Logical agents

First Order Logic

Outline

- Recap – Logical agents and Propositional logic
- Logical Equivalence
- Validity and satisfiability
- Inference rules for proofs
- Inference rules on Wumpus World
- First-order logic syntax
- FOL – Sentences and terms
- FOL – Quantifiers

Recap - Logical agents and Propositional Logic (PL)

- Basic concepts of logic:
  - syntax: formal structure of sentences
  - semantics: truth of sentences w.r.t. models
  - entailment: necessary truth of one sentence given another
  - inference: deriving new sentences from old ones
  - soundness: anything that is provable is actually true
  - completeness: anything that is true is provable
- Logical agents apply inference to a knowledge base to derive new information and make decisions
- Syntax for PL: propositional symbols joined with $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$
- Semantics for PL: truth values, logical consequence of $KB \models F$

Logical Equivalence $\equiv \Leftrightarrow$

- Two sentences α and β are logically equivalent if they are true in the same set of models or if they each entail the other

$\alpha \equiv \beta$ if and only if $\alpha \models \beta$ and $\beta \models \alpha$

$$
\begin{array}{rcll}
(\alpha \land \beta) & \equiv & (\beta \land \alpha) & \text{commutativity of } \land \\
(\alpha \lor \beta) & \equiv & (\beta \lor \alpha) & \text{commutativity of } \lor \\
((\alpha \land \beta) \land \gamma) & \equiv & (\alpha \land (\beta \land \gamma)) & \text{associativity of } \land \\
((\alpha \lor \beta) \lor \gamma) & \equiv & (\alpha \lor (\beta \lor \gamma)) & \text{associativity of } \lor \\
\neg(\neg\alpha) & \equiv & \alpha & \text{double-negation elimination} \\
(\alpha \Rightarrow \beta) & \equiv & (\neg\beta \Rightarrow \neg\alpha) & \text{contraposition} \\
(\alpha \Rightarrow \beta) & \equiv & (\neg\alpha \lor \beta) & \text{implication elimination} \\
(\alpha \Leftrightarrow \beta) & \equiv & ((\alpha \Rightarrow \beta) \land (\beta \Rightarrow \alpha)) & \text{biconditional elimination} \\
\neg(\alpha \land \beta) & \equiv & (\neg\alpha \lor \neg\beta) & \text{De Morgan} \\
\neg(\alpha \lor \beta) & \equiv & (\neg\alpha \land \neg\beta) & \text{De Morgan} \\
(\alpha \land (\beta \lor \gamma)) & \equiv & ((\alpha \land \beta) \lor (\alpha \land \gamma)) & \text{distributivity of } \land \text{ over } \lor \\
(\alpha \lor (\beta \land \gamma)) & \equiv & ((\alpha \lor \beta) \land (\alpha \lor \gamma)) & \text{distributivity of } \lor \text{ over } \land
\end{array}
$$

Modus Ponens

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta} .$$

AND Elimination

$$\frac{\alpha \land \beta}{\alpha} .$$

Inference Rules for Proofs

- Proof = a sequence of inference rules applications
- Modus Ponens (mode that affirms)

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta}.$$

E.g. $(WumpusAhead \wedge WumpusAlive) \Rightarrow Shoot$

$(WumpusAhead \wedge WumpusAlive)$ are given, $Shoot$ can be inferred

- And-Elimination

$$\frac{\alpha \wedge \beta}{\alpha}.$$

E.g. $(WumpusAhead \wedge WumpusAlive)$ are given, $WumpusAlive$ can be inferred

- All logical equivalences can be used as inference rules, e.g. biconditional elimination:

$$\frac{\alpha \Leftrightarrow \beta}{(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)} \quad \text{and} \quad \frac{(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)}{\alpha \Leftrightarrow \beta}.$$

Natural Deduction

- Given $P \rightarrow (Q \wedge R), S \rightarrow R, R \rightarrow P$. Prove $S \rightarrow Q$
  1. $P \rightarrow (Q \wedge R)$ Premise
  2. $S \rightarrow R$ Premise
  3. $R \rightarrow P$ Premise
  4. S Assumption
  5. R Modus Ponens 2,4
  6. P Modus Ponens 3,5
  7. $Q \wedge R$ Modus Ponens 1,6
  8. Q Conjunction ($\wedge$) elimination
  9. $S \rightarrow Q$ Implication introduction 4, 8

Recap: KB for Wumpus World

- $P_{x,y}$ is true if there is a pit in $[x,y]$.
- $W_{x,y}$ is true if there is a wumpus in $[x,y]$.
- $B_{x,y}$ is true if agent perceives breeze in $[x,y]$.
- $S_{x,y}$ is true if agent perceives stench in $[x,y]$.
- There is no pit in [1,1]

  R_1: $\neg P_{1,1}$

- A square is breezy if and only if there is a pit in a neighbouring square

  R_2: $B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$ ☚ apply biconditional elimination

  R_3: $B_{2,1} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$

- Breeze percepts

  R_4: $\neg B_{1,1}$

  R_5: $B_{2,1}$

Inference Rules and Equivalences on Wumpus World

- Biconditional elimination to R_2: $\alpha \Leftrightarrow \beta \equiv \alpha \Rightarrow \beta \wedge \beta \Rightarrow \alpha$

  R_2: $B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

  R_6: $(B_{1,1} \Rightarrow (P_{1,2} \vee P_{2,1})) \wedge ((P_{1,2} \vee P_{2,1}) \Rightarrow B_{1,1})$

- And-Elimination to R_6: $\alpha \wedge \beta$ *then* α

  R_7: $((P_{1,2} \vee P_{2,1}) \Rightarrow B_{1,1})$

- Contraposition to R_7: $\alpha \Rightarrow \beta \equiv \neg\beta \Rightarrow \neg\alpha$

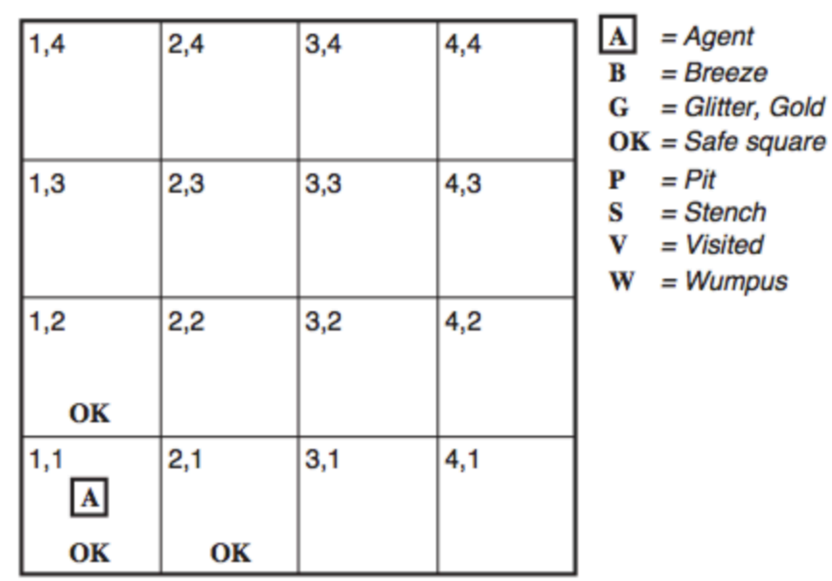
  R_8: $(\neg B_{1,1} \Rightarrow \neg(P_{1,2} \vee P_{2,1}))$

- Modus Ponens to R_8 and percept R_4: $\neg B_{1,1}$ $\alpha \Rightarrow \beta, \alpha$ *then* β

  R_9: $\neg(P_{1,2} \vee P_{2,1})$

- De Morgan's rule to R_9: $\neg(\alpha \vee \beta) \equiv (\neg\alpha \wedge \neg\beta)$

  R_{10}: $\neg P_{1,2} \wedge \neg P_{2,1}$ Neither [1,2] nor [2,1] contains a pit



Validity and Satisfiability

- A sentence is valid if it is true in all models. Also known as tautology

  e.g. $True, \quad A \vee \neg A, \quad A \Rightarrow A, \quad (A \wedge (A \Rightarrow B)) \Rightarrow B$

- A sentence is satisfiable if it is true in some model(s), e.g. in previous lecture on Wumpus World KB $(R_1 \wedge R_2 \wedge R_3 \wedge R_4 \wedge R_5)$ is satisfiable because there are 3 models in which it is true

  e.g. $A \vee B, \quad C$

- A sentence is unsatisfiable (invalid) if it is true in no models

  e.g. $A \wedge \neg A$

- Validity and satisfiability are connected:

  $\alpha \vDash \beta$ if and only if $\alpha \wedge \neg\beta$ is unsatisfiable

Propositional Logic: Pros and Cons

- ★ Declarative: pieces of syntax correspond to facts
- ★ Allows partial/disjunctive/negated information (unlike most data structures and databases)
- ★ Compositional: meaning of a sentence is function of the meaning of its parts, e.g. $B_{1,1} \wedge P_{1,2}$ are derived from the meaning of $B_{1,1}$ and of $P_{1,2}$
- ★ Meaning in PL is context-independent unlike natural language where meaning depends on context
- ◉ PL 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, e.g. $B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

First-order Logic

- Propositional logic looks at the world as containing facts (true/false)
- First-order logic (FOL) like natural language, looks at the world as containing
  - Objects (nouns) people, houses, numbers, theories, colours, Lotteria, baseball games, wars, centuries
  - Relations or properties (verb, verb phrases, adjectives) such as *red, breezy, round, bogus, etc.,* or more general *n*-ary relations such as *sister of, bigger than, inside, part of, has colour, occurred after, etc.*
  - Functions (relations with only one value for a given input) such as *mother of, best friend, third inning of, one more than, etc.*

Logics in General

Language	Ontological Commitment (What exists in the world)	Epistemological Commitment (What an agent believes about facts)
Propositional logic	facts	true/false/unknown
First-order logic	facts, objects, relations	true/false/unknown
Temporal logic	facts, objects, relations, times	true/false/unknown
Probability theory	facts	degree of belief $\in [0, 1]$
Fuzzy logic	facts with degree of truth $\in [0, 1]$	known interval value

Basics of FOL: Syntax

Constants (objects)	$KingJohn, 2, SKKU, \ldots$
Predicates (relations)	$SisterOf, >, \ldots$
Functions	$Sqrt, LeftLegOf, \ldots$
Variables	$x, y, a, b, \ldots$
Connectives	$\wedge \; \vee \; \neg \; \Rightarrow \; \Leftrightarrow$
Equality	$=$
Quantifiers	$\forall \; \exists$

FOL: Atomic Sentences

- Atomic sentence = $predicate(term_1, \ldots, term_n)$

  or $term_1 = term_2$

- Term = $function(term_1, \ldots, term_n)$

  or *constant* or *variable*

- Examples

  $BrotherOf(KingJohn, RichardTheLionheart)$

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

  Often write $>(X, Y)$ as $X > Y$

FOL: Complex Sentences

- Complex sentences are made from atomic sentences using connectives

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

- E.g.

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

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

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

FOL: Universal Quantification

$\forall$ *<variables> <sentence>* *"For all"*

- The square of any number is not negative.

  $\forall x\ (x^2 \geq 0)$

- Everyone at SKKU is smart.

  $\forall x \quad At(x, SKKU) \Rightarrow Smart(x)$

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

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

$\forall x \quad At(x, SKKU) \wedge Smart(x)$

means: "Everyone is at SKKU and everyone is smart"

FOL: Existential Quantification

$\exists$ *<variables> <sentence>* *"there exists"* *"there is at least one"*

"Someone at SKKU is smart"

$\exists x \quad At(x, SKKU) \land Smart(x)$

Typically, $\land$ is the main connective with $\exists$

✦ Common mistake: Using $\Rightarrow$ as the main connective with $\exists$

$\exists x \quad Human(x) \Rightarrow Female(x)$

$Human(Car) \Rightarrow Female(Car)$

is true if there is anything not human!

Remember $F \Rightarrow T/F$ *is TRUE*

Connections between ∀ and ∃

- Connected to each other through negation (De Morgan's laws for quantifiers)

$$
\begin{aligned}
\forall x\ \neg P &\equiv \neg \exists x\ P \\
\neg \forall x\ P &\equiv \exists x\ \neg P \\
\forall x\ P &\equiv \neg \exists x\ \neg P \\
\exists x\ P &\equiv \neg \forall x\ \neg P
\end{aligned}
$$

Statement	True When	False When
$\neg \exists x P(x) \equiv \forall x \neg P(x)$	For every x, $P(x)$ is false.	There is an x for which $P(x)$ is true.
$\neg \forall x P(x) \equiv \exists x \neg P(x)$	There is an x for which $P(x)$ is false.	$P(x)$ is true for every x.

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 (at least) one person who loves everyone in this 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
- $\forall x\ Likes(x, IceCream)$ "Everyone likes ice cream."

nobody $\neg\exists x\ \neg Likes(x, IceCream)$ "There is no one who doesn't like ice cream."

- $\exists x\ Likes(x, Broccoli)$ "Some people like broccoli."

not everyone $\neg\forall x\ \neg Likes(x, Broccoli)$ "Not everyone doesn't like broccoli."

Fun with Sentences I

- $\forall y\ \exists x\ MotherOf(x,y)$

- $\exists x\ \forall y\ MotherOf(x,y)$

- Brothers are siblings

- One’s mother is one’s female parent

Fun with Sentences II

- “Sibling” is symmetric
- A first cousin is a child of a parent’s sibling
- Susan bought everything that Frank bought
- Bob takes R or Python (but not both)

Summary

- Special formula for sentences:
  - valid/tautology True in all interpretations
  - satisfiable True is some interpretations
  - contradictory True in no interpretations
- Inference rules and equivalences can be used within propositional logic (PL) to infer new knowledge in a domain
- First-order logic
  - is more expressive than PL
  - objects and relations are semantic primitives
  - syntax: constants, functions, predicates, equality, quantifiers
- Some ~ ($\exists$), every ~ ($\forall$)

References

- Russel and Norvig, Sections 7.5, 7.51 & Chapter 8 (up to Section 8.32)

Admin

- No class on Wednesday
- Assignment 2 (Minimax) will be out this week
- Please work with your partner and submit only ONE solution script next week Wednesday
- The score will be shared for this assignment so hope you could help each other when working on it

Groups for Assignment 2

1	Michelle Halim	Kevser Basturk (Kay)	Kim Sunny 김선이
2	Royce Lim	Wei Kaichen 위해신	
3	Eugene Oon	Park Chan Il 박찬일	
4	Roosa Rauhala	Isabel Tran	
5	Kang Ju Hee 강주희	Kim Ha Neul 김하늘	
6	Noh Geon Joon 노건준	Choi Jun Seo 최준서	
7	Cho Hyun Young 조현영	Park Ji Woo 박지우	Lim Hyun Joon 임현준
8	Lee Heung Kyu 이흥규	Song Jung Hyun 송정현	
9	Je Ra Hyang 제라향	Lee Danni 이대니	
10	Cha Min Gyeong 차민경	Danae Koh	Liu Xia Tian 유하천