22c:145 Artificial Intelligence

First-Order Logic

Readings: Chapter 8 of Russell & Norvig.

Language of FOL: Grammar

Sentence	::=	AtomicS | ComplexS
AtomicS	::=	**True** | **False** | RelationSymb(Term, ...) | Term = Term
ComplexS	::=	(Sentence) | Sentence Connective Sentence | ¬Sentence
	|	Quantifier Sentence
Term	::=	FunctionSymb(Term, ...) | ConstantSymb | Variable
Connective	::=	$\wedge \mid \vee \mid \Rightarrow \mid \Leftrightarrow$
Quantifier	::=	∀ Variable | ∃ 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	::=	$P \mid Q \mid \cdots \mid Red \mid Brother \mid Apple \mid > \mid \cdots$

Universal quantification

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

Everyone at UIowa is smart:

$\forall x \ \ At(x, UIowa) \implies Smart(x)$

$\forall x \ \ P(x)$ is true in an interpretation m iff $P(x)$ is true with x being each possible object in the domain

Roughly speaking, equivalent to the conjunction of instantiations of P

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

A common mistake to avoid

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

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

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

means “Everyone is at UIowa and everyone 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!

Existential quantification

$\exists\, \langle variable \rangle \;\; \langle sentence \rangle$

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

$\exists\, x \;\; P$ is true in an interpretation I iff P is true with x being *some* possible object in the domain

Roughly speaking, equivalent to the disjunction of instantiations of P

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

Fun with sentences

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

"Sibling" is symmetric

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

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

One's mother is one's female parent
$\forall\, x, y \;\; Mother(x,y) \Leftrightarrow (Female(x) \land Parent(x,y))$.

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

Fun with sentences

Brothers are siblings
$\forall\, x, y \;\; Brother(x,y) \implies Sibling(x,y)$.

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

One's mother is one's female parent

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 $Parent$:
$\forall\, x, y \;\; Sibling(x,y) \Leftrightarrow [\neg(x = y) \land \exists\, m, f \;\; \neg(m = f) \land$

$Parent(m,x) \land Parent(f,x) \land Parent(m,y) \land Parent(f,y)]$

Fun with sentences

Brothers are siblings
$\forall\, x, y \;\; Brother(x,y) \implies Sibling(x,y)$.

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

One's mother is one's female parent
$\forall\, x, y \;\; Mother(x,y) \Leftrightarrow (Female(x) \land Parent(x,y))$.

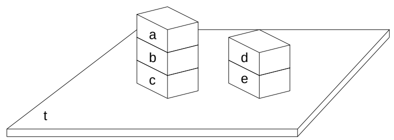
A first cousin is a child of a parent's sibling
$\forall\, x, y \;\; FirstCousin(x,y) \Leftrightarrow \exists\, p, ps \;\; Parent(p,x) \land$
$Sibling(ps,p) \land Parent(ps,y)$

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

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

Function Symbols: $Support$

Relation Symbols: $On, Above, Clear$



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

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

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;

Models for FOL: Lots!

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

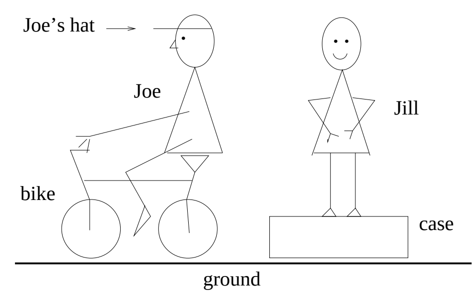
For each number of universe elements n from 1 to ∞
  For each k-ary predicate P_k in the sentence
    For each possible k-ary relation on n objects
      For each constant symbol C in the sentence
        For each one of n objects mapped to C
        . . .

Enumerating models is not going to be easy!

A different interpretation: $\mathcal{B}$

Constant Symbols: A, B, C, D, E, T Function Symbols: $Support$

Relation Symbols: $On, Above, Clear$



$A^{\mathcal{B}} = Joe'sHat, B^{\mathcal{B}} = Joe, C^{\mathcal{B}} = bike, D^{\mathcal{B}} = Jill, E^{\mathcal{B}} = case, T^{\mathcal{B}} = ground$

$$
\begin{aligned}
Support^{\mathcal{B}} &= \{\langle Joe's\ hat, Joe\rangle, \langle Joe, bike\rangle, \langle bike, ground\rangle, \langle Jill, case\rangle, \\
&\quad \langle case, ground\rangle, \langle ground, ground\rangle\} \\
On^{\mathcal{B}} &= \{\langle Joe's\ hat, Joe\rangle, \langle Joe, bike\rangle, \langle Jill, case\rangle, \} \\
Above^{\mathcal{B}} &= \{\langle Joe's\ hat, Joe\rangle, \langle Joe's\ hat, bike\rangle, \ldots\}
\end{aligned}
$$

$Clear^{\mathcal{B}} = \{\langle Jill\rangle$ (no hat), $Joe's\ hat\}$

Example

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

$$\begin{array}{rl} MotherOf^{\mathcal{D}} & \text{is a unary fn mapping people to their mother} \\ SchoolOf^{\mathcal{D}} & \text{is a unary fn mapping people to their school} \\ FchildOf^{\mathcal{D}} & \text{is a binary fn mapping a couple to their first child} \\ Bill^{\mathcal{D}} & \text{is Bill Clinton} \\ \sigma & := \{x \mapsto \text{Chelsea Clinton},\ y \mapsto \text{Hillary Clinton}\} \end{array}$$

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

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

- What is the meaning of $\boxed{SchoolOf(FchildOf(y, Bill))}$?

$$[\![SchoolOf(FchildOf(y, Bill))]\!]^{\mathcal{D}}_{\sigma} = SchoolOf^{\mathcal{D}}(FchildOf^{\mathcal{D}}(\sigma(y), Bill^{\mathcal{D}}) = \text{Sta}$$

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.

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

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):

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

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.

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

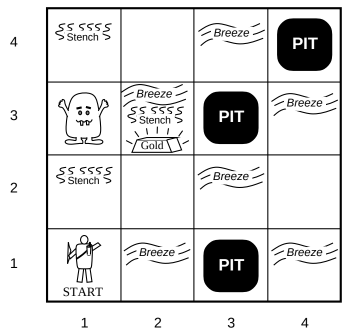
Models, Validity, etc. for Sets of Sentences

- 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.

Models, Validity, etc. for Sentences

- An interpretation $(\mathcal{D}, \sigma)$ **satisfies** a sentence φ, or is a **model** for φ, if $[\![\varphi]\!]^{\mathcal{D}}_{\sigma} = \mathit{True}$.
- 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.
- Valid sentences do not tell us anything about the world. (They are always true!)

The Wumpus World in FOL



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!

Knowledge base for the wumpus world

"Perception"
$\forall b, g, t \;\; Percept([Smell, b, g], t) \implies Smelt(t)$
$\forall s, b, t \;\; Percept([s, b, Glitter], t) \implies AtGold(t)$

Reflex: $\forall t \;\; AtGold(t) \implies Action(Grab, t)$

Reflex with internal state: do we have the gold already?
$\forall t \;\; AtGold(t) \land \neg Holding(Gold, t) \implies Action(Grab, t)$

$Holding(Gold, t)$ cannot be observed
$\Rightarrow$ keeping track of change is essential

Interacting with FOL KBs

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

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

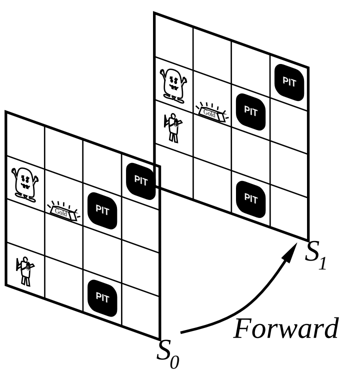
I.e., does the KB entail any particular actions at $t = 5$?

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

Given a sentence S and a substitution σ,
$S\sigma$ denotes the result of plugging σ into S; e.g.,
$S = Smarter(x, y)$
$\sigma = \{x/Hillary, y/Bill\}$
$S\sigma = Smarter(Hillary, Bill)$
$Ask(KB, S)$ returns some/all σ such that $KB \models S\sigma$

Keeping Track of Change

Facts hold in situations, rather than eternally
E.g., $Holding(Gold, Now)$ rather than just $Holding(Gold)$



Deducing Hidden Properties

- Properties of locations:
  - $\forall x, t \;\; At(Agent, x, t) \land Smelt(t) \implies Smelly(x)$
  - $\forall x, t \;\; At(Agent, x, t) \land Breeze(t) \implies Breezy(x)$
- Squares are breezy near a pit:
  - Diagnostic rule—infer cause from effect
    $\forall y \;\; Breezy(y) \implies \exists x \;\; Pit(x) \land Adjacent(x, y)$
  - Causal rule—infer effect from cause
    $\forall x, y \;\; Pit(x) \land Adjacent(x, y) \implies Breezy(y)$
  - 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:
    $\forall y \;\; Breezy(y) \Leftrightarrow [\exists x \;\; Pit(x) \land Adjacent(x, y)]$

Describing Actions

- "Effect" axiom—describe changes due to action
  $\forall s \;\; AtGold(s) \implies Holding(Gold, Result(Grab, s))$
- "Frame" axiom—describe non-changes due to action
  $\forall s \;\; HaveArrow(s) \implies HaveArrow(Result(Grab, s))$

Situation Calculus

Situation calculus is one way to represent change in FOL:
Adds a situation argument to each 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

- 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

- Exmaple: For holding the gold:
  $\forall a, s \;\; Holding(Gold, Result(a, s)) \Leftrightarrow$
  $[(a = Grab) \wedge AtGold(s) \vee$
  $Holding(Gold, s) \wedge (a \neq Release)]$

Frame, Qualifcation, 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, ...

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

Making Plans

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

Summary

- First-order logic:
  – objects and relations are semantic primitives
  – syntax: constants, functions, predicates, equality, quantifiers
- Increased expressive power: sufficient to define wumpus world
- Situation calculus:
  – conventions for describing actions and change in FOL
  – can formulate planning as inference on a situation calculus KB