CSE 473 Logic in AI

Dan Weld

(With some slides from Mausam, Stuart Russell, Dieter Fox, Henry Kautz…)

There is nothing so powerful as truth, and often nothing so strange.

- Daniel Webster (1782-1852)

Overview

– Introduction & Agents
– Search, Heuristics & CSPs
– Adversarial Search
– Logical Knowledge Representation
– Planning & MDPs
– Reinforcement Learning
– Uncertainty & Bayesian Networks
– Machine Learning
– NLP & Special Topics

KR Hypothesis

Any ***intelligent process*** will have ingredients that

1) We as external observers interpret as knowledge
2) This knowledge plays a formal, causal & essential role in guiding the behavior

- Brian Smith (paraphrased)

Some KR Languages

- Propositional Logic
- Predicate Calculus
- Frame Systems
- Rules with Certainty Factors
- Bayesian Belief Networks
- Influence Diagrams
- Semantic Networks
- Concept Description Languages
- Non-monotonic Logic

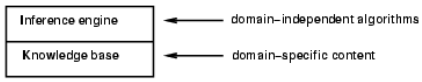
Knowledge Representation

- *represent knowledge in a manner that facilitates inferencing (i.e. drawing conclusions) from knowledge.*
- Typically based on
  – Logic
  – Probability
  – Logic and Probability

Basic Idea of Logic

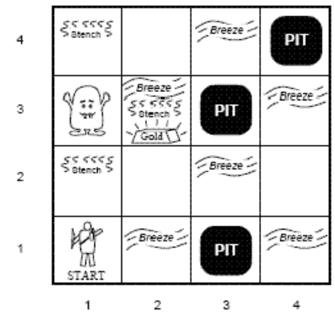
- By starting with true assumptions, you can deduce true conclusions.

Knowledge bases



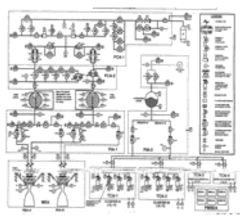
- Knowledge base = set of sentences in a formal language
- Declarative approach to building an agent (or other system):
  - `Tell` it what it needs to know
- Then it can `Ask` itself what to do - answers should follow from the KB
- Agents can be viewed at the knowledge level
  i.e., what they know, regardless of how implemented
- Or at the implementation level
  i.e., data structures in KB and algorithms that manipulate them

When Useful



Deep Space One

- Autonomous diagnosis & repair "Remote Agent"
- Compiled schematic to 7,000 var SAT problem



Muddy Children Problem

- Mom to N children "Don't get dirty"
- While playing, K≥1 get mud on forehead
- Father: "Some of you are dirty!"
- Father: "Raise your hand if you are dirty"
  - Noone raises hand
- Father: "Raise your hand if you are dirty"
  - Noone raises hand

…

- Father: "Raise your hand if you are dirty"
  - All dirty children raise hand

} K-1 times

Components of KR

- Syntax: defines the sentences in the language
- Semantics: defines the "meaning" of sentences
- Inference Procedure
  - Algorithm
  - Sound?
  - Complete?
  - Complexity
- Knowledge Base

Propositional Logic

- Syntax
  - Atomic sentences: P, Q, …
  - Connectives: $\wedge$, $\vee$, $\neg$, $\Rightarrow$
- Semantics
  - Truth Tables
- Inference
  - Modus Ponens
  - Resolution
  - DPLL
  - GSAT
- Complexity

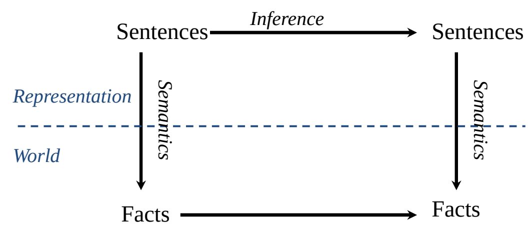
Propositional Logic: Syntax

- Atoms
  - P, Q, R, …
- Literals
  - P, ¬P
- Sentences
  - Any literal is a sentence
  - If S is a sentence
    - Then (S ∧ S) is a sentence
    - Then (S ∨ S) is a sentence
- Conveniences

  P → Q   same as ¬P ∨ Q

Semantics

- ***Syntax***: which arrangements of symbols are *legal*
  - (Def "sentences")
- ***Semantics***: what the symbols *mean* in the world
  - (Mapping between symbols and worlds)



Propositional Logic: **SEMANTICS**

- "Interpretation" (or "possible world")
  - Assignment to each variable either T or F
  - Assignment of T or F to each connective via defns

P ∧ Q

P \ Q	T	F
T	T	F
F	F	F

P ∨ Q

P \ Q	T	F
T	T	T
F	T	F

Satisfiability, Validity, & Entailment

- S is **satisfiable** if it is true in ***some*** world
- S is **unsatisfiable** if it is false ***all*** worlds
- S is **valid** if it is true in ***all*** worlds
- S1 **entails** S2 if ***wherever*** S1 is true S2 is also true

Examples

P → Q

X → X

S ∧ (W ∧ ¬S)

T ∨ ¬T

Notation

⇒
⊃ } **Implication** (syntactic symbol)
→

|— **Proves:** S1 |-ᵢ S2 if inference algo, i, says `S2' from S1

|= **Entails:** S1 |= S2 if wherever S1 is true S2 is also true

Resolution

If the unicorn is mythical, then it is immortal, but if it is not mythical, it is a reptile. If the unicorn is either immortal or a reptile, then it is horned.

$(\neg R \vee H)$ $(\neg I \vee H)$

$(M \vee R)$ $(\neg M \vee I)$

M = mythical
I = immortal
R = reptile
H = horned

Prop. Logic: **Knowledge Engr**

1) One of the women is a biology major
2) Lisa is not next to Dave in the ranking
3) Dave is immediately ahead of Jim
4) Jim is immediately ahead of a bio major
5) Mary or Lisa is ranked first

1. Choose Vocabulary

Universe: Lisa, Dave, Jim, Mary
LD = "Lisa is immediately ahead of Dave"
D = "Dave is a Bio Major"

2. Choose initial sentences (wffs)

Reasoning Tasks

- Model finding
  - KB = background knowledge
  - S = description of problem
  - Show (KB $\wedge$ S) is satisfiable
  - A kind of constraint satisfaction
- Deduction
  - S = question
  - Prove that KB $\models$ S
  - Two approaches:
    - Rules to derive new formulas from old (inference)
    - Show (KB $\wedge \neg$ S) is unsatisfiable

Special Syntactic Forms

- General Form:
  - $((q \wedge \neg r) \rightarrow s)) \wedge \neg (s \wedge t)$
- Conjunction Normal Form (CNF)
  - $(\neg q \vee r \vee s) \wedge (\neg s \vee \neg t)$
  - Set notation: $\{ (\neg q, r, s), (\neg s, \neg t) \}$
  - empty clause () = *false*
- Binary clauses: 1 or 2 literals per clause
  - $(\neg q \vee r)$ $(\neg s \vee \neg t)$
- Horn clauses: 0 or 1 positive literal per clause
  - $(\neg q \vee \neg r \vee s)$ $(\neg s \vee \neg t)$
  - $(q \wedge r) \rightarrow s$ $(s \wedge t) \rightarrow$ *false*

Propositional Logic: **Inference**

A ***mechanical*** process for computing new sentences

1. Backward & Forward Chaining
2. Resolution (Proof by Contradiction)
3. GSAT
4. Davis Putnam

Inference 1: Forward Chaining

Forward Chaining

Based on rule of ***modus ponens***

If know $P_1, \ldots, P_n$ & know $(P_1 \wedge \ldots \wedge P_n) \rightarrow Q$

Then can conclude Q

Backward Chaining: search

start from the query and go backwards

Analysis

- Sound?
- Complete?

  Can you prove
  { } |= Q ∨ ¬Q

- If KB has only Horn clauses & query is a single literal
  - Forward Chaining is complete
  - Runs linear in the size of the KB

Propositional Logic: **Inference**

A *mechanical* process for computing new sentences

1. Backward & Forward Chaining
2. Resolution (Proof by Contradiction)
3. GSAT
4. Davis Putnam

Conversion to CNF

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

1. Eliminate $\Leftrightarrow$, replacing $\alpha \Leftrightarrow \beta$ with $(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$.

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

2. Eliminate $\Rightarrow$, replacing $\alpha \Rightarrow \beta$ with $\neg\alpha \vee \beta$.

   $(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}) \wedge (\neg(P_{1,2} \vee P_{2,1}) \vee B_{1,1})$

3. Move $\neg$ inwards using de Morgan's rules and double-negation:

   $(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}) \wedge ((\neg P_{1,2} \wedge \neg P_{2,1}) \vee B_{1,1})$

4. Apply distributivity law ($\vee$ over $\wedge$) and flatten:

   $(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}) \wedge (\neg P_{1,2} \vee B_{1,1}) \wedge (\neg P_{2,1} \vee B_{1,1})$

Inference 2: Resolution

[Robinson 1965]

$\{ (p \vee \alpha), (\neg p \vee \beta \vee \gamma) \} \mid\!-_R (\alpha \vee \beta \vee \gamma)$

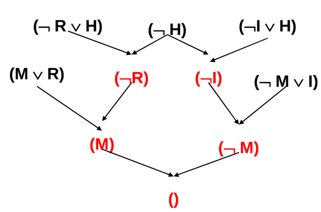
Correctness
  If S1 $\mid\!-_R$ S2 then S1 |= S2
Refutation Completeness:
  If S is unsatisfiable then S $\mid\!-_R$ ()

Resolution

If the unicorn is mythical, then it is immortal, but if it is not mythical, it is a reptile. If the unicorn is either immortal or a reptile, then it is horned.

Prove: the unicorn is horned.

M = mythical
I = immortal
R = reptile
H = horned



Resolution as Search

- States?
- Operators