First-Order Logic

Russell and Norvig:
Chapter 8, Sections 8.1-8.3

Outline

- Why FOL?
- Syntax and semantics of FOL
- Using FOL
- Wumpus world in FOL

Propositional logic, pros and cons

- ☺ Propositional logic is declarative
- ☺ Propositional logic 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 $P_{1,2}$

Propositional logic, pros and cons

- ☺ 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

Why not use Natural Language?

- It serves a different purpose:
  - Communication rather than representation
- It is not compositional
  - Context matters
- It can be ambiguous
  - Again, context matters

Create a new language

- Builds on propositinal logic
- But is inspired by natural language!

First-order logic

- Whereas propositional logic assumes the world contains facts,
- first-order logic (like natural language) assumes the world contains
  - Objects: people, houses, numbers, colors, baseball games, wars, …
  - Relations: red, round, prime, brother of, bigger than, part of, comes between, …
  - Functions: father of, best friend, one more than, plus, …

Syntax of FOL: Basic elements

- Constants KingJohn, 2, NUS,...
- Predicates Brother, >,...
- Functions Sqrt, LeftLegOf,...
- Variables x, y, a, b,...
- Connectives $\neg, \Rightarrow, \wedge, \vee, \Leftrightarrow$
- Equality $=$
- Quantifiers $\forall, \exists$

Atomic sentences

Term = *function* ($term_1$,...,$term_n$) or *constant* or *variable*

Atomic sentence = *predicate* ($term_1$,...,$term_n$) or $term_1 = term_2$

Examples:

Brother(KingJohn,RichardTheLionheart)

>(Length(LeftLegOf(Richard)),Length(LeftLegOf(KingJohn)))

Complex sentences

Complex sentences = Made from atomic sentences using connectives

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

Examples:

$Sibling(KingJohn,Richard) \Rightarrow Sibling(Richard,KingJohn)$

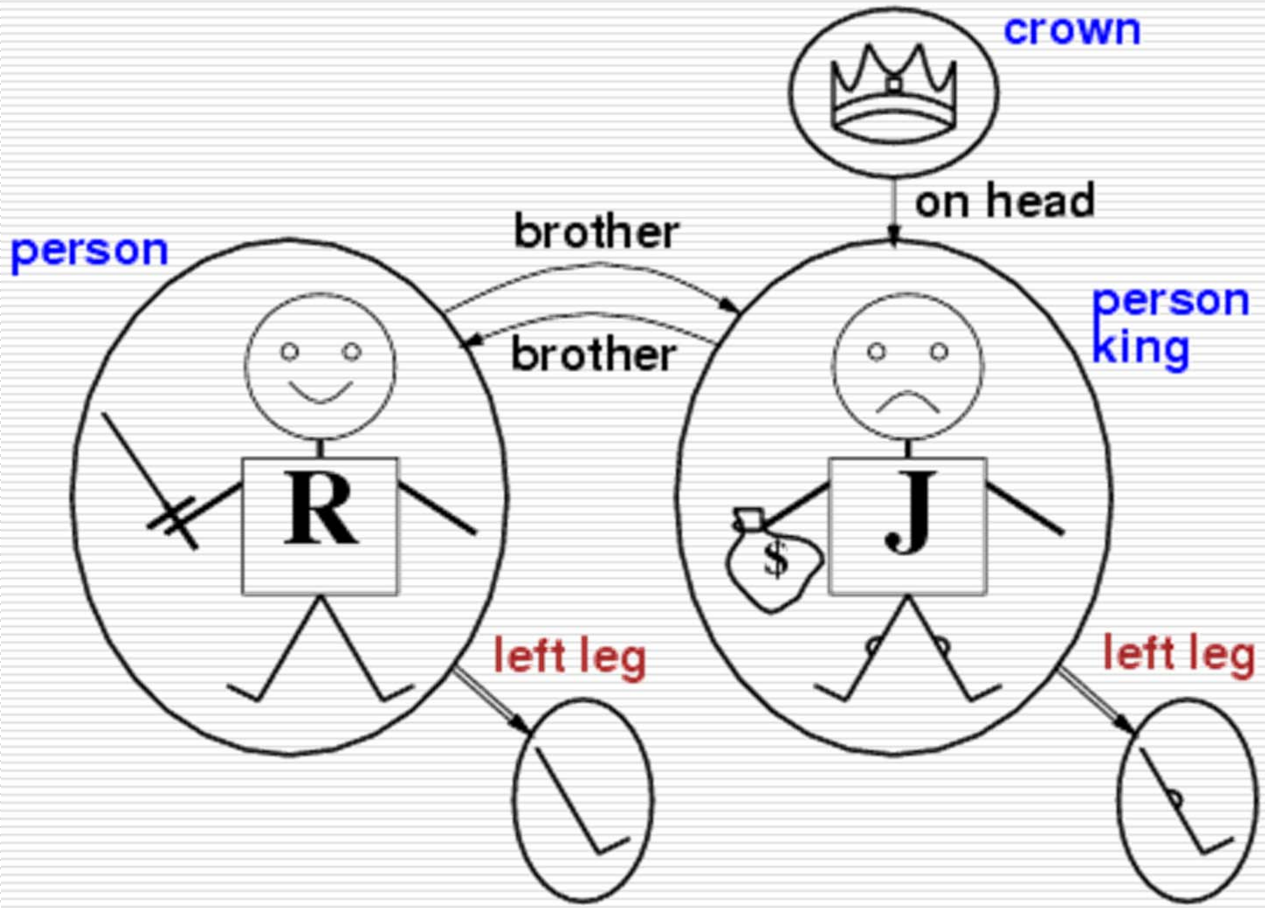
$>(1,2) \vee \leq (1,2)$

$<(1,2) \wedge \neg >(1,2)$

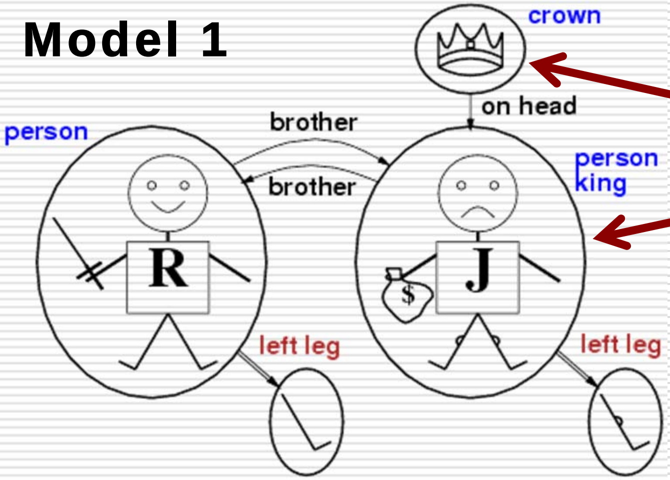
Truth in first-order logic

- Sentences are true with respect to a model
- A model contains objects (domain elements) and relations among them
- A model specifies an interpretation (referents) for
  constant symbols → objects
  predicate symbols → relations
  function symbols → functional relations
- An atomic sentence $predicate(term_1,...,term_n)$ is true iff the objects referred to by $term_1,...,term_n$ are in the relation referred to by *predicate*

Models and Interpretations



Models and Interpretations



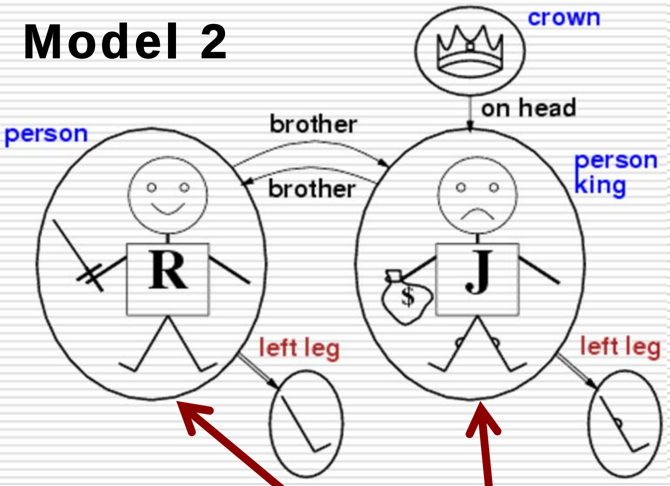
Interpretation I.

Constant: A

Constant: B

Atomic sentence: Brother(A, B)? ➔ **False**

Models and Interpretations



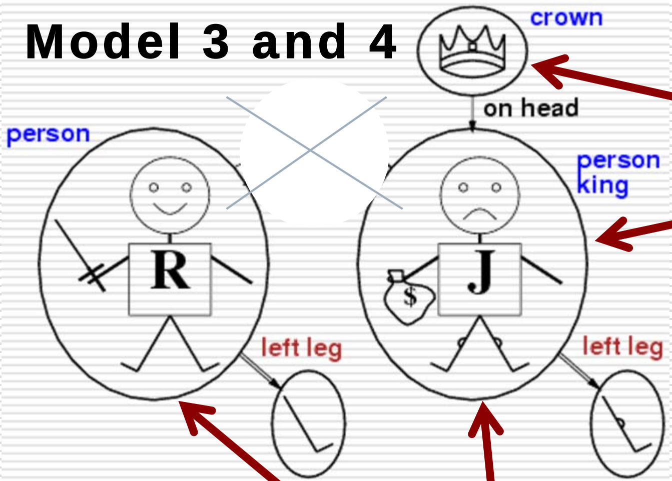
Atomic sentence: Brother(A, B)? ➔ **True**

Interpretation II.

Constant: A

Constant: B

Models and Interpretations



Interpretation I.

Constant: A

Constant: B

Atomic sentence: Brother(A, B)? ➔ **False**

Interpretation II.

Constant: A

Constant: B

Universal quantification

- $\forall$<*variables*> <*sentence*>

Everyone in HR is smart:
$\forall x\ At(x,HR) \Rightarrow Smart(x)$

- $\forall x$ *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*

$$At(KingJohn,HR) \Rightarrow Smart(KingJohn)$$
$$\land\ At(Richard,HR) \Rightarrow Smart(Richard)$$
$$\land\ At(HR,HR) \Rightarrow Smart(HR)$$
$$\land\ \ldots$$

A common mistake to avoid

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

  $\forall x\ At(x,HR) \wedge Smart(x)$

  means "Everyone is at HR and everyone is smart"

Existential quantification

- $\exists$*<variables> <sentence>*

  Someone at HR is smart:
  $\exists x$ At(x,HR) $\wedge$ 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*

  At(KingJohn,HR) $\wedge$ Smart(KingJohn)
  $\vee$ At(Richard,HR) $\wedge$ Smart(Richard)
  $\vee$ At(HR,HR) $\wedge$ Smart(HR)
  $\vee$ ...

Another mistake to avoid

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

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

$$\exists x\ At(x,HR) \Rightarrow Smart(x)$$

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

Properties of quantifiers

- $\forall x\ \forall y$ is the same as $\forall y\ \forall x$
- $\exists x\ \exists y$ is the same as $\exists y\ \exists x$

- $\exists x\ \forall y$ 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) $\quad \neg\exists x\ \neg$Likes(x,IceCream)

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

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., definition of *Sibling* in terms of *Parent*:

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

Using FOL

The kinship domain:

- Brothers are siblings

  $\forall$x,y *Brother(x,y)* $\Leftrightarrow$ *Sibling(x,y)*

- One's mother is one's female parent

  $\forall$m,c *Mother(c)* = m $\Leftrightarrow$ *(Female(m)* $\wedge$ *Parent(m,c))*

- "Sibling" is symmetric

  $\forall$x,y *Sibling(x,y)* $\Leftrightarrow$ *Sibling(y,x)*

Some sentences are ***Axioms*** *(i.e. definitions, facts) while others are* ***Theorems*** *derived from those.*

Wumpus World

- Perceives STENCH adjacent to WUMPUS
- Perceives BREEZE adjacent to PIT
- Perceives GLITTER in GOLD room
- Perceives BUMP when hitting wall
- Can move forwards, turn left, turn right or shoot an arrow. Arrow flies in facing direction until hitting a wall or killing a WUMPUS
- Perceives SCREAM if WUMPUS gets killed
- Can pick up GOLD if in same room

Interacting with FOL KBs

- Suppose a wumpus-world agent is using an FOL KB and perceives STENCH and BREEZE (but no GLITTER) at $t=5$:

  `Tell`(KB,Percept([STENCH,BREEZE,None],5))
  `Ask`(KB,∃a BestAction(a,5))

- I.e., does the KB entail some best action at $t=5$?

- Answer: *Yes*, {*a/Shoot*}   ← substitution (binding list)

- Given a sentence *S* and a substitution q,
- *S*q denotes the result of plugging q into *S*; e.g.,
  S = Smarter(x,y)
  q = {x/Hillary,y/Bill}
  *S*q = Smarter(Hillary,Bill)

- `Ask`(KB,S) returns some/all q such that KB ⊨ Sq

KB for the wumpus world

- Perception
  - $\forall$t,s,b Percept([s,b,GLITTER],t) $\Rightarrow$ Glitter(t)
- Reflex
  - $\forall$t Glitter(t) $\Rightarrow$ BestAction(Grab,t)

Deducing hidden properties

- $\forall x,y,a,b$ *Adjacent*([x,y],[a,b]) $\Leftrightarrow$ [a,b] $\in$ {[x+1,y], [x-1,y],[x,y+1],[x,y-1]}

Properties of squares:

- $\forall s,t$ *At*(Agent,s,t) $\wedge$ Breeze(t) $\Rightarrow$ Breezy(s)

Squares are breezy near a pit:

- Diagnostic rule---infer cause from effect
  $\forall s$ Breezy(s) $\Rightarrow$ $\exists r$ Adjacent(r,s) $\wedge$ Pit(r)
- Causal rule---infer effect from cause
  $\forall r$ Pit(r) $\Rightarrow$ [$\forall s$ Adjacent(r,s) $\Rightarrow$ Breezy(s)]

Summary

- First-order logic:
  - objects and relations are semantic primitives
  - syntax: constants, functions, predicates, equality, quantifiers
- Increased expressive power: sufficient to define Wumpus world
  - We did not have to write sentence for every square!