothers are male
3. Every one is male or female but not both
4. Every married person has a spouse
5. Married people have spouses
6. Only married people have spouses
7. People cannot be married to their siblings
8. Not everybody has a spouse
9. Everybody has a mother
10. Everybody has a mother and only one

More fun with sentences

1. $\forall x \ \neg Siblings(x,x)$
2. $\forall x,y \quad (Sisters(x,y) \Rightarrow Female(x) \land Female(y)) \land$
   $(Brothers(x,y) \Rightarrow Male(x) \land Male(y))$
3. $\forall x \ \ Person(x) \Rightarrow (Male(x) \lor Female(x)) \land$
   $\neg(Male(x) \land Female(x))$
4. $\forall x \ \ (Person(x) \land Married(x)) \Rightarrow \exists y \ \ Spouse(x,y)$
5. $\forall x \ \ (Person(x) \land Married(x)) \Rightarrow \exists y \ \ Spouse(x,y)$
6. $\forall x,y \ \ (Person(x) \land Person(y) \land Spouse(x,y)) \Rightarrow Married(x) \land Married(y)$
7. $\forall x,y \ \ Spouse(x,y) \Rightarrow \neg Siblings(x,y)$
8. $\neg\forall x \ \ Person(x) \Rightarrow \exists y \ \ Spouse(x,y)$
   Alter.: $\exists x \ \ Person(x) \land \neg\exists y \ \ Spouse(x,y)$
9. $\forall x \ \ Person(x) \Rightarrow \exists y \ \ Mother(y,x)$
10. $\forall x \ \ Person(x) \Rightarrow \exists y \ \ Mother(y,x) \land$
    $\neg\exists z \ \ \neg(y = z) \land Mother(z,x)$

The Wumpus World in FOL



Interacting with FOL KBs

Suppose a wumpus-world agent is using an FOL KB
and perceives a stench and a breeze (but no glitter) at time $t = 5$:

$Tell(KB, Percept([Stench, Breeze, None], 5))$
$Ask(KB, \exists a\ Action(a, 5))$

I.e., does the KB entail any particular actions at time $t = 5$?

Answer: Yes, $\{a/Shoot\}$ $\leftarrow$ substitution (binding list)

Given a sentence φ and a substitution σ,
$\varphi\sigma$ denotes the result of plugging σ into φ

Ex: $\varphi = Smarter(x, y)$ $\sigma = \{x/Bart,\ y/Homer\}$
$\varphi\sigma = Smarter(Bart, Homer)$

$AskVar(KB, \exists x\ \varphi)$ returns some/all σ such that $KB \models \varphi\sigma$

Knowledge base for the wumpus world

"Perception"

$\forall\, b, g, t \;\; Percept([Stench, b, g], t) \Rightarrow Smelt(t)$
$\forall\, s, b, t \;\; Percept([s, b, Glitter], t) \Rightarrow AtGold(t)$

Reflex:

$\forall\, t \;\; AtGold(t) \Rightarrow Action(Grab, t)$

Reflex with internal state: do we have the gold already?

$\forall\, t \;\; AtGold(t) \wedge \neg Holding(Gold, t) \Rightarrow Action(Grab, t)$

Note: $Holding(Gold, t)$ cannot be observed, hence keeping track of change is essential

Deducing hidden properties

Properties of locations:

$\forall\, x, t \;\; At(Agent, x, t) \wedge Smelt(t) \Rightarrow Smelly(x)$
$\forall\, x, t \;\; At(Agent, x, t) \wedge Breeze(t) \Rightarrow Breezy(x)$

Squares are breezy near a pit:

- Diagnostic rule — infer cause from effect

  $\forall\, y \;\; Breezy(y) \Rightarrow \exists\, x \;\; Pit(x) \wedge Adjacent(x, y)$

- Causal rule — infer effect from cause

  $\forall\, x, y \;\; Pit(x) \wedge Adjacent(x, y) \Rightarrow Breezy(y)$

- Neither of these is complete — e.g., the causal rule doesn't say whether squares far away from pits can be breezy

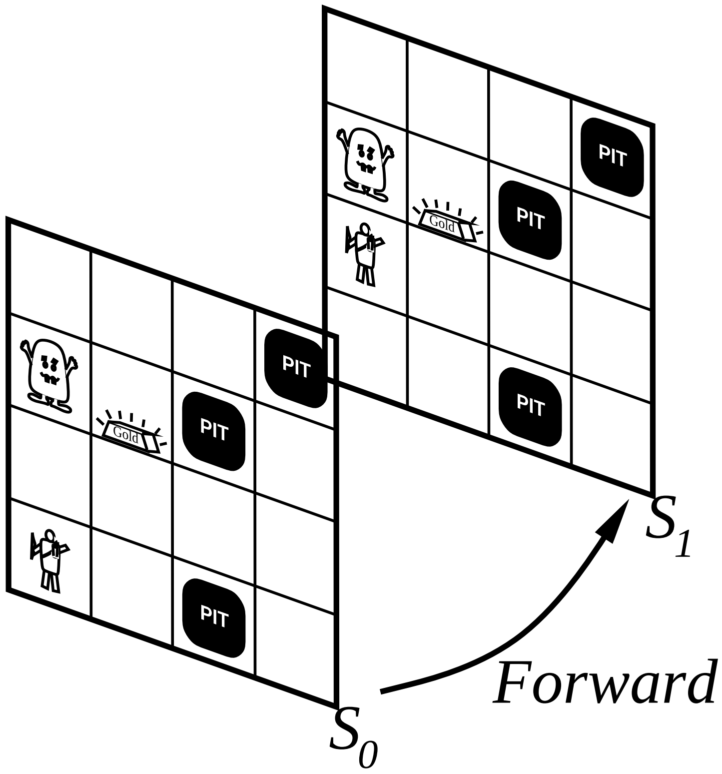
- Definition for the $Breezy$ predicate:

  $\forall\, y \;\; Breezy(y) \;\Leftrightarrow\; (\exists\, x \;\; Pit(x) \wedge Adjacent(x, y))$

Keeping Track of Change

Some facts hold in *situations*, rather than eternally

E.g.,
$Holding(Gold, Now)$ rather than just $Holding(Gold)$
$At(Agent, [1, 1], t)$ rather than just $At(Agent, [1, 1])$



Situation Calculus

Situation calculus is one way to represent change in FOL:
Adds a situation argument to each *fluent*, i.e., non-eternal predicate

E.g.,
Now in $Holding(Gold, Now)$ denotes a *situation* (or a *time stamp*)

Situations are connected by the $Result$ function:

$Result(a, s)$ is the situation that results from doing a in s

Describing Actions

Effect axioms: describe changes due to action

$\forall\, s \;\; AtGold(s) \Rightarrow Holding(Gold, Result(Grab, s))$

Frame axiom: describe non-changes due to action

$\forall\, s \;\; HaveArrow(s) \Rightarrow HaveArrow(Result(Grab, s))$

Frame, Qualification, and Ramification

Frame problem: find an elegant way to handle non-change

- representation—avoid frame axioms
- inference—avoid repeated “copy-overs” to keep track of state

Qualification problem: true descriptions of real actions require endless caveats—what if gold is slippery or nailed down or . . .

Ramification problem: real actions have many secondary consequences—what about the dust on the gold, wear and tear on gloves, . . .

Describing Actions

Successor-state axioms solve the representational frame problem

Each axiom is about a predicate, not an action per se:

P true afterwards $\Leftrightarrow$ (an action made P true $\vee$
P true already and no action made P false)

Example: For holding the gold:

$$\forall\, a, s \;\; Holding(Gold, Result(a, s)) \;\Leftrightarrow$$
$$[AtGold(s) \wedge (a = Grab) \vee$$
$$Holding(Gold, s) \wedge (a \neq Release)]$$

Making Plans

Example

Initial condition in KB:
$At(Agent, [1,1], S_0)$
$At(Gold, [2,1], S_0)$

Query: $Ask(KB, \exists s\, Holding(Gold, s))$
i.e., in what situation will I be holding the gold?

Answer: $\{s/Result(Grab, Result(Forward, S_0))\}$
i.e., go forward and then grab the gold

This assumes that the agent is interested in plans starting at S_0 and that S_0 is the only situation described in the KB

Making plans: A better way

Represent plans as action sequences $[a_1, a_2, \ldots, a_n]$

- $PlanResult(p, s)$ is the result of executing p in s
- Then the query
  $Ask(KB, \exists\, p \;\; Holding(Gold, PlanResult(p, S_0)))$
  has the solution $\{p/[Forward, Grab]\}$
- Definition of $PlanResult$ in terms of $Result$:
  $\forall\, s \;\; PlanResult([\,], s) = s$
  $\forall\, a, p, s \;\; PlanResult(a :: p, s) = PlanResult(p, Result(a, s))$

Planning systems are special-purpose reasoners designed to do this type of inference more efficiently than a general-purpose reasoner