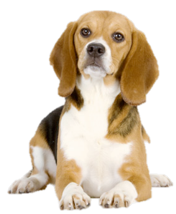
First-Order Logic & Inference

AI Class 19 (Ch. 8.1–8.3, 9)



Material from Dr. Marie desJardin, Some material adopted from notes by Andreas Geyer-Schulz and Chuck Dyer

Bookkeeping

- Midterms returned today
- HW4 due 11/7 @ 11:59

First-Order Logic

Chapter 8

First-Order Logic

- First-order logic (FOL) models the world in terms of
  - **Objects,** which are things with individual identities
  - **Properties** of objects that distinguish them from other objects
  - **Relations** that hold among sets of objects
  - **Functions,** which are a subset of relations where there is only one "value" for any given "input"
- Examples:
  - Objects: Students, lectures, companies, cars ...
  - Relations: Brother-of, bigger-than, outside, part-of, has-color, occurs-after, owns, visits, precedes, ...
  - Properties: blue, oval, even, large, ...
  - Functions: father-of, best-friend, second-half, one-more-than ...

Sentences: Terms and Atoms

- A **term** (denoting a real-world individual) is:
  - A constant symbol: *John*, or
  - A variable symbol: *x*, or
  - An n-place function of n terms

    x and $f(x_1, ..., x_n)$ are terms, where each x_i is a term

    is-a(John, Professor)
  - A term with no variables is a **ground term**.
- An **atomic sentence** is an n-place predicate of n terms
  - Has a truth value (*t* or *f*)

Sentences: Terms and Atoms

- A **complex sentence** is formed from atomic sentences connected by the logical connectives:

  $\neg P$, $P \vee Q$, $P \wedge Q$, $P \rightarrow Q$, $P \leftrightarrow Q$ where P and Q are sentences

  has-a(x, Bachelors) ∧ *is-a(x, human)*

  does NOT SAY everyone with a bachelors' is human

  has-a(John, Bachelors) ∧ *is-a(John, human)*

  has-a(Mary, Bachelors) ∧ *is-a(Mary, human)*

Quantifiers

- **Universal quantification**
  - ∀x P(x) means that P holds for **all** values of *x* in its domain
  - States universal truths
  - E.g.: $\forall x\ dolphin(x) \rightarrow mammal(x)$
- **Existential quantification**
  - ∃x P(x) means that P holds for **some** value of x in the domain associated with that variable
  - Makes a statement about some object without naming it
  - E.g., $\exists x\ mammal(x) \wedge lays\text{-}eggs(x)$

Sentences: Quantification

- **Quantified sentences** adds quantifiers ∀ and ∃

$$\forall x\ has\text{-}a(x, Bachelors) \rightarrow is\text{-}a(x, human)$$

$$\exists x\ has\text{-}a(x, Bachelors)$$

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

Everyone who has a bachelors' is human.

There exists some who has a bachelors'.

Everybody loves somebody.

Sentences: Well-Formedness

- A **well-formed formula (wff)** is a sentence containing no "free" variables. That is, all variables are "bound" by universal or existential quantifiers.
- $(\forall x)P(x,y)$ has x bound as a universally quantified variable, but y is free.

Quantifiers: Uses

- Universal quantifiers **often** used with "implies" to form "rules":

  $(\forall x)$ student(x) $\rightarrow$ smart(x)

  "All students are smart"
- Universal quantification **rarely*** used to make blanket statements about every individual in the world:

  $(\forall x)$student(x)$\wedge$smart(x)

  "Everyone in the world is a student and is smart"

*Deliberately, anyway

Quantifiers: Uses

- Existential quantifiers are **usually** used with "and" to specify a list of properties about an individual:

  $(\exists x)$ student(x) $\wedge$ smart(x)

  "There is a student who is smart"

- A common mistake is to represent this English sentence as the FOL sentence:

  $(\exists x)$ student(x) $\rightarrow$ smart(x)

  - But what happens when there is a person who is *not* a student?

Quantifier Scope

- Switching the order of universal quantifiers *does not* change the meaning:
  - $(\forall x)(\forall y)P(x,y) \leftrightarrow (\forall y)(\forall x)\ P(x,y)$
- Similarly, you can switch the order of existential quantifiers:
  - $(\exists x)(\exists y)P(x,y) \leftrightarrow (\exists y)(\exists x)\ P(x,y)$
- Switching the order of universals and existentials *does* change meaning:
  - Everyone likes someone: $(\forall x)(\exists y)$ likes(x,y)
  - Someone is liked by everyone: $(\exists y)(\forall x)$ likes(x,y)

Connections between All and Exists

We can relate sentences involving ∀ and ∃ using De Morgan's laws:

$(\forall x)\ \neg P(x) \leftrightarrow \neg(\exists x)\ P(x)$

$\neg(\forall x)\ P \leftrightarrow (\exists x)\ \neg P(x)$

$(\forall x)\ P(x) \leftrightarrow \neg\ (\exists x)\ \neg P(x)$

$(\exists x)\ P(x) \leftrightarrow \neg(\forall x)\ \neg P(x)$

Quantified Inference Rules

- Universal instantiation
  - $\forall x\ P(x)\ \therefore\ P(A)$
- Universal generalization
  - $P(A) \wedge P(B) \ldots\ \therefore\ \forall x\ P(x)$
- Existential instantiation
  - $\exists x\ P(x)\ \therefore P(F)$ ← **skolem constant F**
- Existential generalization
  - $P(A)\ \therefore\ \exists x\ P(x)$

Universal Instantiation (a.k.a. Universal Elimination)

- If (∀x) P(x) is true, then P(C) is true, where C is *any* constant in the domain of x
- Example:
  (∀x) eats(Ziggy, x) ⇒ eats(Ziggy, IceCream)
- The variable symbol can be replaced by any ground term, i.e., any constant symbol or function symbol applied to ground terms only

Existential Instantiation (a.k.a. Existential Elimination)

- Variable is replaced by a **brand-new constant**
  - I.e., not occurring in the KB
- From (∃x) P(x) infer P(c)
  - Example:
    - (∃x) eats(Ziggy, x) → eats(Ziggy, Stuff)
  - "Skolemization"
- *Stuff* is a **skolem constant**
- Easier than manipulating the existential quantifier

Existential Generalization
(a.k.a. Existential Introduction)

- If P(c) is true, then (∃x) P(x) is inferred.
- Example

  eats(Ziggy, IceCream) ⇒ (∃x) eats(Ziggy, x)
- All instances of the given constant symbol are replaced by the new variable symbol
- Note that the variable symbol cannot already exist anywhere in the expression

Translating English to FOL

Whiteboard time!

Every gardener likes the sun.

∀x gardener(x) → likes(x,Sun)

You can fool some of the people all of the time.

∃x ∀t person(x) ∧time(t) → can-fool(x,t)

You can fool all of the people some of the time.

∀x ∃t (person(x) → time(t) ∧can-fool(x,t))

∀x (person(x) → ∃t (time(t) ∧can-fool(x,t))

Equivalent

All purple mushrooms are poisonous.

∀x (mushroom(x) ∧ purple(x)) → poisonous(x)

Translating English to FOL

No purple mushroom is poisonous.

$\neg\exists x\ purple(x) \wedge mushroom(x) \wedge poisonous(x)$

$\forall x\ (mushroom(x) \wedge purple(x)) \rightarrow \neg poisonous(x)$

Equivalent

There are exactly two purple mushrooms.

$\exists x\ \exists y\ mushroom(x) \wedge purple(x) \wedge mushroom(y) \wedge purple(y) \wedge \neg(x=y) \wedge \forall z\ (mushroom(z) \wedge purple(z)) \rightarrow ((x=z) \vee (y=z))$

Clinton is not tall.

$\neg tall(Clinton)$

X is above Y iff X is on directly on top of Y or there is a pile of one or more other objects directly on top of one another starting with X and ending with Y.

$\forall x\ \forall y\ above(x,y) \leftrightarrow (on(x,y) \vee\ \exists z\ (on(x,z) \wedge above(z,y)))$

Semantics of FOL

- **Domain M:** the set of all objects in the world (of interest)
- **Interpretation I:** includes
  - Assign each constant to an object in M
  - Define each function of n arguments as a mapping $M^n \Rightarrow M$
  - Define each predicate of n arguments as a mapping $M^n \Rightarrow \{T, F\}$
  - Therefore, every ground predicate with any instantiation will have a truth value
  - In general there is an infinite number of interpretations because |M| is infinite
- **Define logical connectives:** **~, ^, v, =>, <=>** as in PL
- **Define semantics of (∀x) and (∃x)**
  - (∀x) P(x) is true iff P(x) is true under all interpretations
  - (∃x) P(x) is true iff P(x) is true under some interpretation

- **Model:** an interpretation of a set of sentences such that every sentence is *True*
- **A sentence is**
  - **satisfiable** if it is true under some interpretation
  - **valid** if it is true under all possible interpretations
  - **inconsistent** if there does not exist any interpretation under which the sentence is true
- **Logical consequence:** S |= X if all models of S are also
  models of X

Axioms, Definitions and Theorems

- **Axioms** are facts and rules that attempt to capture all of the (important) facts and concepts about a domain; axioms can be used to prove **theorems**
  - Mathematicians don't want any unnecessary (dependent) axioms –ones that can be derived from other axioms
  - Dependent axioms can make reasoning faster, however
  - Choosing a good set of axioms for a domain is a kind of design problem
- A **definition** of a predicate is of the form "p(X) ↔ …" and can be decomposed into two parts
  - **Necessary** description: "p(x) → …"
  - **Sufficient** description "p(x) ← …"
  - Some concepts don't have complete definitions (e.g., person(x))

More on Definitions

- Examples: define father(x, y) by parent(x, y) and male(x)
  - parent(x, y) is a necessary (**but not sufficient**) description of father(x, y)
    - father(x, y) → parent(x, y)
  - parent(x, y) ^ male(x) ^ age(x, 35) is a **sufficient** (**but not necessary**) description of father(x, y):

    father(x, y) ← parent(x, y) ^ male(x) ^ age(x, 35)
  - parent(x, y) ^ male(x) is a **necessary and sufficient** description of father(x, y)

    parent(x, y) ^ male(x) ↔ father(x, y)

Higher-Order Logic

- FOL only allows to quantify over variables, and variables can only range over objects.
- HOL allows us to quantify over relations
- Example: (quantify over functions)
  "two functions are equal iff they produce the same value for all arguments"
  $\forall f \, \forall g \, (f = g) \leftrightarrow (\forall x \; f(x) = g(x))$
- Example: (quantify over predicates)
  $\forall r \; \text{transitive}(r) \rightarrow (\forall xyz) \; r(x,y) \wedge r(y,z) \rightarrow r(x,z))$
- More expressive, but undecidable.

Expressing Uniqueness

- Sometimes we want to say that there is a single, unique object that satisfies a certain condition
- "There exists a unique x such that king(x) is true"
  - $\exists x\ king(x) \wedge \forall y\ (king(y) \rightarrow x=y)$
  - $\exists x\ king(x) \wedge \neg\exists y\ (king(y) \wedge x \neq y)$
  - $\exists!\ x\ king(x)$
- "Every country has exactly one ruler"
  - $\forall c\ country(c) \rightarrow \exists!\ r\ ruler(c,r)$
- Iota operator: "ι x P(x)" means "the unique x such that p(x) is true"
  - "The unique ruler of Freedonia is dead"
  - $dead(\iota\ x\ ruler(freedonia,x))$