COMP9414: Artificial Intelligence

Lecture 4a: Knowledge Representation

Wayne Wobcke

e-mail:w.wobcke@unsw.edu.au

This Lecture

- Knowledge Representation and Logic
- Logical Arguments
- Propositional Logic
  - Syntax
  - Semantics
- Validity, Equivalence, Satisfiability, Entailment
- Inference by Natural Deduction

The Knowledge Level

Knowledge Level Hypothesis. There exists a distinct computer systems level, lying immediately above the symbol level, which is characterized by knowledge as the medium and the principle of rationality as the law of behavior.

Principle of Rationality. If an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action.

Knowledge. Whatever can be ascribed to an agent, such that its behavior can be computed according to the principle of rationality.

"The Knowledge Level" (Newell, 1982)

Knowledge Representation

- Any agent can be described on different levels
  - Knowledge level (knowledge ascribed to agent)
  - Logical level (algorithms for manipulating knowledge)
  - Implementation level (how algorithms are implemented)
- Knowledge Representation is concerned with expressing knowledge explicitly in a computer-tractable way (for use by an agent in reasoning) – not the same as Newell's view
- Reasoning attempts to take this knowledge and draw inferences (e.g. answer queries, determine facts that follow from the knowledge, decide what to do, etc.) – as part of the agent architecture

Knowledge Representation and Reasoning

- A knowledge-based agent has at its core a knowledge base
- A knowledge base is an explicit set of sentences about some domain expressed in a suitable formal representation language
  - Sentences express facts (true) **or** non-facts (false)
  - So the "knowledge base" is better called a "belief base"
- **Fundamental Questions**
  - How do we write down knowledge about a domain/problem?
  - How do we automate reasoning to deduce new facts or ensure consistency of a knowledge base?

Motivating Example – Ontologies

AfPak Ontology

- Ashraf Ghani is President Ghani – equality
- Ashraf Ghani is the President of Afghanistan – role
- Ashraf Ghani is in the government – part of
- Nangarhar is a province – a kind of
- Nangarhar is in Afghanistan – part of
- Bombing implies Attacking – linguistic meaning/semantics

Ontology = Set of such facts

Why Formal Languages – not English?

- Natural languages exhibit ambiguity

  "The fisherman went to the bank" (lexical)
  "The boy saw a girl with a telescope" (structural)
  "The table won't fit through the doorway because it is too [wide/narrow]" (co-reference)
- Ambiguity makes it difficult to interpret meaning of phrases/sentences
  - But also makes inference harder to define and compute
- Symbolic logic is a syntactically unambigious language (originally developed in an attempt to formalize mathematical reasoning)

Syntax vs Semantics

Syntax Describes the legal sentences in a knowledge representation language (e.g. in the language of arithmetic expressions $x < 4$)

Semantics Refers to the meaning of sentences. Relates sentences (and sentence fragments) to aspects of the world the sentence is about. Semantics refers to a sentence's relationship to the "real world" or to some model of the world. Semantic properties of sentences include truth and falsity (e.g. $x < 4$ is true for $x = 3$ and false when $x = 5$). Semantic properties of names and descriptions include referents.

Note: The meaning of a sentence is not intrinsic to that sentence. An interpretation is required to determine sentence meanings. Interpretations are agreed amongst a linguistic community.

Propositions

- Propositions are entities (facts or non-facts) that can be true or false
- Expressed using ordinary declarative sentences (not questions)
  - "The sky is blue" expresses the proposition that the sky is blue (here and now). Is this proposition true?
- Examples

  "Socrates is bald" (assumes 'Socrates', 'bald' are well defined)
  "The car is red" (requires 'the car' to be identified)
  "Socrates is bald and the car is red" (complex proposition)
- In Propositional Logic, use single letters to represent propositions, a scheme of abbreviation, e.g. P: Socrates is bald
- Important: Reasoning is independent of propositional substructure!

Logical Arguments

An argument relates a set of premises to a conclusion

– valid if the conclusion necessarily follows from the premises

All humans have 2 eyes
Jane is a human
Therefore Jane has 2 eyes

All humans have 4 eyes
Jane is a human
Therefore Jane has 4 eyes

- **Both** are (logically) correct **valid** arguments
- Which statements are true/false?

Logical Arguments

An argument relates a set of premises to a conclusion

– invalid if the conclusion can be false when the premises are all true

All humans have 2 eyes
Jane has 2 eyes
Therefore Jane is human

No human has 4 eyes
Jane has 2 eyes
Therefore Jane is not human

- **Both** are (logically) incorrect **invalid** arguments
- Which statements are true/false?

Propositional Logic

- Use letters to stand for "basic" propositions; combine them into more complex sentences using operators for not, and, or, implies, iff
- Propositional connectives:

$\neg$	negation	$\neg P$	"not P"
$\wedge$	conjunction	$P \wedge Q$	"P and Q"
$\vee$	disjunction	$P \vee Q$	"P or Q"
$\rightarrow$	implication	$P \rightarrow Q$	"If P then Q"
$\leftrightarrow$	bi-implication	$P \leftrightarrow Q$	"P if and only if Q"

From English to Propositional Logic

- "It is not the case that the sky is blue": $\neg B$
  (alternatively "the sky is not blue")
- "The sky is blue and the grass is green": $B \wedge G$
- "Either the sky is blue or the grass is green": $B \vee G$
- "If the sky is blue, then the grass is not green": $B \rightarrow \neg G$
- "The sky is blue if and only if the grass is green": $B \leftrightarrow G$
- "If the sky is blue, then if the grass is not green, the plants will not grow": $B \rightarrow (\neg G \rightarrow \neg P)$

Improving Readability

- $(P \rightarrow (Q \rightarrow (\neg(R)))$ vs $P \rightarrow (Q \rightarrow \neg R)$
- Rules for omitting brackets
  - Omit brackets where possible (except maybe last example below!)
  - Precedence from highest to lowest is: $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$
  - All binary operators are left associative (so $P \rightarrow Q \rightarrow R$ abbreviates $(P \rightarrow Q) \rightarrow R$)
- **Questions**
  - Is $(P \vee Q) \vee R$ (always) the same as $P \vee (Q \vee R)$?
  - Is $(P \rightarrow Q) \rightarrow R$ (always) the same as $P \rightarrow (Q \rightarrow R)$?

Truth Table Semantics

- The semantics of the connectives can be given by truth tables

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
True	True	False	True	True	True	True
True	False	False	False	True	False	False
False	True	True	False	True	True	False
False	False	True	False	False	True	True

- One row for each possible assignment of True/False to variables
- **Important:** P and Q are **any** sentences, including complex sentences

Example – Complex Sentence

R	S	$\neg R$	$R \wedge S$	$\neg R \vee S$	$(R \wedge S) \rightarrow (\neg R \vee S)$
True	True	False	True	True	True
True	False	False	False	False	True
False	True	True	False	True	True
False	False	True	False	True	True

Thus $(R \wedge S) \rightarrow (\neg R \vee S)$ is a tautology

Definitions

- A sentence is valid if it is True under all possible assignments of True/False to its variables (e.g. $P \vee \neg P$)
- A tautology is a valid sentence
- Two sentences are equivalent if they have the same truth table, e.g. $P \wedge Q$ and $Q \wedge P$
  - So P is equivalent to Q if and only if $P \leftrightarrow Q$ is valid
- A sentence is satisfiable if there is some assignment of True/False to its variables for which the sentence is True
- A sentence is unsatisfiable if it is not satisfiable (e.g. $P \wedge \neg P$)
  - Sentence is False for all assignments of True/False to its variables
  - So P is a tautology if and only if $\neg P$ is unsatisfiable

Material Implication

- $P \rightarrow Q$ evaluates to False only when P is True and Q is False
- $P \rightarrow Q$ is equivalent to $\neg P \vee Q$: material implication
- English usage often suggests a causal connection between antecedent (P) and consequent (Q) – this is not reflected in the truth table
- Examples
  - $(P \wedge Q) \rightarrow Q$ is a tautology for any Q
  - $P \rightarrow (P \vee Q)$ is a tautology for any Q
  - $(P \wedge \neg P) \rightarrow Q$ is a tautology for any Q

Logical Equivalences – All Valid

Commutativity:	$p \wedge q \leftrightarrow q \wedge p$	$p \vee q \leftrightarrow q \vee p$
Associativity:	$p \wedge (q \wedge r) \leftrightarrow (p \wedge q) \wedge r$	$p \vee (q \vee r) \leftrightarrow (p \vee q) \vee r$
Distributivity:	$p \wedge (q \vee r) \leftrightarrow (p \wedge q) \vee (p \wedge r)$	$p \vee (q \wedge r) \leftrightarrow (p \vee q) \wedge (p \vee r)$
Implication:	$(p \rightarrow q) \leftrightarrow (\neg p \vee q)$	
Idempotent:	$p \wedge p \leftrightarrow p$	$p \vee p \leftrightarrow p$
Double negation:	$\neg\neg p \leftrightarrow p$	
Contradiction:	$p \wedge \neg p \leftrightarrow \text{FALSE}$	
Excluded middle:		$p \vee \neg p \leftrightarrow \text{TRUE}$
De Morgan:	$\neg(p \wedge q) \leftrightarrow (\neg p \vee \neg q)$	$\neg(p \vee q) \leftrightarrow (\neg p \wedge \neg q)$

Proof of Equivalence

Let $P \Leftrightarrow Q$ mean "P is equivalent to Q" ($P \Leftrightarrow Q$ is **not** a formula)

Then $P \wedge (Q \rightarrow R) \quad \Leftrightarrow \quad \neg(P \rightarrow Q) \vee (P \wedge R)$

$$
\begin{array}{rcll}
P \wedge (Q \rightarrow R) & \Leftrightarrow & P \wedge (\neg Q \vee R) & \text{[Implication]} \\
& \Leftrightarrow & (P \wedge \neg Q) \vee (P \wedge R) & \text{[Distributivity]} \\
& \Leftrightarrow & (\neg\neg P \wedge \neg Q) \vee (P \wedge R) & \text{[Double negation]} \\
& \Leftrightarrow & \neg(\neg P \vee Q) \vee (P \wedge R) & \text{[De Morgan]} \\
& \Leftrightarrow & \neg(P \rightarrow Q) \vee (P \wedge R) & \text{[Implication]}
\end{array}
$$

Assumes substitution: if $A \Leftrightarrow B$, replace A by B in any subformula

Assumes equivalence is transitive: if $A \Leftrightarrow B$ and $B \Leftrightarrow C$ then $A \Leftrightarrow C$

Entailment

- S entails P ($S \models P$) if whenever all formulae in S are True, P is True
  - Semantic definition – concerns truth (not proof)
- Compute whether $S \models P$ by calculating a truth table for S and P
  - Syntactic notion – concerns computation/proof
  - Not always this easy to compute (how inefficient is this?)
- A tautology is a special case of entailment where S is the empty set
  - All rows of the truth table are True

Simple Entailments

Write $P \models Q$ for $\{P\} \models Q$

$P \wedge Q \models P$ $\qquad$ $P \wedge Q \models Q$

$P \models P \vee Q$ $\qquad$ $Q \models P \vee Q$

$P \models \neg\neg P$ $\qquad$ $\neg\neg P \models P$

$\{P, P \rightarrow Q\} \models Q$

If $P \models Q$ then $\models P \rightarrow Q$

Entailment Example

P	Q	$P \rightarrow Q$	Q
True	True	True	True
True	False	False	False
False	True	True	True
False	False	True	False

Therefore $\{P, P \rightarrow Q\} \models Q$

- In the only row where both P and $P \rightarrow Q$ are True (row 1), Q is also True (here S is the set $\{P, P \rightarrow Q\}$)

Note: The column for $P \rightarrow Q$ is calculated from that for P and Q using the truth table definition, and Q is used again to check the entailment

Entailment – Tautology

R	S	$\neg R$	$R \wedge S$	$\neg R \vee S$	$(R \wedge S) \rightarrow (\neg R \vee S)$
True	True	False	True	True	True
True	False	False	False	False	True
False	True	True	False	True	True
False	False	True	False	True	True

Therefore $\models (R \wedge S) \rightarrow (\neg R \vee S)$

Models

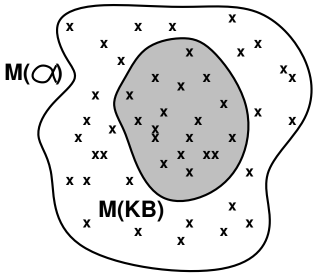
Can also think in terms of models, formally structured interpretations with respect to which truth is evaluated

- For Propositional Logic, a model is one row of the truth table

A model M is a model of a sentence α if α is True in M

Let $M(\alpha)$ be the set of all models of α

Then $KB \models \alpha$ if and only if $M(KB) \subseteq M(\alpha)$



Natural Deduction Proofs

Logical Rules of Inference

Negations

$$\frac{\varphi \Rightarrow \psi \quad \varphi \Rightarrow \neg\psi}{\neg\varphi} \qquad \frac{\neg\neg\varphi}{\varphi}$$

Implications

$$\frac{\varphi \vdash \psi}{\varphi \Rightarrow \psi} \qquad \frac{\varphi \Rightarrow \psi \quad \varphi}{\psi}$$

Biconditionals

$$\frac{\varphi \Rightarrow \psi \quad \psi \Rightarrow \varphi}{\varphi \Leftrightarrow \psi} \qquad \frac{\varphi \Leftrightarrow \psi}{\varphi \Rightarrow \psi \quad \psi \Leftrightarrow \varphi}$$

Conjunctions

$$\frac{\varphi_1 \quad \ldots \quad \varphi_n}{\varphi_1 \wedge \ldots \wedge \varphi_n} \qquad \frac{\varphi_1 \wedge \ldots \wedge \varphi_n}{\varphi_1 \quad \ldots \quad \varphi_n}$$

Disjunctions

$$\frac{\varphi_i}{\varphi_1 \vee \ldots \vee \varphi_1} \qquad \frac{\varphi_1 \vee \ldots \vee \varphi_n \quad \varphi_1 \Rightarrow \psi \quad \ldots \quad \varphi_n \Rightarrow \psi}{\psi}$$

Notes: $\vdash$ means proof; $\Rightarrow$ is our $\rightarrow$

Natural Deduction Example

$$\dfrac{\dfrac{\dfrac{\dfrac{\overline{A \to (B \to C)}^{\,2} \qquad \dfrac{\overline{A \wedge B}^{\,1}}{A}}{B \to C} \qquad \dfrac{\overline{A \wedge B}^{\,1}}{B}}{C}}{A \wedge B \to C}\;1}{(A \to (B \to C)) \to (A \wedge B \to C)}\;2$$

Conclusion

- Ambiguity of natural languages avoided with formal languages
- Enables formalization of (truth preserving) entailment
- Propositional Logic: Simplest logic of truth and falsity
- Knowledge Based Systems: First-Order Logic
- Automated Reasoning: How to compute entailment (inference)
- Many many logics not studied in this course