UIT2504
Artificial Intelligence

Derivations in Propositional Logic

Logic in General

- Sentences written in logic must be well-formed formula and follow a grammar
- There are several possible interpretations for a set of sentences KB
- Interpretations in which KB evaluates to true are called as models of KB
- Given a new sentence α, KB logically entails α (written as $KB \vDash \alpha$) iff every model of KB is also a model of α
- We write $KB \vdash \alpha$ if α can be derived from KB using syntactic derivation rules
- Sentences in logic are usually written in a normal form
- There may be several strategies for effective application of the derivation rules

Derivations in Logic

- KB $\vdash$ α if α can be **derived** from KB using syntactic **inference rules**

$$x + y = 4$$

$$x + y - y = 4 - y$$

$$x = 4 - y$$

- Inference procedure is **sound** if every α derivable is entailed by KB
- Inference procedure is **complete** if every α that is entailed by KB can be derived from KB
- When we have sound and complete inference procedure, KB $\overset{?}{\models}$ α can be reduced to KB $\overset{?}{\vdash}$ α

Properties of Sentences

- A sentence KB is satisfiable if it has a model
  (unsatisfiable if it has no model)
- KB is a valid sentence (tautology) if it is true in all the interpretations
  (invalid if it is false in at least one interpretation)
- Two sentences KB_1 and KB_2 are equivalent if they have same set of models

Some Popular Equivalences

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

Derivations

- Given a set of sentences KB, new sentences can be derived using inference rules
- Equivalences shown before can be used as inference rules
- Example: Double Negation Elimination

$$\frac{\neg\neg\alpha}{\alpha}$$

Derivations

- Given a set of sentences KB, new sentences can be derived using inference rules
- Equivalences shown before can be used as inference rules
- Example: Double Negation Elimination

$$\frac{\neg\neg\alpha}{\alpha}$$

Already available

Derivations

- Given a set of sentences KB, new sentences can be derived using inference rules
- Equivalences shown before can be used as inference rules
- Example: Double Negation Elimination

Already available

$$\frac{\neg\neg\alpha}{\alpha}$$

Can be derived

Inference Rules

- Modus Ponens ($\Rightarrow$-Elimination)

$$\frac{P \Rightarrow Q \qquad\qquad P}{Q}$$

Inference Rules

- Modus Ponens ($\Rightarrow$-Elimination)

$$\frac{P \Rightarrow Q \qquad P}{Q}$$

- $\wedge$-elimination and $\wedge$-intro

$$\frac{P \wedge Q}{P} \qquad \frac{P \wedge Q}{Q}$$

$$\frac{P \qquad Q}{P \wedge Q}$$

Inference Rules

- $\Leftrightarrow$-elim and $\Leftrightarrow$-intro

$$\frac{P \Leftrightarrow Q}{P \Rightarrow Q} \qquad \frac{P \Leftrightarrow Q}{Q \Rightarrow P} \qquad \frac{P \Rightarrow Q \quad Q \Rightarrow P}{P \Leftrightarrow Q}$$

Inference Rules

- $\Leftrightarrow$-elim and $\Leftrightarrow$-intro

$$\frac{P \Leftrightarrow Q}{P \Rightarrow Q} \qquad \frac{P \Leftrightarrow Q}{Q \Rightarrow P} \qquad \frac{P \Rightarrow Q \quad\quad Q \Rightarrow P}{P \Leftrightarrow Q}$$

- $\vee$-intro and $\vee$-elim

$$\frac{P}{P \vee Q}$$

$$\frac{P \vee Q \quad\quad \text{Assume } P\text{, Derive } R \quad\quad \text{Assume } Q\text{, Derive } R}{R}$$

Inference Rules

- $\Rightarrow$-intro

$$\frac{\text{Assume } P \text{ and derive } Q}{P \Rightarrow Q}$$

Inference Rules

- $\Rightarrow$-intro

$$\frac{\text{Assume P and derive Q}}{P \Rightarrow Q}$$

- $\neg$-intro

$$\frac{\text{Assume P and derive contradiction}}{\neg P}$$

Inference Rules

- $\Rightarrow$–intro

$$\frac{\text{Assume P and derive Q}}{P \Rightarrow Q}$$

- $\neg$–intro

$$\frac{\text{Assume P and derive contradiction}}{\neg P}$$

- $\neg$–elim

$$\frac{\text{Assume } \neg P \text{ and derive contradiction}}{P}$$

Inference Rules

- $\Rightarrow$–intro

$$\frac{\text{Assume P and derive Q}}{P \Rightarrow Q}$$

- $\neg$–intro

$$\frac{\text{Assume P and derive contradiction}}{\neg P}$$

- $\neg$–elim

$$\frac{\text{Assume } \neg P \text{ and derive contradiction}}{P}$$

- Contradiction

$$\frac{P \qquad \neg P}{\square}$$

Derivations

- Given $(P \Rightarrow W)$, $(W \Rightarrow H)$, W
- Derive H

Derivations

- Given $(P \Rightarrow W)$, $(W \Rightarrow H)$, W
- Derive H

1. $P \Rightarrow W$	Premise
2. $W \Rightarrow H$	Premise
3. W	Premise

Derivations

- Given $(P \Rightarrow W)$, $(W \Rightarrow H)$, W
- Derive H

1. $P \Rightarrow W$	Premise
2. $W \Rightarrow H$	Premise
3. W	Premise
4. H	Modus Ponens, 2, 3

Derivations

- Given $(P \Rightarrow Q)$, $(Q \Rightarrow R)$
- Derive $P \Rightarrow R$

Derivations

- Given $(P \Rightarrow Q), (Q \Rightarrow R)$
- Derive $P \Rightarrow R$

1. $P \Rightarrow Q$	Premise
2. $Q \Rightarrow R$	Premise

Derivations

- Given $(P \Rightarrow Q)$, $(Q \Rightarrow R)$
- Derive $P \Rightarrow R$

1. $P \Rightarrow Q$ Premise
2. $Q \Rightarrow R$ Premise
3. Assume P, Derive R

4. $P \Rightarrow R$ $\Rightarrow$-intro, 3

Derivations

- Given $(P \Rightarrow Q)$, $(Q \Rightarrow R)$
- Derive $P \Rightarrow R$

1. $P \Rightarrow Q$ Premise
2. $Q \Rightarrow R$ Premise
3. Assume P, Derive R

3.1 P Assumption

4. $P \Rightarrow R$ $\Rightarrow$–intro, 3

Derivations

- Given $(P \Rightarrow Q)$, $(Q \Rightarrow R)$
- Derive $P \Rightarrow R$

1. $P \Rightarrow Q$	Premise
2. $Q \Rightarrow R$	Premise
3. Assume P, Derive R	
3.1 P	Assumption
3.2 Q	Modus Ponens, 3.1, 1
4. $P \Rightarrow R$	$\Rightarrow$–intro, 3

Derivations

- Given $(P \Rightarrow Q)$, $(Q \Rightarrow R)$
- Derive $P \Rightarrow R$

1. $P \Rightarrow Q$	Premise
2. $Q \Rightarrow R$	Premise
3. Assume P, Derive R	
3.1 P	Assumption
3.2 Q	Modus Ponens, 3.1, 1
3.3 R	Modus Ponens, 3.2, 2
4. $P \Rightarrow R$	$\Rightarrow$–intro, 3

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise
2. Assume $(P \wedge \neg Q)$, Derive Contradiction

3. $\neg(P \wedge \neg Q)$ $\neg$-intro, 2

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise

2. Assume $(P \wedge \neg Q)$, Derive Contradiction

2.1 $P \wedge \neg Q$ Assumption

3. $\neg(P \wedge \neg Q)$ $\neg$-intro, 2

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise

2. Assume $(P \wedge \neg Q)$, Derive Contradiction

2.1 $P \wedge \neg Q$ Assumption

2.2 P $\wedge$-elim, 2.1

3. $\neg(P \wedge \neg Q)$ $\neg$-intro, 2

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise

2. Assume $(P \wedge \neg Q)$, Derive Contradiction

2.1 $P \wedge \neg Q$ Assumption

2.2 P $\wedge$-elim, 2.1

2.3 $\neg Q$ $\wedge$-elim, 2.1

3. $\neg(P \wedge \neg Q)$ $\neg$-intro, 2

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise
2. Assume $(P \wedge \neg Q)$, Derive Contradiction

2.1 $P \wedge \neg Q$ Assumption

2.2 P $\wedge$-elim, 2.1

2.3 $\neg Q$ $\wedge$-elim, 2.1

2.4 Q Modus Ponens, 1, 2.2

3. $\neg(P \wedge \neg Q)$ $\neg$-intro, 2

Derivations

- Given $(P \Rightarrow Q)$
- Derive $\neg(P \wedge \neg Q)$

1. $P \Rightarrow Q$ Premise

2. Assume $(P \wedge \neg Q)$, Derive Contradiction

2.1 $P \wedge \neg Q$ Assumption

2.2 P $\wedge$–elim, 2.1

2.3 $\neg Q$ $\wedge$–elim, 2.1

2.4 Q Modus Ponens, 1, 2.2

2.5 $\square$ Contradiction, 2.3, 2.4

3. $\neg(P \wedge \neg Q)$ $\neg$–intro, 2

Normal Forms

- **Disjunctive normal form**
  - Disjunction of conjunction of literals

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

Normal Forms

- **Disjunctive normal form**
  - Disjunction of conjunction of literals

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

- **Conjunctive normal form**
  - Conjunction of disjunction of literals

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

Normal Forms

- **Disjunctive normal form**
  - Disjunction of conjunction of literals

  $(P \land Q) \lor (\neg Q \land R) \lor (\neg S \land T)$

- **Conjunctive normal form**
  - Conjunction of disjunction of literals

  $(\neg P \lor Q) \land (P \lor R) \land (S \lor Q)$

- Any well-formed sentence can be transformed into its equivalent CNF

Conversion to CNF

- Eliminate $\Leftrightarrow$

  $P \Leftrightarrow Q$ is same as $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$
- Eliminate $\Rightarrow$

  $P \Rightarrow Q$ is same as $\neg P \vee Q$
- Move $\neg$ close to atoms

  Use double-negation elimination and de Morgan laws
- Use the distributive laws to rearrange the sentences in CNF

CNF: Example

Example: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

CNF: Example

Example: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$

$(B_{1,1} \Rightarrow (P_{1,2} \lor P_{2,1})) \land ((P_{1,2} \lor P_{2,1}) \Rightarrow B_{1,1})$

$(\neg B_{1,1} \lor P_{1,2} \lor P_{2,1}) \land (\neg(P_{1,2} \lor P_{2,1}) \lor B_{1,1})$

$(\neg B_{1,1} \lor P_{1,2} \lor P_{2,1}) \land ((\neg P_{1,2} \land \neg P_{2,1}) \lor B_{1,1})$

$(\neg B_{1,1} \lor P_{1,2} \lor P_{2,1}) \land (\neg P_{1,2} \lor B_{1,1}) \land (\neg P_{2,1} \lor B_{1,1})$

Clauses

- A **clause** is a disjunction of literals
- A sentence in conjunctive normal form (conjunction of clauses) can be expressed as a set of clauses

  $KB = \{\neg P \vee Q, P\}$

- A **Horn Clause** is a clause with at most one positive literal
- A Horn clause with exactly one positive literal is a **definite clause**
- A clause may also be written in implicative form. For example,

  $KB = \{ P \Rightarrow Q, P\}$

Questions?