knowledge
  knowledge representation, knowledge base, types of knowledge

wumpus world
  example of knowledge-based agents

knowledge representation language
  syntax, semantics, interpretation

inference
  sound, complete

logic
  syntax, semantics, limitations
Knowledge

and agents

world model
contains knowledge the agent has about
the world

inference mechanism
draws conclusions from current
knowledge

actions
are taken based on conclusions

learning
allows adaptations of the world model
Knowledge

and its meaning

ontology
study of the nature of being or existence:
 vocabulary of the domain

epistemology
study of knowledge:
nature, structure, origins

a priori knowledge
known to be true in advance of
experience
does not require evidence for its
validation

a posteriori knowledge
empirical, open to revision
requires evidence for its validation
**Types of Knowledge**

**procedural**

knowing how to do something
algorithm

**declarative**

statements that can be true or false
specification

**tacit** also: unconscious

can’t be expressed in language
skills

also other classifications of knowledge
Knowledge Hierarchy

meta-knowledge
knowledge about knowledge
selects applicable knowledge

knowledge
information items and their relationships
usually loosely structured

information
processed data

data
items of potential interest
usually rigidly structured

noise
irrelevant items, of no interest
often obscure data
knowledge about knowledge
information items and their relationships
items with pre-defined relationships
unprocessed input
Knowledge-Based Agent

reason about representations of the world

tasks
accept new tasks through explicit goals

competence
acquire knowledge by being told or learning

flexibility
adapt to changes by updating relevant knowledge

knowledge required about
current state of the world
infer inaccessible properties of the world
keep track of changes in the world
consequences of actions
Description Levels

for knowledge-based agents

**knowledge level** or epistemological level
most abstract level
used for exchanging knowledge via Tell, Ask

**logical level**
encoding of knowledge into logical sentences

**implementation level**
runs on the agent architecture
physical representations of the sentences in a computer
important for efficient performance
Wumpus World

endangered cave-dwelling agents

world
cave consisting of rooms connected by passageways

wumpus
beast that eats anyone entering its room disperses stench into adjacent rooms gives out a penetrating scream if killed

pits
bottomless traps generate a breeze in adjoining rooms

gold
reward for the agent perceived as a glitter

walls
surround the cave
result in a bump if the agent walks into it
Wumpus World Agent

formal representation

percepts
[Stench, Breeze, Glitter, Bump, Scream]

actions
[Forward, Right, Left, Grab, Shoot]

goal
find the goal and bring it back to the start as quickly as possible without getting killed

environment
grid of squares with agents and objects
Knowledge Representation Language

*express knowledge in computer-tractable form*

**syntax**

describes admissible sentences

**semantics**

relates sentences to the real world

**inference rules** logic, proof theory

describe the generation of new sentences from existing ones
Logic

and knowledge

knowledge representation
formal method to describe knowledge via logical sentences

inference mechanism
generally accepted rules of reasoning
often with strict formal properties,
e.g. correctness, completeness
Inference

in computers

interpretation
   is usually only known to the designer or user of a model

real world
   no real-world knowledge except for the knowledge base

valid sentences
   can be checked by a computer
   may be very complex
   are independent of their interpretation
Formal Logic

for knowledge representation and reasoning

syntax

defines the language for statements

a well-formed formula (wff) is a legitimate expression

semantics

establishes the connection between the language and the problem domain

provides an interpretation of a formula

axioms

represent the basic assumptions

inference rules

specify when a new formula can be derived from existing ones
calculus

set of rules for the derivation of new
formulae (*theorems*)

**proof** of a theorem

sequence of rule applications during the
derivation of a theorem
Logic Systems

and their properties

interpretation
{assignment of truth values to a wff}

model
{interpretation in which the wff is true}

satisfiability
{there is an interpretation which makes the wff true}

validity
{the wff is true in all interpretation}

correctness of a calculus
{only semantically valid formulae can be deduced syntactically}

completeness of a calculus
{each semantically valid formula can also...}
be deduced syntactically
Propositional Logic

manipulation of propositions

knowledge representation
logical variables represent propositions
propositions can be either true or false
logical connectives for constructing
compound sentences

inference
specified by a calculus
allows the evaluation of a sentence to true
or false

limited ability to express knowledge
not adequate for many statements about the world
Propositional Logic

*logical treatment of simple statements*

**syntax**
propositional symbols, logical connectives

**semantics**
a truth value is assigned to each symbol
(interpretation)

**evaluation**
truth tables, semantic trees, etc.
decidable: there are systematic procedures to check the validity of any propositional formula

**limitations**
expressiveness: no quantifiers, variables, terms, functions
Predicate Logic

manipulation of predicates and terms

predicates
express relationships between objects

terms
used for the specification of objects

- constants stand for one specific object
- variables represent currently unspecified objects
- functions map arguments (terms) from one domain to another
quantifiers
  restrict the scope of variables

unification
  computes proper substitutions for
  matching predicate logic expressions

much more powerful than propositional logic
still some restrictions in its basic form (first order predicate logic)
Predicate Logic

logical treatment of complex statements

syntax quantifiers, predicates, constants, variables, functions, terms
several notational variants (normal forms, clause form)

semantics a mapping is defined between objects in a domain and symbols (interpretation)
far more complex than for propositional logic

evaluation undecidable: there can be no systematic procedures to check the validity of an arbitrary predicate logic formula
various calculi and proof methods, especially for limited subsets (Horn
clause logic, first order predicate logic)

limitations efficiency, understandability
Inference Methods

ways to come to conclusions

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)

**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
Metaknowledge

knowledge about knowledge

abstraction
similarities or patterns in the knowledge itself are found

evaluation
the computation process is observed, and knowledge about it is gathered and applied

verification
new knowledge is in the correct form
“Am I doing things right?”

validation
a chain of correct inference steps leads to the correct answer
“Am I doing the right thing?”
Important Concepts

non-monotonicity
axioms can be retracted, and new ones introduced

truth maintenance systems
maintain the integrity of the knowledge base
intermediate conclusions based on retracted facts are withdrawn

closed world assumption
if something is not explicitly stated as an axiom, it is assumed to be false

refutation “reductio ad absurdum”
a statement is proven by assuming that it is false, and showing that this leads to a contradiction

frame problem
recognition of changes over time
inspired by movies as sequences of frames
Advantages of logic

correctness
  consistency can be checked automatically

completeness
  all possible solutions are guaranteed to be found

expressiveness
  in principle, all formalisms can be translated into logic
  higher order logic might be required

declarative style
  does not require implementation-dependent details
Limitations

of logic

efficiency
  evaluation time unknown, often no
  intermediate results

formalization
  can be tedious

uncertainty
  only true and false

control
  heuristics for evaluation either are
  extra-logical or meta-level concepts

nonmonotonicity
  not for deductive approaches
Summary

Knowledge-Based Agents

knowledge
  knowledge representation, knowledge base, types of knowledge

wumpus world
  example of knowledge-based agents

knowledge representation language
  syntax, semantics, interpretation

inference
  sound, complete

logic
  syntax, semantics, limitations