Knowledge - Lecture 1

knowledge-based agents: agents that reason by operating on internal representations of knowledge

sentence: an assertion about the world in a knowledge representation language

propositional logic:

propositional symbols: P, Q, R - facts

logical connectives: $\neg$ (not) $\wedge$ (and) $\vee$ (or) $\rightarrow$ (implication) $\leftrightarrow$ (biconditional)

not

P	$\neg P$
false	true
true	false

biconditional

P	Q	$P \leftrightarrow Q$
false	false	true
false	true	false
true	false	false
true	true	true

and

P	Q	$P \wedge Q$
false	false	false
false	true	false
true	false	false
true	true	true

or

P	Q	$P \vee Q$
false	false	false
false	true	true
true	false	true
true	true	true

implication

P	Q	$P \rightarrow Q$
* false	false	true
* false	true	true
true	false	false
true	true	true

* makes no claim

model: assignment of a truth value to every propositional symbol (a "possible world")

knowledge base: a set of sentences known by a knowledge based agent

entailment: in every model in which sentence α is true, β is also true

$$\alpha \vDash \beta$$

inference: the process of deriving new sentences from old ones

Model Checking

- To determine if $\mathrm{KB} \vDash \alpha$:
  - Enumerate all possible models.
  - If in every model where KB is true, α is true, then KB entails α.
  - Otherwise, KB does not entail α.

P: It is a Tuesday. Q: It is raining. R: Harry will go for a run.

KB: $(P \wedge \neg Q) \rightarrow R$ $\quad P \quad \neg Q$

Query: R

P	Q	R	KB
false	false	false	F
false	false	true	F
false	true	false	F
false	true	true	F
true	false	false	F
true	false	true	T
true	true	false	F
true	true	true	F

Knowledge Engineering: describe real world problems with logical symbols

Logic puzzles can be solved by models

Model checking isn't very efficient

Inference Rules

Modus Ponens	$\alpha \to \beta ; \alpha$	β
And Elimination	$\alpha \land \beta$	α
Double Negation Elimination	$\neg(\neg\alpha)$	α
Implication Elimination	$\alpha \to \beta$	$\neg\alpha \lor \beta$
Biconditional Elimination	$\alpha \leftrightarrow \beta$	$(\alpha \to \beta) \land (\beta \to \alpha)$
De Morgan's Law	$\neg(\alpha \land \beta)$	$\neg\alpha \lor \neg\beta$
Distributive Law	$(\alpha \land (\beta \lor \gamma))$ $(\alpha \lor (\beta \land \gamma))$	$(\alpha \land \beta) \lor (\alpha \land \gamma)$ $(\alpha \lor \beta) \land (\alpha \lor \gamma)$
	$P \lor Q, \neg P$	Q
	$P \lor Q, \neg P \lor R$	$Q \lor R$

clause: a disjunction of literals

connected with OR conjunction = AND

conjunctive normal form: logical sentence that is a conjunction of clauses

Theorem Proving

- initial state: starting knowledge base
- actions: inference rules
- transition model: new knowledge base after inference
- goal test: check statement we're trying to prove
- path cost function: number of steps in proof

Conversion to CNF conjuctive normal form

- Eliminate biconditionals
  - turn $(\alpha \leftrightarrow \beta)$ into $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$
- Eliminate implications
  - turn $(\alpha \rightarrow \beta)$ into $\neg\alpha \vee \beta$
- Move $\neg$ inwards using De Morgan's Laws
  - e.g. turn $\neg(\alpha \wedge \beta)$ into $\neg\alpha \vee \neg\beta$
- Use distributive law to distribute $\vee$ wherever possible

$(P \vee Q) \rightarrow R$

$(\neg P \vee R) \wedge (\neg Q \vee R)$

$P \vee Q \vee S \quad \neg P \vee R \vee S$

$(Q \vee R \vee S)$

eliminate duplicate
eliminate opposite

P | ¬P | ( ) empty

Inference by Resolution

- To determine if $\mathrm{KB} \vDash \alpha$:
  - Convert $(\mathrm{KB} \wedge \neg\alpha)$ to Conjunctive Normal Form.
  - Keep checking to see if we can use resolution to produce a new clause.
    - If ever we produce the **empty** clause (equivalent to False), we have a contradiction, and $\mathrm{KB} \vDash \alpha$. ← yaay :3
    - Otherwise, if we can't add new clauses, no entailment.

Does $(A \vee B) \wedge (\neg B \vee C) \wedge (\neg C)$ entail A?

$$(A \vee B) \wedge (\neg B \vee C) \wedge (\neg C) \wedge (\neg A)$$

$$(A \vee B) \quad (\neg B \vee C) \quad (\neg C) \quad (\neg A) \quad (\neg B) \quad (A) \quad (\,)$$

← yaay :3

First Order Logic

$Person(Minerva)$	Minerva is a person.
$House(Gryffindor)$	Gryffindor is a house.
$\neg House(Minerva)$	Minerva is not a house.
$BelongsTo(Minerva, Gryffindor)$	Minerva belongs to Gryffindor.

Universal Quantification

$\forall x.\ BelongsTo(x, Gryffindor) \rightarrow \neg BelongsTo(x, Hufflepuff)$

For all objects x, if x belongs to Gryffindor, then x does not belong to Hufflepuff.

Anyone in Gryffindor is not in Hufflepuff.

Existential Quantification

$\exists x.\ House(x) \wedge BelongsTo(Minerva, x)$

There exists an object x such that x is a house and Minerva belongs to x.

Minerva belongs to a house.