KBS : Logic and Reasoning

- Motivation
- Objectives
- Knowledge and Reasoning
  - logic as prototypical reasoning system
  - syntax and semantics
  - validity and satisfiability
  - logic languages
- Reasoning Methods
  - propositional and predicate calculus
  - inference methods
- Reasoning in Knowledge-Based Systems
  - shallow and deep reasoning
  - forward and backward chaining
  - alternative inference methods
  - meta-knowledge
- Important Concepts and Terms
- Chapter Summary

Dilbert on Reasoning

Motivation

- without reasoning, knowledge-based systems would be practically worthless
  - derivation of new knowledge
  - examination of the consistency or validity of existing knowledge
- reasoning in KBS can perform certain tasks better than humans
  - reliability, availability, speed
  - also some limitations
    - common-sense reasoning
    - complex inferences

Objectives

- be familiar with the essential concepts of logic and reasoning
  - sentence, operators, syntax, semantics, inference methods
- appreciate the importance of reasoning for knowledge-based systems
  - generating new knowledge
  - explanations
- understand the main methods of reasoning used in KBS
  - shallow and deep reasoning
  - forward and backward chaining
- evaluate reasoning methods for specific tasks and scenarios
- apply reasoning methods to simple problems

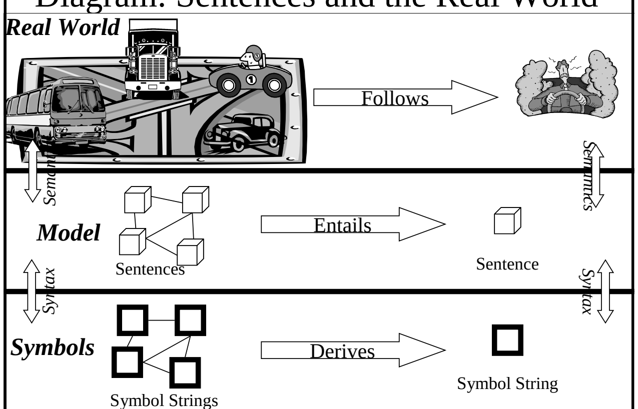
Knowledge Representation Languages

- syntax
  - sentences of the language that are built according to the syntactic rules
  - some sentences may be nonsensical, but syntactically correct
- semantics
  - refers to the facts about the world for a specific sentence
  - interprets the sentence in the context of the world
  - provides **meaning** for sentences
- languages with precisely defined syntax and semantics can be called **logics**

Sentences and the Real World

- *syntax*
  - describes the principles for constructing and combining sentences
    - e.g. BNF grammar for admissible sentences
    - inference rules to derive new sentences from existing ones
- *semantics* Sentences Sentence
  - establishes the relationship between a sentence and the aspects of the real world it describes
  - can be checked directly by comparing sentences with the corresponding objects in the real world
    - not always feasible or practical
  - complex sentences can be checked by examining their individual parts

Diagram: Sentences and the Real World



Introduction to Logic

- expresses knowledge in a particular mathematical notation

  `All birds have wings --> ¥x. Bird(x) -> HasWings(x)`
- rules of inference
  - guarantee that, given true facts or premises, the new facts or premises derived by applying the rules are also true

  `All robins are birds --> ¥x Robin(x) -> Bird(x)`
- given these two facts, application of an inference rule gives:

  `¥x Robin(x) -> HasWings(x)`

Logic and Knowledge

- rules of inference act on the superficial structure or syntax of the first 2 formulas
  - doesn't say anything about the meaning of birds and robins
  - could have substituted mammals and elephants etc.
- major advantages of this approach
  - deductions are guaranteed to be correct to an extent that other representation schemes have not yet reached
  - easy to automate derivation of new facts
- problems
  - computational efficiency
  - uncertain, incomplete, imprecise knowledge

Summary of Logic Languages

- propositional logic
  - facts
  - true/false/unknown
- first-order logic
  - facts, objects, relations
  - true/false/unknown
- temporal logic
  - facts, objects, relations, times
  - true/false/unknown
- probability theory
  - facts
  - degree of belief [0..1]
- fuzzy logic
  - degree of truth
  - degree of belief [0..1]

Propositional Logic

- Syntax
- Semantics
- Validity and Inference
- Models
- Inference Rules
- Complexity

Syntax

- symbols
  - logical constants True, False
  - propositional symbols P, Q, …
  - logical connectives
    - conjunction $\wedge$, disjunction $\vee$,
    - negation $\neg$,
    - implication $\Rightarrow$, equivalence $\Leftrightarrow$
  - parentheses (,)
- sentences
  - constructed from simple sentences
  - conjunction, disjunction, implication, equivalence, negation

BNF Grammar Propositional Logic

Sentence → *AtomicSentence* | *ComplexSentence*
AtomicSentence → `True` | `False` | `P` | `Q` | `R` | ...
ComplexSentence → (*Sentence*)
| *Sentence Connective Sentence*
| ¬ *Sentence*

Connective → ∧ | ∨ | ⇒ | ⇔

ambiguities are resolved through precedence ¬ ∧ ∨ ⇒ ⇔ or parentheses

e.g. ¬ *P* ∨ *Q* ∧ *R* ⇒ *S* is equivalent to (¬ *P*) ∨ (*Q* ∧ *R*)) ⇒ *S*

Semantics

- interpretation of the propositional symbols and constants
  - symbols can be any arbitrary fact
    - sentences consisting of only a propositional symbols are satisfiable, but not valid
  - the constants True and False have a fixed interpretation
    - True indicates that the world is as stated
    - False indicates that the world is not as stated
- specification of the logical connectives
  - frequently explicitly via truth tables

Validity and Satisfiability

- a sentence is *valid* or necessarily true if and only if it is true under all possible interpretations in all possible worlds
  - also called a *tautology*
  - since computers reason mostly at the syntactic level, valid sentences are very important
    - interpretations can be neglected
- a sentence is *satisfiable* iff there is some interpretation in some world for which it is true
- a sentence that is not satisfiable is *unsatisfiable*
  - also known as a *contradiction*

Truth Tables for Connectives

P	***Q***	**¬ *P***	***P* ∧ *Q***	***P* ∨ *Q***	***P* ⇒ *Q***	***P* ⇔ *Q***
False	*False*	*True*	*False*	*False*	*True*	*True*
False	*True*	*True*	*False*	*True*	*True*	*False*
True	*False*	*False*	*False*	*True*	*False*	*False*
True	*True*	*False*	*True*	*True*	*True*	*True*

Validity and Inference

- truth tables can be used to test sentences for validity
  - one row for each possible combination of truth values for the symbols in the sentence
  - the final value must be True for every sentence

Propositional Calculus

- properly formed statements that are either True or False
- syntax
  - logical constants, True and False
  - proposition symbols such as P and Q
  - logical connectives: and ^, or V, equivalence <=>, implies => and not ~
  - parentheses to indicate complex sentences
- sentences in this language are created through application of the following rules
  - True and False are each (atomic) sentences
  - Propositional symbols such as P or Q are each (atomic) sentences
  - Enclosing symbols and connective in parentheses yields (complex) sentences, e.g., (P ^ Q)

Complex Sentences

- Combining simpler sentences with logical connectives yields complex sentences
  - conjunction
    - sentence whose main connective is and: P ^ (Q V R)
  - disjunction
  - sentence whose main connective is or: A V (P ^ Q)
  - implication (conditional)
    - sentence such as (P ^ Q) => R
    - the left hand side is called the premise or antecedent
    - the right hand side is called the conclusion or consequent
    - implications are also known as rules or if-then statements
  - equivalence (biconditional)
    - (P ^ Q) <=> (Q ^ P)
  - negation
    - the only unary connective (operates only on one sentence)
    - e.g., ~P

Syntax of Propositional Logic

- A BNF (Backus-Naur Form) grammar of sentences in propositional logic

```
Sentence -> AtomicSentence | ComplexSentence

  AtomicSentence -> True | False | P | Q | R | ...

  ComplexSentence -> (Sentence)
                   | Sentence Connective Sentence
                   | ~Sentence

  Connective -> ^ | V | <=> | =>
```


Semantics

- propositions can be interpreted as any facts you want
  - e.g., P means "robins are birds", Q means "the wumpus is dead", etc.
- meaning of complex sentences is derived from the meaning of its parts
  - one method is to use a truth table
  - all are easy except P => Q
    - this says that if P is true, then I claim that Q is true; otherwise I make no claim;
    - P is true and Q is true, then P => Q is true
    - P is true and Q is false, then P => Q is false
    - P is false and Q is true, then P => Q is true
    - P is false and Q is false, then P => Q is true

Exercise Semantics and Truth Tables

- Use a truth table to prove the following:
  - P represents the fact "Wally is in location [1, 3]": W[1,3]
  - H represents the fact "Wally is in location [2, 2]": W[2,2]
  - We know that Wally is either in [1,3] or [2,2]: (P V H)
  - We learn that Wally is not in [2,2]: ~H
  - Can we prove that Wally is in [1,3]: ((P V H) ^ ~H) => P
    - This says that if the agent has some premises, and a possible conclusion, it can determine if the conclusion is true (i.e., all the rows of the truth table are true)
- Inference Rules : more efficient than truth tables

Modus Ponens

- eliminates =>

```
(X => Y),   X
_______________
       Y
```

  - If it rains, then the streets will be wet.
  - It is raining.
  - Infer the conclusion: The streets will be wet. (affirms the antecedent)

Modus tollens

(X => Y), ~Y

____

¬ X

- If it rains, then the streets will be wet.
- The streets are not wet.
- Infer the conclusion: It is not raining.

- NOTE: Avoid the fallacy of affirming the consequent:
  - If it rains, then the streets will be wet.
  - The streets are wet.
  - *cannot* conclude that it is raining.

  - If Bacon wrote Hamlet, then Bacon was a great writer.
  - Bacon was a great writer.
  - cannot conclude that Bacon wrote Hamlet.

Syllogism

- chain implications to deduce a conclusion

(X => Y), (Y => Z)

____

(X => Z)

More Inference Rules

- and-elimination
- and-introduction
- or-introduction
- double-negation elimination
- unit resolution

Resolution

(X v Y), (~Y v Z)

(X v Z)

- basis for the inference mechanism in the Prolog language and some theorem provers

Complexity issues

- truth table enumerates 2^n rows of the table for any proof involving n symbol
  - it is complete
  - computation time is exponential in n
- checking a set of sentences for satisfiability is NP-complete
  - but there are some circumstances where the proof only involves a small subset of the KB, so can do some of the work in polynomial time
  - if a KB is monotonic (i.e., even if we add new sentences to a KB, all the sentences entailed by the original KB are still entailed by the new larger KB), then you can apply an inference rule locally (i.e., don't have to go checking the entire KB)

Inference Methods 1

- deduction — sound
  - conclusions must follow from their premises; prototype of logical reasoning
- induction — unsound
  - inference from specific cases (examples) to the general
- abduction — unsound
  - reasoning from a true conclusion to premises that may have caused the conclusion
- resolution — sound
  - find two clauses with complementary literals, and combine them
- generate and test — unsound
  - a tentative solution is generated and tested for validity
  - often used for efficiency (trial and error)

Inference Methods 2

- default reasoning — unsound
  - general or common knowledge is assumed in the absence of specific knowledge
- analogy — unsound
  - a conclusion is drawn based on similarities to another situation
- heuristics — unsound
  - rules of thumb based on experience
- intuition — unsound
  - typically human reasoning method
- nonmonotonic reasoning — unsound
  - new evidence may invalidate previous knowledge
- autoepistemic — unsound
  - reasoning about your own knowledge

Predicate Logic

- new concepts (in addition to propositional logic)
  - complex objects
    - terms
  - relations
    - predicates
    - quantifiers
  - syntax
  - semantics
  - inference rules
  - usage

Objects

- distinguishable things in the real world
  - people, cars, computers, programs, ...
- frequently includes concepts
  - colors, stories, light, money, love, ...
- properties
  - describe specific aspects of objects
    - green, round, heavy, visible,
  - can be used to distinguish between objects

Relations

- establish connections between objects
- relations can be defined by the designer or user
  - neighbor, successor, next to, taller than, younger than, …
- functions are a special type of relation
  - non-ambiguous: only one output for a given input

Syntax

- also based on sentences, but more complex
  - sentences can contain terms, which represent objects
- constant symbols: `A`, `B`, `C`, `Franz`, $\texttt{Square}_{1,3}$, …
  - stand for unique objects (in a specific context)
- predicate symbols: Adjacent-To, Younger-Than, ...
  - describes relations between objects
- function symbols: Father-Of, Square-Position, …
  - the given object is related to exactly one other object

Semantics

- provided by interpretations for the basic constructs
  - usually suggested by meaningful names
- constants
  - the interpretation identifies the object in the real world
- predicate symbols
  - the interpretation specifies the particular relation in a model
  - may be explicitly defined through the set of tuples of objects that satisfy the relation
- function symbols
  - identifies the object referred to by a tuple of objects
  - may be defined implicitly through other functions, or explicitly through tables

BNF Grammar Predicate Logic

Sentence	→ *AtomicSentence*			
	| *Sentence Connective Sentence*			
	| *Quantifier Variable, ... Sentence*			
	| ¬ *Sentence* | (*Sentence*)			
AtomicSentence	→ *Predicate(Term, ...)*	| *Term = Term*		
Term	→ *Function(Term, ...)*	| *Constant*	| *Variable*	
Connective	→ ∧ | ∨ | ⇒ | ⇔			
Quantifier	→ ∀ | ∃			
Constant	→ *A, B, C,* X_1, X_2, *Jim, Jack*			
Variable	→ *a, b, c,* x_1, x_2, *counter, position*			
Predicate	→ *Adjacent-To, Younger-Than,*			
Function	→ *Father-Of, Square-Position, Sqrt, Cosine*			

ambiguities are resolved through precedence or parentheses

Terms

- logical expressions that specify objects
- constants and variables are terms
- more complex terms are constructed from function symbols and simpler terms, enclosed in parentheses
  - basically a complicated name of an object
- semantics is constructed from the basic components, and the definition of the functions involved
  - either through explicit descriptions (e.g. table), or via other functions

Unification

- an operation that tries to find consistent variable bindings (substitutions) for two terms
  - a substitution is the simultaneous replacement of variable instances by terms, providing a “binding” for the variable
  - without unification, the matching between rules would be restricted to constants
  - often used together with the resolution inference rule
  - unification itself is a very powerful and possibly complex operation
    - in many practical implementations, restrictions are imposed
      - e.g. substitutions may occur only in one direction (“matching”)

Atomic Sentences

- state facts about objects and their relations
- specified through predicates and terms
  - the predicate identifies the relation, the terms identify the objects that have the relation
- an atomic sentence is `true` if the relation between the objects holds
  - this can be verified by looking it up in the set of tuples that define the relation

Complex Sentences

- logical connectives can be used to build more complex sentences
- semantics is specified as in propositional logic

Quantifiers

- can be used to express properties of collections of objects
  - eliminates the need to explicitly enumerate all objects
- predicate logic uses two quantifiers
  - universal quantifier $\forall$
  - existential quantifier $\exists$

Universal Quantification

- states that a predicate P is holds for all objects x in the universe under discourse
  $\forall x\ P(x)$
- the sentence is true if and only if all the individual sentences where the variable x is replaced by the individual objects it can stand for are true

Existential Quantification

- states that a predicate P holds for some objects in the universe
  $\exists x\ P(x)$
- the sentence is true if and only if there is at least one true individual sentence where the variable *x* is replaced by the individual objects it can stand for

Horn clauses or sentences

- class of sentences for which a polynomial-time inference procedure exists
  - $P_1 \wedge P_2 \wedge ... \wedge P_n \Rightarrow Q$
    where P_i and Q are non-negated atomic sentences
- not every knowledge base can be written as a collection of Horn sentences
- Horn clauses are essentially rules of the form
  - If $P_1 \wedge P_2 \wedge ... \wedge P_n$ then Q

Shallow and Deep Reasoning

- shallow reasoning
  - also called experiential reasoning
  - aims at describing aspects of the world heuristically
  - short inference chains
  - possibly complex rules
- deep reasoning
  - also called causal reasoning
  - aims at building a model of the world that behaves like the “real thing”
  - long inference chains
  - often simple rules that describe cause and effect relationships

Examples Shallow and Deep Reasoning

- shallow reasoning

```
IF a car has
        a good battery
        good spark plugs
        gas
        good tires
  THEN the car can move
```

- deep reasoning

```
IF the battery is good
THEN there is electricity

IF there is electricity AND good spark plugs
THEN the spark plugs will fire

IF the spark plugs fire AND
        there is gas
THEN the engine will run

IF the engine runs AND
        there are good tires
THEN the car can move
```


Forward Chaining

- given a set of basic facts, we try to derive a conclusion from these facts
- example: What can we conjecture about Clyde?

```
IF elephant(x) THEN mammal(x)
IF mammal(x) THEN animal(x)
elephant (Clyde)
```

modus ponens:

IF p THEN q
p
―――
q

unification:

find compatible values for variables

Forward Chaining Example

```
IF elephant(x) THEN mammal(x)
IF mammal(x) THEN animal(x)
elephant(Clyde)
```

unification: find compatible values for variables

modus ponens:
IF p THEN q
p
―――
q

IF elephant(x) THEN mammal(x)

elephant (Clyde)

Forward Chaining Example

```
IF elephant(x) THEN mammal(x)
IF mammal(x) THEN animal(x)
elephant(Clyde)
```

unification: find compatible values for variables

modus ponens:
IF p THEN q
p
―――
q

animal(Clyde)

IF mammal(Clyde) THEN animal(Clyde)

IF elephant(Clyde) THEN mammal(Clyde)

elephant (Clyde)

Backward Chaining

- try to find supportive evidence (i.e. facts) for a hypothesis
- example: Is there evidence that Clyde is an animal?

```
IF elephant(x) THEN mammal(x)
IF mammal(x) THEN animal(x)
elephant (Clyde)
```

modus ponens:

```
IF p THEN q
p
---
q
```

unification:

find compatible values for variables

Backward Chaining Example

```
IF elephant(x) THEN mammal(x)
IF mammal(x) THEN animal(x)
elephant(Clyde)
```

unification:
find compatible values for variables

modus ponens:

```
IF p THEN q
p
---
q
```

animal(Clyde)

?

IF mammal(x) THEN animal(x)

Backward Chaining Example

```
IF elephant(x) THEN mammal(x)
IF mammal(x) THEN animal(x)
elephant(Clyde)
```

unification:
find compatible values for variables

modus ponens:

```
IF p THEN q
p
---
q
```

animal(Clyde)

IF mammal(Clyde) THEN animal(Clyde)

IF elephant(Clyde) THEN mammal(Clyde)

elephant (Clyde)

Forward vs. Backward Chaining

Forward Chaining	*Backward Chaining*
planning, control	diagnosis
data-driven	goal-driven (hypothesis)
bottom-up reasoning	top-down reasoning
find possible conclusions supported by given facts	find facts that support a given hypothesis
similar to breadth-first search	similar to depth-first search
antecedents (LHS) control evaluation	consequents (RHS) control evaluation

Alternative Inference Methods

- theorem proving
  - emphasis on mathematical proofs, not so much on performance and ease of use
- probabilistic reasoning
  - integrates probabilities into the reasoning process
- fuzzy reasoning
  - enables the use of ill-defined predicates

Metaknowledge

- deals with "knowledge about knowledge"
  - e.g. reasoning about properties of knowledge representation schemes, or inference mechanisms
  - usually relies on higher order logic
    - in (first order) predicate logic, quantifiers are applied to variables
    - second-order predicate logic allows the use of quantifiers for function and predicate symbols
      - equality is an important second order axiom
        - » two objects are equal if all their properties (predicates) are equal
    - may result in substantial performance problems

Summary Reasoning

- reasoning relies on the ability to generate new knowledge from existing knowledge
  - implemented through inference rules
    - related terms: inference procedure, inference mechanism, inference engine
- computer-based reasoning relies on syntactic symbol manipulation (derivation)
  - inference rules prescribe which combination of sentences can be used to generate new sentences
  - ideally, the outcome should be consistent with the meaning of the respective sentences ("sound" inference rules)
- logic provides the formal foundations for many knowledge representation schemes
  - rules are frequently used in expert systems

Important Concepts and Terms

- and operator
- atomic sentence
- backward chaining
- existential quantifier
- expert system shell
- forward chaining
- higher order logic
- Horn clause
- inference
- inference mechanism
- If-Then rules
- implication
- knowledge
- knowledge base
- knowledge-based system
- knowledge representation
- matching
- meta-knowledge
- not operator
- or operator
- predicate logic
- propositional logic
- production rules
- quantifier
- reasoning
- rule
- satisfiability
- semantics
- sentence
- symbol
- syntax
- term
- validity
- unification
- universal quantifier