

Artificial Intelligence

Automated Reasoning

Andrea Torsello

Automated Reasoning

Very important area of AI research

“Reasoning” usually means deductive reasoning

- New facts are deduced logically from old ones

Inductive reasoning (later in course)

- Guessing facts from old ones and from evidence

Two main aspects of deductive reasoning

- Logical representations (thousands of them)
- Rules of deduction (how to deduce new things)

Applications of Automated Reasoning

Automated theorem proving

- Automated mathematics
- Axioms(A) are given, theorem statement(T) is given
- Reasoning agent searches from A to T (or from T to A)
 - Using rules of deduction to move around the search space

Automated verification

- Hardware and Software verification
 - That they perform as specified
- Remember the Intel chip fiasco?
 - Intel now have lots of people working on automated verification

Automating Deductive Reasoning

Aims of automated deduction

- Deduce new knowledge from old
- Prove/disprove some open conjectures

Theorem proving

- Search for a path from axioms to theorem statement
- Operators are (sound) inference rules

Applications:

- Agents that use deductive inference
- Mechanising and automating mathematics
- Verifying hardware and software specifications
- The semantic web

Inference Rules

A entails B iff

- B is true when A is true
- Any model of A is a model of B

Then this is a sound inference rule

$$\frac{A}{B}$$

Axioms $\Rightarrow$ C $\Rightarrow$ D $\Rightarrow$... $\Rightarrow$ Z $\Rightarrow$ Theorem

- Each step is application of inference rule
- Theorem is entailed by the axioms

Tautologies

S: $(X \rightarrow (Y \wedge Z)) \leftrightarrow ((X \rightarrow Y) \wedge (X \rightarrow Z))$

- Show that no matter what truth values for X, Y and Z
 - The statement S is always true

X	Y	Z	$Y \wedge Z$	$X \rightarrow Y$	$X \rightarrow Z$	$X \rightarrow (Y \wedge Z)$	$((X \rightarrow Y) \wedge (X \rightarrow Z))$	S
true	true	true	true	true	true	true	true	true
true	true	false	false	true	false	false	false	true
true	false	true	false	false	true	false	false	true
true	false	false	false	false	false	false	false	true
false	true	true	true	true	true	true	true	true
false	true	false	false	true	true	true	true	true
false	false	true	false	true	true	true	true	true
false	false	false	false	true	true	true	true	true

Columns 7 and 8 are always the same

Inference with Tautologies

$P \wedge Q \leftrightarrow Q \wedge P$ is obviously true

- Regardless of meaning or truth values of P and Q
- This is content-free and a tautology

One way to define a rule of inference:

- We can replace $P \wedge Q$ with $Q \wedge P$, and vice versa
- They are true for same models
- Replacing one for other preserves soundness

Equivalence Rules

A and B are logically equivalent (write $A \equiv B$)

- Same models for each
- Can replace any instance of A with an instance of B without affecting models

Formalised as rewrite rule $A \Rightarrow B$

- Also $B \Rightarrow A$
- Must avoid looping $A \Rightarrow B \Rightarrow A \Rightarrow B \Rightarrow \dots$
- Choose one direction, or always loop-check

Rewrite rules used for inference

- Showing theorem and axioms are equivalent
- Preprocessing theorem/axioms into a particular format

Properties

Commutativity

$$P \wedge Q \equiv Q \wedge P$$

$$P \vee Q \equiv Q \vee P$$

$$P \leftrightarrow Q \equiv Q \leftrightarrow P$$

Associativity

$$(P \wedge Q) \wedge R \equiv P \wedge (Q \wedge R)$$

$$(P \vee Q) \vee R \equiv P \vee (Q \vee R)$$

Distributivity over 'and' and 'or':

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

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

Distributivity over implication

$$P \rightarrow (Q \vee R) \equiv (P \rightarrow Q) \vee (P \rightarrow R)$$

$$P \rightarrow (Q \wedge R) \equiv (P \rightarrow Q) \wedge (P \rightarrow R)$$

Properties

De Morgan's Law (refers to either)

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

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

Contraposition: imagine the opposite is true

$$P \rightarrow Q \equiv \neg Q \rightarrow \neg P$$

Often useful in mathematics proof

Remove implication or equivalence (very useful)

$$P \rightarrow Q \equiv \neg P \vee Q$$

$$P \leftrightarrow Q \equiv (P \rightarrow Q) \wedge (Q \rightarrow P)$$

Reduce to truth value

$$P \wedge \neg P \equiv \text{False}$$

$$P \vee \neg P \equiv \text{True}$$

An Example Deduction

$$(P \leftrightarrow Q) \wedge (\neg P \wedge Q)$$

Show that this sentence is false

- Show that this rewrites to False
- This proves the negation

1. Using the double negation rewrite: $P \Rightarrow \neg\neg P$

$$(P \leftrightarrow Q) \wedge (\neg P \wedge \neg\neg Q)$$

2. Using De Morgan's Law: $\neg P \wedge \neg Q \Rightarrow \neg(P \vee Q)$

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

3. Using the commutativity of $\vee$: $P \vee Q \Rightarrow Q \vee P$

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

4. Using 'replace implication' from right to left: $\neg P \vee Q \Rightarrow P \rightarrow Q$

$$(P \leftrightarrow Q) \wedge \neg(Q \rightarrow P)$$

5. Using 'replace equivalence' from left to right: $P \leftrightarrow Q \Rightarrow (P \rightarrow Q) \wedge (Q \rightarrow P)$

$$((P \rightarrow Q) \wedge (Q \rightarrow P)) \wedge \neg(Q \rightarrow P)$$

6. Using the associativity of $\wedge$: $(P \wedge Q) \wedge R \Rightarrow P \wedge (Q \wedge R)$

$$(P \rightarrow Q) \wedge ((Q \rightarrow P) \wedge \neg(Q \rightarrow P))$$

7. Using the consistency equivalence above: $P \wedge \neg P \Rightarrow \text{False}$

$$(P \rightarrow Q) \wedge \text{False}$$

8. Using the definition of $\wedge$:

$$\text{False}$$

Propositional Inference Rules

Rewrite rules are good for bidirectional search

- But we don't need equivalence, just entailment

Classic example

- All men are mortal, socrates is a man
- Therefore: Socrates is mortal

This is an instance of an inference rule

- Known as **Modus Ponens** (Aristotle)

$$\frac{A \rightarrow B, \quad A}{B}$$

Above line: what we know, below: what we can deduce

Soundness of Modus Ponens

A	B	$A \rightarrow B$	Top: $A \rightarrow B, A$	Bottom: B
True	True	True	True	True
True	False	False	False	False
False	True	True	False	True
False	False	True	False	True

And-Elimination & -Introduction

And-Elimination:

$$\frac{A_1 \wedge A_2 \wedge \dots \wedge A_n}{A_i}$$

$$[1 \leq i \leq n]$$

And-Introduction:

$$\frac{A_1, A_2, \dots, A_n}{A_1 \wedge A_2 \wedge \dots \wedge A_n}$$

The sentences may be from different places

- Selected from the database

Or-Introduction & Unit Resolution

Or-introduction

$$\frac{A_i}{A_1 \vee A_2 \vee \dots \vee A_n}$$

$$[1 \leq i \leq n]$$

Unit resolution

$$\frac{(A \vee B), \neg B}{A}$$

Basis for resolution theorem proving

Substitution & Instantiation

FOL sentences have quantified variables

- Substitute into a variable by assigning a particular value
- Replace with given term, remove quantifier

Instantiation (grounding) is a kind of substitution

- Must substitute a ground term

Example: $\exists X.Y.\text{likes}(X,Y)$ becomes $\text{likes}(\text{tony}, \text{george})$

We write:

- $\text{Subst}(\{X/\text{tony}, Y/\text{george}\}, \text{likes}(X,Y)) = \text{likes}(\text{tony}, \text{george})$

Universal Elimination

Given a sentence, A

- Containing a universally quantified variable V
- Then we can replace V by any ground term g

$$\frac{\forall V.A}{\text{Subst}(\{V/g\}, A)}$$

Remember to remove quantifier

- Not as complicated as it looks:

$\forall X \text{ likes}(X, \text{ice_cream})$ becomes $\text{likes}(\text{ben}, \text{ice_cream})$

Existential Elimination

Given a sentence, A

- Containing an existentially quantified variable, V
- Then we can replace V by any constant, k
- As long as k is not mentioned anywhere else

$$\frac{\exists V.A}{\text{Subst}(\{V/k\}, A)}$$

For the sake of argument, let's call it...

Existential Introduction

Given a sentence, A

- And a variable, V , **which is not used in A**
- Then any ground term, g , in A can be substituted by V
 - As long as g does not appear in A also

A

$\exists V. \text{Subst}(\{g/V\}, A)$

Exercise: find sentence where V is in A such that this inference rule is not sound

Chains of Inference

Remember the problem we're trying to solve

- Search for a path from axioms, A , to theorem, T

Three approaches

- Forward chaining
- Backward chaining
- Proof by contradiction

Specification of a search problem:

- Representation of states (first-order logic sentences)
- Initial state (depends...)
- Operators (inference rules, including equivalences)
- Goal state (depends...)

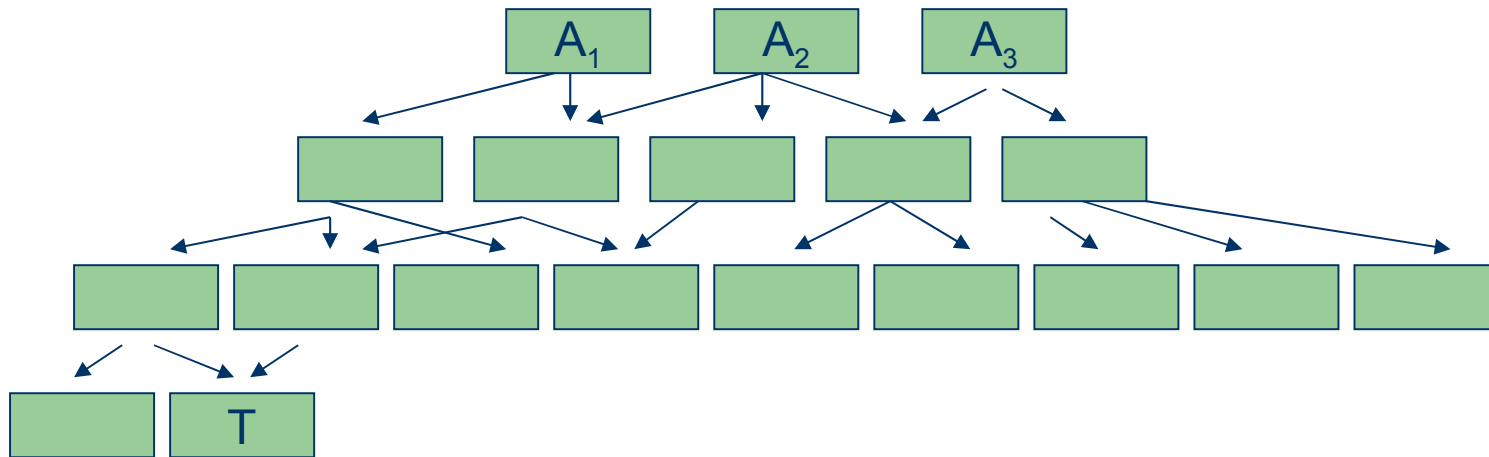
Forward Chaining

Deduce new facts from axioms

- Deduce new facts from these, etc.,...

Hopefully end up deducing the theorem statement

Can take a long time: not using the goal to direct search



Backward Chaining

Start with the theorem state and work backwards

- Hope to end up at the axioms

Each step asks: given the state that I'm at...

- Which operator could have been applied to which state to produce the state (sentence) I'm at

No problem when using equivalences

- Can also use a bidirectional search (from both ends)

Difficult when using general inference rules

- Many possible ways to invert operators

“Reductio ad absurdum”

Assume theorem is false

- then show that the assumption contradicts the axioms
- which proves that the theorem is true

Add negated theorem to the set of axioms

- See if we can deduce the ‘False’ sentence

Advantage: using the theorem statement from start

- Can look to see how close we are to the false statement
- Possibilities for a heuristic search!

Basis for resolution theorem proving (next week)