Einstein's Puzzle in Logic

- We used propositinal variables to specify everything:
  x_1 = "house #1 is red";
  x_2 = "house #2 is red";
  ...
  x_{26} = "Brit lives in house #1";
  ...
- Can we use one symbol with parameters for "the color of house $\#x$ is y", i.e., *IsColor*(x, y)?
- Or simply, *color*$(x) = y$?
- Yes, we can do this in First-Order Logic.

22c:145 Artificial Intelligence

First-Order Logic

Readings: Chapter 8 of Russell & Norvig.

Ontological Commitments

1. The world is made of **objects**, *things* with individual identities and **properties** that distinguish them.
2. Various **relations** hold among objects. Some of these relations are **functional**.
3. Every fact involving objects and their relations is either true or false.

First-Order Logic

- Rather powerful representation and reasoning system.
- Very well understood and extensively studied (a couple of thousand years!).
- Many fancy knowledge representation formalisms—semantic nets, frames, scripts—are basically sugar-coated variants of (part of) it.

Language of FOL: Symbols

- Variables, constant symbols, and function symbols are used to build **terms**:
  $x, Bill, 6, FatherOf(x), Height(FatherOf(Bill)), Log(3+y), \ldots$
- Relational symbols are applied to terms to build **predicates**:
  $Even(x), Married(Bill, Hillary), Loves(x, MotherOf(x)), \ldots$
- Predicates and logical constants are used to build **sentences**:
  $Even(3), \exists x\, \neg Loves(x, MotherOf(x)), \forall x\, Even(x) \Rightarrow Odd(x+1), \ldots$

The Language of First-Order Logic

Symbols

- Variables
- Constant symbols
- Functional symbols (with *arities*)
- Relational symbols (with *arities*)
- Logical constants: As in Propositional Logic plus

  $\forall$ (for all) $\exists$ (there exists) $=$ (equal)

Language of FOL: Sentences

- $\mathbf{True}$ ($\mathbf{False}$) is a sentence.
- If t_1, t_2 are terms, $t_1 = t_2$ is a sentence.
- If p is an n-ary relation symbol and $t_1, \ldots, t_n$ are n terms, $p(t_1, \ldots, t_n)$ is a sentence.
- If φ is a sentence, $\neg\varphi$, $\exists x\, \varphi$, $\forall x\, \varphi$ are sentences.
- If φ_1, φ_2 are sentences, $\varphi_1 \wedge \varphi_2$, $\varphi_1 \vee \varphi_2$, $\varphi_1 \Rightarrow \varphi_2$, $\varphi_1 \Leftrightarrow \varphi_2$ are sentences.
- Nothing else is a sentence.

Language of FOL: Terms

- A variable is a term.
- A constant symbol is a term.
- If f is an n-ary function symbol and $t_1, \ldots, t_n$ are n terms, $f(t_1, \ldots, t_n)$ is a term.
- Nothing else is a term.

A Note on the Syntax of FOL

As defined, the syntax of FOL is ambiguous:

$$\forall x \, \neg P(x) \vee Q(x) \wedge R(x) \Rightarrow S(x) \tag{1}$$

It can be disambiguated by using parentheses or fixing a precedence ordering on the logical connectives. We will use the following ordering:

$$(\neg, \; \wedge, \; \vee, \; \Rightarrow, \; \Leftrightarrow, \; \{\forall, \exists\})$$

With this ordering, (1) actually is

$$\forall x \; (((\neg P(x)) \vee (Q(x) \wedge R(x))) \Rightarrow S(x))$$

When in doubt, use parentheses!

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	::=	$\wedge$ | $\vee$ | $\Rightarrow$ | $\Leftrightarrow$
Quantifier	::=	∀ Variable | ∃ Variable
Variable	::=	a | b | $\cdots$ | x | y | $\cdots$
ConstantSymb	::=	A | B | $\cdots$ | $John$ | 0 | 1 | $\cdots$ | π | $\ldots$
FunctionSymb	::=	F | G | $\cdots$ | $Cosine$ | $Height$ | $FatherOf$ | $+$ | $\ldots$
RelationSymb	::=	P | Q | $\cdots$ | Red | $Brother$ | $Apple$ | $>$ | $\cdots$

Constant Symbols and Variables

- Denote (stand for) objects/individuals.
  Example:
  - $Dog143$ is a constant symbol denoting a particular dog.
  - $peach$ is a variable ranging over peaches.
- The symbol⟶object association is arbitrary!
- Under another interpretation $Dog143$ could denote another dog, or maybe something else altogether. Similarly, $peach$ might as well range over apples.

Semantics of First-Order Logic

Interprets

- constant symbols and variables as *objects*;
- functional symbols as *functions* from objects to objects;
- relational symbols as *relations* over objects;
- $=$ as equality (i.e., the *identity* relation).

Semantics of FOL: Examples

	Sentence	Intuitive meaning
1.	$Above(A,B)$	object A is above object B
2.	$British(Elisabeth)$	an “object” called Elisabeth is British
3.	$\forall x\ Apple(x) \Rightarrow Red(x)$	...
4.	$\forall x\ Apple(x) \land Red(x)$	...
5.	$\exists x\ Apple(x) \lor Red(x)$	...
6.	$\forall x\ \exists y\ y > x$	...
6b.	$\forall x\ Number(x) \Rightarrow \exists y\ y > x$	...
7.	$(\forall x\ Apple(x)) \lor \exists y\ Pear(y)$	...
7b.	$(\forall x\ Apple(x)) \lor \exists x\ Pear(x)$	...
8.	$\forall x\ \exists y\ Loves(x,y)$	...
8b.	$\exists y\ \forall x\ Loves(x,y)$	...
9.	$\exists x\ \forall y\ Loves(x,y)$	...

Semantics of First-Order Logic

Interprets

- the universal quantifier (essentially) as an *infinite conjunction*;
  $(\forall x\ Red(x) \equiv Red(Obj_1) \land Red(Obj_2) \land Red(Obj_3) \land Red(Obj_4) \land \cdots)$
- the existential quantifier (essentially) as an *infinite disjunction*;
  $(\exists x\ Red(x) \equiv Red(Obj_1) \lor Red(Obj_2) \lor Red(Obj_3) \lor Red(Obj_4) \lor \cdots)$
- $\mathbf{True}, \mathbf{False}, \land, \lor, \neg, \Rightarrow, \Leftrightarrow$ as in propositional logic.

Free and Bound Variables

An occurrence of variable in a sentence is **free** if it is not in the scope of any quantifier with the same variable. A non-free variable occurrence is **bound**.

$\forall x\ P(x) \land Q(x)$	both occurrences of x bound by $\forall x$
$(\forall x\ P(x)) \land Q(x)$	first occurrence of x bound by $\forall x$, second free
$\forall x\ R(x,z) \land \exists y\ S(x,y)$	both occurrences of x bound by $\forall x$, occurrence of y bound by $\exists y$, occurrence of z free

A sentence is **open** if it contains free variables; it is **closed** otherwise.

Semantics of FOL: Examples

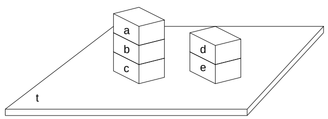
	Sentence	Intuitive meaning
10.	$\forall x\ (Bird(x) \land \neg Ostrich(x)) \Rightarrow Flies(x)$	...
11.	$\forall x\ (Fish(x) \land \neg Ostrich(x)) \Rightarrow Flies(x)$	...
12.	$\forall x\ Flies(x) \Leftrightarrow (Bird(x) \lor Plane(x))$	...
13.	$\forall x\ Person(x) \Rightarrow (W(x) \lor M(x)) \land \neg(W(x) \land M(x))$	...
14.	$\forall x\ Age(x) < Age(FatherOf(x))$	...

Note: The **scope** of a quantifier is the sentence it applies to:

$$\forall x\ \underbrace{P(x) \land Q(x)}_{\text{scope of } \forall x} \qquad \forall x\ \underbrace{R(x,z) \land \underbrace{\exists y\ S(x,y)}_{\text{scope of } \exists y}}_{\text{scope of } \forall x}$$

Relations

- Let U be the universe of discourse.
- A **relation** of arity n is a subset of the Cartesian product $\underbrace{U \times \cdots \times U}_{n \text{ times}}$.
- Example: Some relations in the Blocks World



$$\begin{aligned} On &:= \{\langle a,b\rangle, \langle b,c\rangle, \langle c,t\rangle, \langle d,e\rangle, \langle e,t\rangle\} \\ Clear &:= \{\langle a\rangle, \langle d\rangle\} \end{aligned}$$

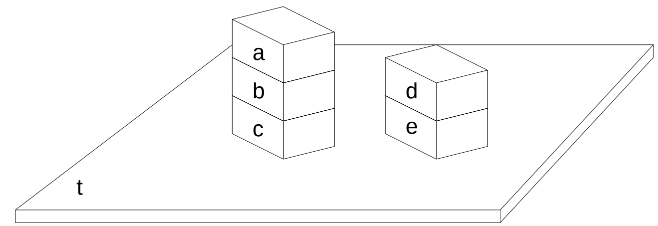
- Example: "Less Than" over the naturals

$$\begin{aligned} < \ &:= \{\langle x,y\rangle \mid x,y \in \mathbb{N}, \exists z \neq 0 \in \mathbb{N} \text{ s.t. } x+z=y\} \\ &:= \{\langle 0,1\rangle, \langle 0,2\rangle, \ldots, \langle 1,2\rangle, \langle 1,3\rangle, \ldots\} \end{aligned}$$

Formalizing Knowledge in FOL

First Step: Fix a **Universe of Discourse**, that is, the set of objects of interest.

- Example: The Blocks World

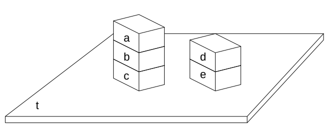


  Here the universe is $\{$ the table, block a, block b, . . . , block e $\}$.

- Note: The universe may be infinite, even uncountable! (the integers, the reals, . . .)

Examples of Functions

- Most mathematical functions $+, \times, e^x, sine, \ldots$
- Some relations in the Blocks World



$$\begin{aligned} Color &:= \{\langle a, blue\rangle, \langle b, red\rangle, \langle c, blue\rangle, \ldots\} \\ Support &:= \{\langle a,b\rangle, \langle b,c\rangle, \langle c,t\rangle, \langle d,e\rangle, \langle e,t\rangle\} \end{aligned}$$

- We may define function *color*(x) such that *color*$(a) = blue$, *color*$(b) = red$, ...
- **Note:** Functions in FOL are always **total**, i.e., defined over the entire universe of discourse.

Functions

- A **function** in one variable is simply a *functional* binary relation.
- A relation f over $U \times U$ is functional if for all $x \in U$ there is *one and only one* $y \in U$ such that $\langle x,y\rangle \in f$.
- If f is a function

$$(\langle x,y\rangle \in f) \wedge (\langle x,z\rangle \in f) \quad \text{implies} \quad y = z$$

  Hence, instead of writing $\langle x,y\rangle \in f$ we write $y = f(x)$.

- Functions in $n > 1$ variables are defined similarly.

Functions and Names

- The functional notation can be used to create several names for the *same* object.
- Examples:

$$3,\ \sqrt{9},\ 1+2,\ age(Baby)$$

$$George\_Bush,\ PresidentOf(US),\ SonOf(HusbandOf(Barbara))$$

Relations and Predicates

- The notation $Mother(Jane, Bill)$ is intended to mean that

$$\langle Jane, Bill\rangle \in Mother$$

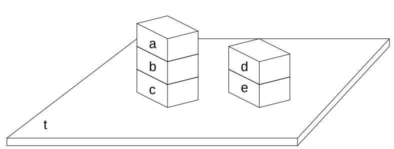
- A relation implicitly defines (or provides the meaning of) a **predicate**, i.e., a function ranging over $\{True, False\}$.
- *Predicate* is a syntactic way of representing a relation (in the style of a function).

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;

Semantics of First-Order Logic

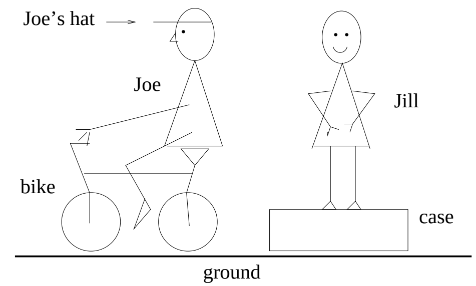
- Let $(\mathcal{D}, \sigma)$ be an interpretation and E an expression of FOL. We write $[\![E]\!]^{\mathcal{D}}_{\sigma}$ to denote the *meaning of* E *in the domain* $\mathcal{D}$ *under the variable assignment* σ.
- The meaning $[\![t]\!]^{\mathcal{D}}_{\sigma}$ of a term t is an object of $\mathcal{D}$.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

A different interpretation: $\mathcal{B}$

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

Relation Symbols: $On, Above, Clear$



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

$Support^{\mathcal{B}}$	$=$	$\{\langle Joe's\ hat, Joe\rangle, \langle Joe, bike\rangle, \langle bike, ground\rangle, \langle Jill, case\rangle, \langle case, ground\rangle, \langle ground, ground\rangle\}$
$On^{\mathcal{B}}$	$=$	$\{\langle Joe's\ hat, Joe\rangle, \langle Joe, bike\rangle, \langle Jill, case\rangle, \}$
$Above^{\mathcal{B}}$	$=$	$\{\langle Joe's\ hat, Joe\rangle, \langle Joe's\ hat, bike\rangle, \ldots\}$
$Clear^{\mathcal{B}}$	$=$	$\{\langle Jill\rangle$ (no hat), $Joe's\ hat\}$

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' coincic with σ except maybe for x

Example

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

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

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

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

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

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

Models, Validity, etc. for Sentences

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

Semantics of First-Order Logic

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

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

- 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):

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

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!

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.