Knowledge Representation

and

Reasoning

First Order Logic

Deepak Khemani
Plaksha University

Sentences, Denotation, and Truth

- A logic is a formal language that defines a set of sentences $\{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \ldots\}$
- Each sentence has a meaning, ascribed from outside.
- Each sentence has a truth value, ascribed from outside
- Compound sentences have truth values that are determined completely by its constituents and logical operators
- Logic has a notion of Entailment
- Logic also has Rules of Inference
- Logic has a machinery or algorithms for producing new sentences from old by applying the rules of inference
- The goal of building the machinery is to mechanically produce all and only the true sentences (conclusions) – given a set of true sentences (axioms, premises)

Deconstructing Sentences

Deduction is possible because of *connections* between *constituents* of sentences

$$\frac{(\alpha \supset \beta) \qquad \alpha}{\beta}$$

Modus Ponens

All men are mortal
Socrates is a man

Socrates is mortal

The form of the argument captures the connections

Propositional Logic cannot make these connections

Aristotelian Logic: Categories and Individuals

Aristotle was concerned with making statements about categories

- *Universal Affirmation*:	All S are P	All Men are Mortal
- *Particular Affirmation*:	Some S are P	Some Children are Tall
- *Universal Denial*:	No S is P	No Priest is Immortal
- *Particular Denial*:	Some S are not P	Some Birds are not Priests

Individuals could belong to categories. For example, “X is a P”

The Socratic Argument is *one of* the fourteen valid forms of the Syllogism

All Men are Mortal

Socrates is a Man

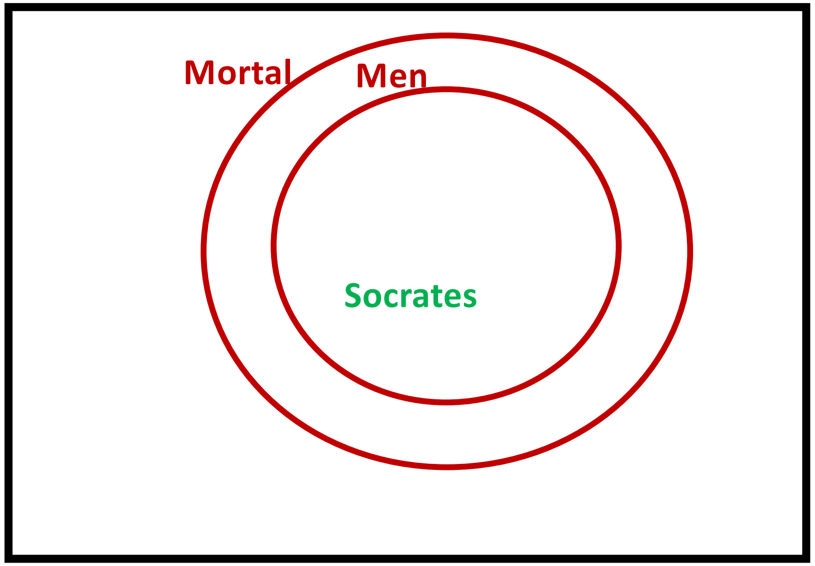
Therefore, Socrates is Mortal

See https://plato.stanford.edu/entries/aristotle-logic/

Relations on a Domain

A Domain D or a Universe of Discourse is the ”world” of First Order Logic

All men are mortal
Socrates is a man
Socrates is mortal

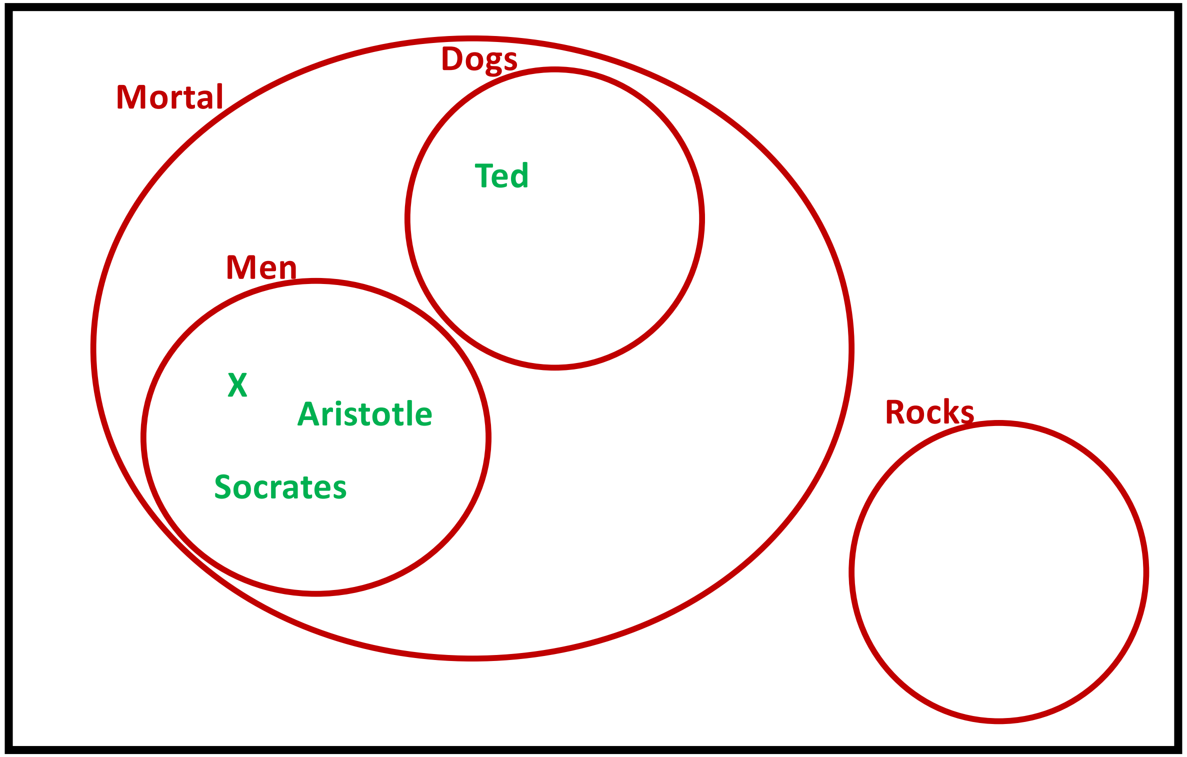


Categories are unary relations or subsets of D

First Order Logic / Predicate Calculus

- The meaning or semantics in PL was externally defined.
- The semantics of First Order Logic (FOL) is defined over a domain.
- The domain D is a set and has individuals
  - Individuals may be named and identified
  - Individuals maybe unknown or variables
- FOL is also called Predicate Calculus because of the use of predicates.
  - Unary predicates are interpreted as subsets of the domain D
  - Binary predicates are binary relations. Subsets of DxD.
- FOL is a logic that breaks down a sentence into constituents and relations between them.
- FOL also has the vocabulary to talk of “all” and “some”

Predicates and Individuals



Individuals
Constants (named) – Aristotle
Variables (unknown) – X

Predicates
Unary relations
Man(Socrates)
Man(Aristotle)
Man(X)
Dog(Ted)
Binary relations
Teacher(Socrates, Aristotle)
Pet(Socrates,Ted)

First Order Logic (FOL): Syntax

The *logical* part of the vocabulary

- Symbols that stand for connectives or operators
  - "∧", "∨", "¬", and "⊃"…
- Brackets "(", ")", "{", "}"…
- The constant symbols "⊥" and "⊤".

- A set of variable symbols $V = \{v_1, v_2, v_3, \ldots\}$
  - commonly used $\{x, y, z, x_1, y_1, z_1, \ldots\}$

- Quantifiers: "∀" read as "for all", and "∃" read as "there exists".
  The former is the *universal quantifier* and
  the latter the *existential quantifier*.
- The symbol "=" read as "equals".

FOL Syntax (contd)

The non-logical part of *FOL vocabulary* constitutes of three sets.

- A set of predicate symbols P = {P_1, P_2, P_3, …}. We also use the symbols {*P*, *Q*, R, …}. More commonly we use words like "Man", "Mortal", "GreaterThan". Each symbol has an arity associated with it.
- A set of function symbols F = {f_1, f_2, f_3, …}. We commonly used the symbols {f, g, h…} or words like "Successor" and "Sum". Each function symbol has an arity that denotes the number of argument it takes.
- A set of constant symbols C = {c_1, c_2, c_3, …}. We often used symbols like "0", or "Socrates", or "Darjeeling" that are meaningful to us.

The three sets define a language L(*P*,*F*,*C*) or L(*R*,*F*,*C*)

Terms of L(*P*,*F*,*C*)

The basic constituents of *FOL* expressions are *terms*. The set of terms $\mathfrak{I}$ of L(*P*,*F*,*C*) is defined as follows. The constants and the variables are terms by definition. More terms are defined using the function symbols.

If $t \in V$ then $t \in \mathfrak{I}$

If $t \in C$ then $t \in \mathfrak{I}$

If $t_1, t_2, \ldots, t_n \in \mathfrak{I}$ and $f \in F$ is an n-place function symbol then $f(t_1,t_2, \ldots,t_n) \in \mathfrak{I}$

Atomic Formulas of L(P,F,C)

The set of formulas is defined using terms and predicate symbols. By default the logical symbols "$\bot$" and "$\top$" are also formulas. The set of well formed formulas *F* of L(*P*,*F*,*C*) is defined as follows.

Atomic formulas $\mathcal{A}$

$\bot \in \mathcal{A}$

$\top \in \mathcal{A}$

If $t_1, t_2 \in \mathfrak{I}$ then $(t_1=t_2) \in \mathcal{A}$

If $t_1, t_2, \ldots, t_n \in \mathfrak{I}$
and $P \in \mathrm{P}$ is an n-place predicate symbol
then $P(t_1,t_2,\ldots,t_n) \in \mathcal{A}$

Formulas of L(P,F,C)

The set of formulas of L(P,F,C) $\mathcal{F}$ is defined as follows

If $\alpha \in \mathcal{A}$ then $\alpha \in \mathcal{F}$

If $\alpha \in \mathcal{F}$ then $\neg\alpha \in \mathcal{F}$

If $\alpha, \beta \in \mathcal{F}$ then $(\alpha \wedge \beta) \in \mathcal{F}$

If $\alpha, \beta \in \mathcal{F}$ then $(\alpha \vee \beta) \in \mathcal{F}$

If $\alpha, \beta \in \mathcal{F}$ then $(\alpha \supset \beta) \in \mathcal{F}$

Universal and Existential Quantifiers

If $\alpha \in \mathcal{F}$ and $x \in V$ then $\forall x\ (\alpha) \in \mathcal{F}$

$\forall x\ (\alpha)$ is read as "for all x (α)"

If $\alpha \in \mathcal{F}$ and $x \in V$ then $\exists x\ (\alpha) \in \mathcal{F}$

$\exists x\ (\alpha)$ is read as "there exists x (α)"

We will also use the notation (forall (x) (α)) and (exists (x) (α)) as given in the book Artificial Intelligence by Eugene Charniak and Drew McDermott.

Makes representation for use in programs simpler.

List notation

Standard mathematical notation

1. $\forall x\ (Man(x) \supset Human(x))$: all men are human beings
2. $Happy(suresh) \vee Rich(suresh)$: Suresh is rich or happy
3. $\forall x\ (CitrusFruit(x) \supset \neg Human(x))$: all citrus fruits are non-human
4. $\exists x\ (Man(x) \wedge Bright(x))$: some men are bright

List notation (a la Charniak & McDermott, "Artificial Intelligence")

1.(forall (x) (if (man x) (human x)))

2.(or (happy suresh) (rich suresh))

3.(forall (x) (if (citrusFruit x) (not (human x))))

4.(exists (x) (and (man x) (bright x)))

Sentences of L(P,F,C)

A variable within the scope of a quantifier is said to be *bound*.

If a variable is not bound then it is *free*.

Example: (forall (x) (and (exists (y) (loves-to-read x y)) (book y))

bound free

A formula of L(P,F,C) without free variables is a sentence of L(P,F,C)

FOL: Rules of Inference

The propositional logic rules we saw earlier are valid in *FOL* as well. In addition we need new rules to handle quantified statements. The two commonly used rules of inference are,

$$\frac{\forall x\ P(x)}{P(a)} \quad \text{where } a \in C \quad \text{Universal Instantiation (UI)}$$

$$\frac{P(a)}{\exists x\ P(x)} \quad \text{where } a \in C \quad \text{Generalization}$$

Examples:

$$\frac{\forall x\ (Man(x) \supset Mortal(x))}{(Man(Socrates) \supset Mortal(Socrates))}$$

$$\frac{(Man(Socrates) \supset Mortal(Socrates))}{\exists x\ (Man(x) \supset Mortal(x))}$$

FOL: Rules of Substitution

The following rules of substitution are also useful,

$\neg\forall x\ \alpha$	$\equiv$	$\exists x\ \neg\alpha$	DeMorgan's law
$\neg\exists x\ \alpha$	$\equiv$	$\forall x\ \neg\alpha$	DeMorgan's law
$\forall x\ \forall y\ \alpha$	$\equiv$	$\forall y\ \forall x\ \alpha$	
$\exists x\ \exists y\ \alpha$	$\equiv$	$\exists y\ \exists x\ \alpha$	

Semantics (First Order Logic)

Difficult to express universal statements meaningfully in Propositional Logic. Consider,

Alice likes mathematics and she likes stories. If she likes mathematics she likes algebra. If she likes algebra and likes physics she will go to college. …

The statements in red colour are specific to Alice. We often want to make these as general statements - *If SOMEONE likes mathematics she likes algebra. If SOMEONE likes algebra and likes physics she will go to college.*

To make such general statements and reason with them we need the notion of variables that FOL gives us, and the universal and existential quantifiers.

The variables take values from a domain, and thus we have the notion of Interpretations in FOL where we choose a domain and interpret the language L(P,F,C) over the domain.

Semantics: Interpretations for L(P,F,C)

An Interpretation ϑ = <D, I> of a FOL language L(P,F,C) constitutes of a domain (or Universe of Discourse) *D* and a mapping *I* from the language *L* to the domain *D*.

Each of the elements of the sets *P, F* and *C* are interpreted over *D*. Each of them is understood or gets meaning from the domain *D*.

Predicate symbols are mapped to relations on *D*

Function symbols are mapped to functions on *D*

Constant symbols are mapped to individuals in *D*

Interpretation ϑ = <D,I> of L(P,F,C)

For every predicate symbol $Q \in P$ of arity N,

$I(Q) = Q^I$ where Q^I is the image of Q and $Q^I \subseteq DxDx\ldots xD$

For every function symbol $f \in F$ of arity N,

$I(f) = f^I$ where f^I is the image of f and f^I: $DxDx\ldots xD \rightarrow D$

For every constant symbol $c \in C$

$I(c) = c^I$ where c^I is the image of c and $c^I \in D$

In addition we have an assignment $A: V \rightarrow D$

from the set of variables of L(P,F,C) to the domain.

$A(v) = v^A$ where $v^A \in D$

Interpretation of Terms of L(P,F,C)

Terms in FOL *denote* elements in the domain.

A term $t \in \mathfrak{I}$ mapped to the element of the domain D as follows.

If $t \in V$ then $t^{IA} = t^{A}$

If $t \in C$ then $t^{IA} = t^{I}$

If $t = f(t_1, t_2, \ldots, t_n)$ and $f \in F$ then $t^{IA} = f^{I}(t_1^{IA}, t_2^{IA}, \ldots, t_n^{IA})$

Variables are mapped by the assignment A.

For example, x → 12

Constants are interpreted by the mapping I.

For example, sifar → 0

Functions denote elements too.

For example, plus(3,8) → 11

Interpretations

Consider the following KB

{(O A B), (O B C), (not (M A)), (M C)}

or $\{O(A, B), O(B, C), \neg M(A), \neg M(C)\}$

Where O is a binary predicate symbol

M is a unary predicate symbol

A, B and C are constant symbols

What is the above KB describing?

Observe that we have not used helpful names!

Remember – meaning lies in the mind of the beholder

Interpretation 1

{(O A B), (O B C), (not (M A)), (M C)}

Domain: Blocks World

Predicate symbols

O → On

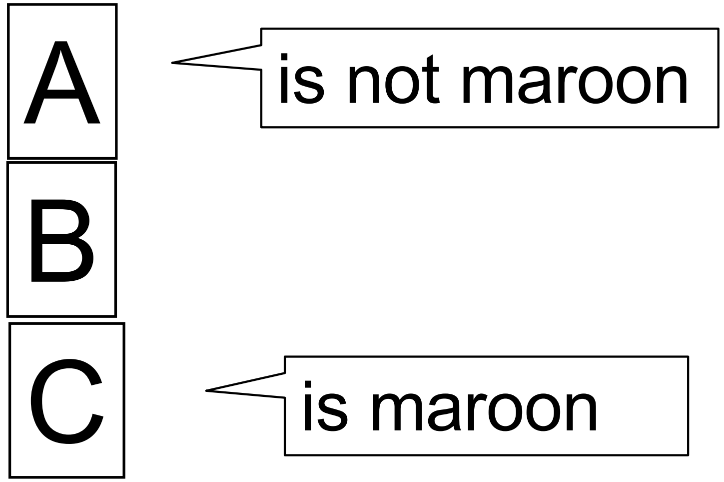
M → Maroon

Constant Symbols

A, B, C → blocks

A is on B

B is on C



Interpretation 2

{(O A B), (O B C), (not (M A)), (M C)}

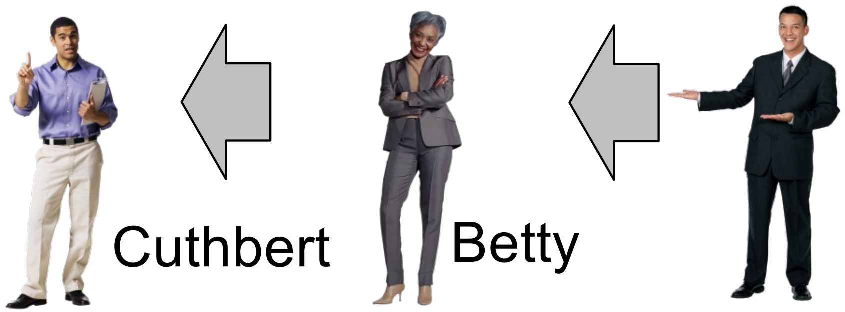
Domain: People

Predicate symbols
O → LookingAt
M → Married

Constant Symbols
A → Archie
B → Betty
C → Cuthbert

Betty is looking at Cuthbert

Archie is looking at Betty



is married

is not married

Truth Assignment to Atomic Formulas of FOL

A valuation function Val: $\mathcal{F} \rightarrow \{true, false\}$

$$Val(\top) = true$$

$$Val(\bot) = false$$

$$Val(t_1{=}t_2)^{IA} = true \quad \text{iff} \quad t_1^{IA} = t_2^{IA}$$

$$Val(Q(t_1,t_2,\ldots,t_n))^{IA} = \text{true iff } \langle t_1^{IA}, t_2^{IA}, \ldots , t_n^{IA}\rangle \in Q^I$$

For example,

- colour(lily) = white is *true* iff both refer to the same colour
- president(usa) = commander(us_army) is *true* iff both are the same person
- LessThan(5,17) is *true* iff <5,17> ∈ < relation on Natural Numbers
- Brother(suresh, ramesh) is *true* iff
  <suresh, ramesh> ∈ Brother relation on the set of people

<u>Truth Assignment to Formulas of FOL</u>

Logical connectives are interpreted in the standard way

If Val(α) = *true* then Val($\neg\alpha$) = *false*

If Val(α) = *false* then Val($\neg\alpha$) = *true*

If Val(α) = *false* and Val(β) = *false* then Val($\alpha\vee\beta$) = *false*
else Val($\alpha\vee\beta$) = *true*

If Val(α) = *true* and Val(β) = *false* then Val($\alpha\supset\beta$) = *false*
else Val($\alpha\supset\beta$) = *true*

If Val(α) = *true* and Val(β) = *true* then Val($\alpha\wedge\beta$) = *true*
else Val($\alpha\wedge\beta$) = *false*

Truth Assignment to Quantified Formulas of FOL

A formula of the form $\exists x(\alpha)$ is true if there is some value of *x* for which the formula is true. A universally quantified formula $\forall x\ (\alpha)$ is true for all possible values of *x*. Formally,

$(\exists x\ (\alpha))^{IA}$ = true iff α^{IB} is true for *some* assignment B that is an x-variant of A.

In other words the formula α is true for some value of *x*.

$(\forall x\ (\alpha))^{IA}$ = true iff α^{IB} is true for *all* assignments B that are x-variants of A.

In other words the formula α is true for all values of *x*.

An assignment *B* is said to be an x-variant of an assignment *A* if they agree on the value of all variables except *x*.

Truth Assignment to Sentences of FOL

- A sentence in FOL is a formula without any free variables.
- This means that all variables in the formula are quantified.
- As a consequence the sentences are *true* or *false* independent of the assignment mapping.

The meaning of the terms and sentences of a set of *FOL* sentences is given by an *interpretation* ϑ = <D, I>, where *D* is a domain and *I* is an interpretation mapping.

An interpretation M = <D, I> of set of sentences or a theory in a language L(*P*,*F*,*C*) is a ***model*** if all the sentences in the set are *true* in the interpretation.

Tautologies, Satisfiable and Unsatisfiable formulas

- A **tautology** is a formula of L(P,F,C) that is *true* in all interpretations.
  - For example Happy(suresh) $\vee \neg$ Happy(suresh)
- A formula of L(P,F,C) is **satisfiable** iff it is *true* in at least one interpretation.
  - For example $\forall$x (Man(x) $\supset$ Human(x))
    (depends on the meaning of Man and Human)
- A formula of L(P,F,C) is **unsatisfiable** iff it is *true* in no interpretation.
  - For example Happy(suresh) $\wedge \neg$ Happy(suresh)