Knowledge Representation I (Propositional Logic)

CSE 473

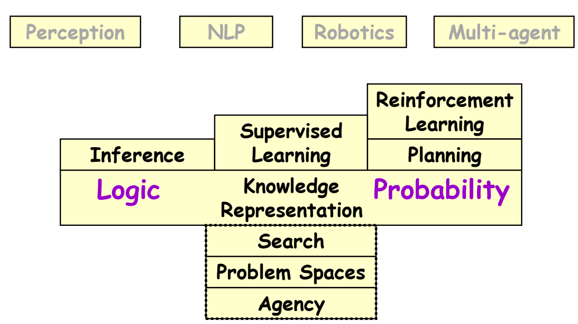
Some KR Languages

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

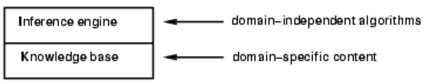
In Fact…

- All popular knowledge representation systems are equivalent to (or a subset of)
  - Logic
    - Either Propositional Logic
    - Or Predicate Calculus
  - Probability Theory

473 Topics



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

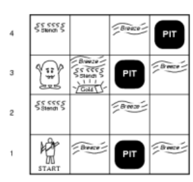
A simple knowledge-based agent

```
function KB-AGENT( percept) returns an action
  static: KB, a knowledge base
          t, a counter, initially 0, indicating time
  TELL(KB, MAKE-PERCEPT-SENTENCE( percept, t))
  action ← ASK(KB, MAKE-ACTION-QUERY(t))
  TELL(KB, MAKE-ACTION-SENTENCE(action, t))
  t ← t + 1
  return action
```

- The agent must be able to:
  Represent states, actions, etc.
  Incorporate new percepts
  Update internal representations of the world
  Deduce hidden properties of the world
  Deduce appropriate actions

Wumpus World PEAS description

- Performance measure
  gold +1000, death -1000
  -1 per step, -10 for using the arrow
- Environment
  Squares adjacent to wumpus are smelly
  Squares adjacent to pit are breezy
  Glitter iff gold is in the same square
  Shooting kills wumpus if you are facing it
  Shooting uses up the only arrow
  Grabbing picks up gold if in same square
  Releasing drops the gold in same square



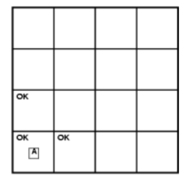
Wumpus world characterization

- Fully Observable?
- Deterministic?
- Episodic?
- Static?
- Discrete?
- Single-agent?

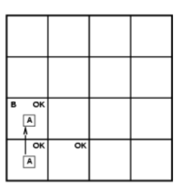
Wumpus world characterization

- Fully Observable No – only local perception
- Deterministic Yes – outcomes exactly specified
- Episodic No – sequential at the level of actions
- Static Yes – Wumpus and Pits do not move
- Discrete Yes
- Single-agent Yes – Wumpus is essentially a natural feature

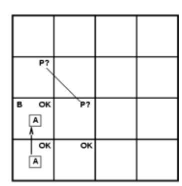
Exploring a wumpus world



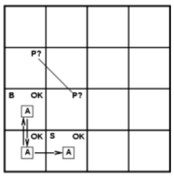
Exploring a wumpus world



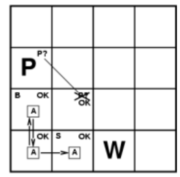
Exploring a wumpus world



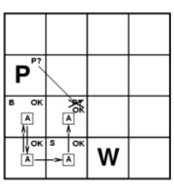
Exploring a wumpus world



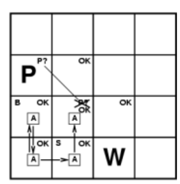
Exploring a wumpus world



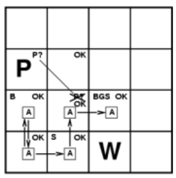
Exploring a wumpus world



Exploring a wumpus world



Exploring a wumpus world



Logic in general

- Logics are formal languages for representing information such that conclusions can be drawn
-
- Syntax defines the sentences in the language
-
- Semantics define the "meaning" of sentences;
-

  i.e., define truth of a sentence in a world

Entailment

- Entailment means that one thing follows from another:

  $KB \models \alpha$

- Knowledge base *KB* entails sentence α if and only if α is true in all worlds where *KB* is true

  E.g., the KB containing "the Giants won" and "the Reds won" entails "Either the Giants won or the Reds won"

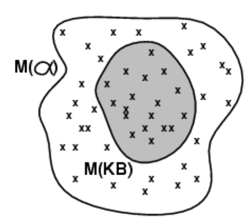
  E.g., x+y = 4 entails 4 = x+y

  Entailment is a relationship between sentences (i.e., syntax) that is based on semantics

Models

- Logicians typically think in terms of models, which are formally structured worlds with respect to which truth can be evaluated
- We say *m* is a model of a sentence α if α is true in *m*
- *M(α)* is the set of all models of α
-
- Then $KB \models \alpha$ iff $M(KB) \subseteq M(\alpha)$
-

  E.g. *KB* = Giants won and Reds won
  α = Giants won

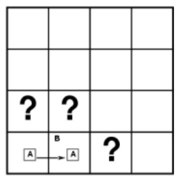


Entailment in the wumpus world

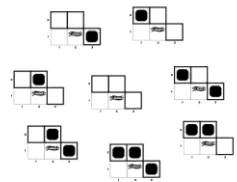
Situation after detecting nothing in [1,1], moving right, breeze in [2,1]

Consider possible models for *KB* (only pits)

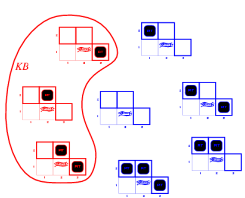
3 Boolean choices $\Rightarrow$ 8 possible models



Wumpus models

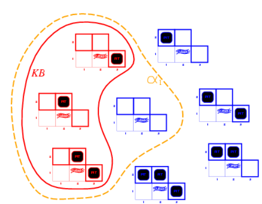


Wumpus models



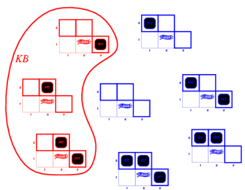
- *KB* = wumpus-world rules + observations

Wumpus models



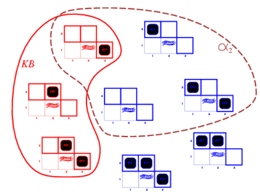
- *KB* = wumpus-world rules + observations
- α_1 = "[1,2] is safe", *KB* $\models \alpha_1$, proved by model checking
-
-

Wumpus models



- *KB* = wumpus-world rules + observations

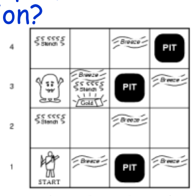
Wumpus models



- *KB* = wumpus-world rules + observations
- α_2 = "[2,2] is safe", $KB \not\models \alpha_2$

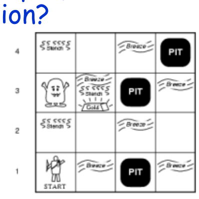
Missing Elements

- How does an agent reason about the wumpus world?
- How do we map truth/information between the real (wumpus) world and our representation?



Missing Elements

- How does an agent reason about the wumpus world?
  Inference
- How do we map truth/information between the real (wumpus) world and our representation?
  Semantics



Inference

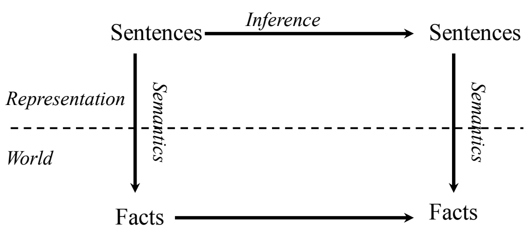
- *KB* $\vdash_i \alpha$ = sentence α can be derived from *KB* by procedure *i*
- Soundness: *i* is sound if whenever *KB* $\vdash_i \alpha$, it is also true that *KB* $\models \alpha$
- Completeness: *i* is complete if whenever *KB* $\models \alpha$, it is also true that *KB* $\vdash_i \alpha$
- Preview: we will define a logic (first-order logic) which is expressive enough to say almost anything of interest, and for which there exists a sound and complete inference procedure.
- That is, the procedure will answer any question whose answer follows from what is known by the *KB*.

Inference

- *KB* $\vdash_i \alpha$ = sentence α can be derived from *KB* by procedure *i*
- Soundness: *i* is sound if whenever *KB* $\vdash_i \alpha$, it is also true that *KB* $\models \alpha$
  "Prodecure *i* only infers things that are true."
- Completeness: *i* is complete if whenever *KB* $\models \alpha$, it is also true that *KB* $\vdash_i \alpha$
  "If something is true, procedure *i* will infer it."
- Preview: we will define a logic (first-order logic) which is expressive enough to say almost anything of interest, and for which there exists a sound and complete inference procedure.
- That is, the procedure will answer any question whose answer follows from what is known by the *KB*.

Semantics

- ***Syntax***: a description of the *legal* arrangements of symbols (Def "sentences")
- ***Semantics***: what the arrangement of symbols *means* in the world



Propositional Logic

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

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

		Q	
		T	F
P	T	T	T
	F	T	F

$P \vee Q$

		Q	
		T	F
P	T		
	F		

$P \wedge Q$

		Q	
		T	F
P	T		
	F		

$\neg P$

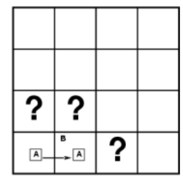
Wumpus world sentences

- Let $P_{i,j}$ be true if there is a pit in [i, j].
- Let $B_{i,j}$ be true if there is a breeze in [i, j].

$\neg P_{1,1}$
$\neg B_{1,1}$
$B_{2,1}$

- "Pits cause breezes in adjacent squares"

$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$
$B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$



Truth tables for inference

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	KB	α_1
false	*false*	*false*	*false*	*false*	*false*	*false*	*false*	*true*
false	*false*	*false*	*false*	*false*	*false*	*true*	*false*	*true*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
false	*true*	*false*	*false*	*false*	*false*	*false*	*false*	*true*
false	*true*	*false*	*false*	*false*	*false*	*true*	<u>*true*</u>	<u>*true*</u>
false	*true*	*false*	*false*	*false*	*true*	*false*	<u>*true*</u>	<u>*true*</u>
false	*true*	*false*	*false*	*false*	*true*	*true*	<u>*true*</u>	<u>*true*</u>
false	*true*	*false*	*false*	*true*	*false*	*false*	*false*	*true*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
true	*true*	*true*	*true*	*true*	*true*	*true*	*false*	*false*

α1 = "[1,2] is safe"

Validity and satisfiability

A sentence is valid if it is true in all models,
e.g., *True*, $A \vee \neg A$, $A \Rightarrow A$, $(A \wedge (A \Rightarrow B)) \Rightarrow B$

Validity is connected to inference via the Deduction Theorem:
$KB \models \alpha$ if and only if $(KB \Rightarrow \alpha)$ is valid

A sentence is satisfiable if it is true in some model
e.g., $A \vee B$, C

A sentence is unsatisfiable if it is true in no models
e.g., $A \wedge \neg A$

Satisfiability is connected to inference via the following:
$KB \models \alpha$ if and only if $(KB \wedge \neg \alpha)$ is unsatisfiable