Lecture 5: Introduction to Knowledge Representation

Dr. Roman V Belavkin

BIS4410

Contents

1 Knowledge Engineering

Knowledge Engineering

Definition 1 (Knowledge Engineering)**.** The process of designing knowledge-based systems (e.g. expert systems, data-driven models).

It consist of three stages:

Knowledge acquisition : the process of obtaining the knowledge from experts.

Knowledge representation : selecting the most appropriate data types and structures to represent and encode the knowledge in explicit form (e.g. lists, sets, scripts, decision trees, object-attribute-value triplets, etc).

Knowledge validation : testing that the knowledge-based system for correctness and completeness.

2 Knowledge Acquisition

Knowledge Acquisition

Definition 2 (Knowledge Acquisition / Elicitation)**.** The process of gathering know-how from experts or other sources.

Definition 3 (Knowledge Acquisition / Elicitation)**.** A transfer from knowledge *source* (e.g. human experts, documents) to a knowledge *repository* (e.g. expert systems).

Typical techniques used:

- Structured interviews

- Protocol or talk aloud analysis
- Questionnaires
- Surveys
- Observation
- Simulation

Declarative and Procedural Knowledge
Knowledge can be classified into:

Declarative or descriptive knowledge — propositions or facts describing *what* (e.g. which facts are known to be true).

Procedural knowledge — logical rules describing *how*.

Remark 1. *Evidence suggests that these types of knowledge are encoded in different parts of the brain or even use different mechanisms. Procedural knowledge is usually difficult to describe, but also harder to forget.*

Symbolic and Sub-Symbolic Representations

Symbolic representations convey information by discrete units (explicit), usually expressed in formal language (i.e. words, sentences).

Sub-symbolic representations convey information by properties (implicit) of many objects (e.g. data, attribute values, probability distributions, weights in neural networks, etc).

Hybrid systems involve both types of knowledge representation.

Plane

Train Car

000001100000
000110011000
001100001100
011111111110
110000000011

Tacit Knowledge Capture
Main approaches:

- Interviewing experts (structured interviews and stories).
- Learning by being told.
- Learning by observation.

Remark 2. *Tacit knowledge often cannot be made explicit due to sub-symbolic information. For example, see Doyle (1988) for problems with capturing probabilities.*

3 Knowledge Representation

Knowledge Representation Paradigms
According to McCalla and Cercone (1983):

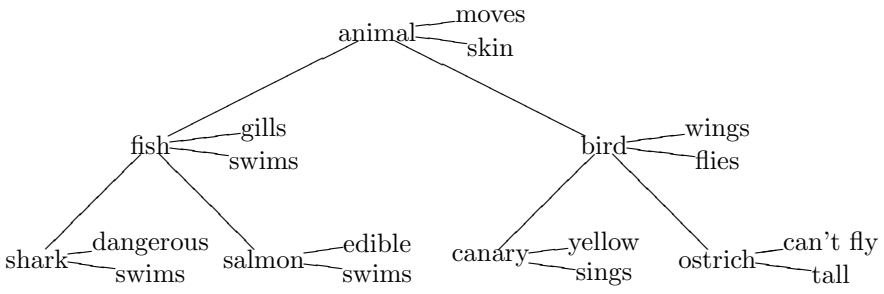
Semantic networks are concept graphs connected by relations, such as IS-A relation (*Dog IS-A mammal*).

First-order logic is useful for manipulation of facts (e.g. automatic theorem-proving).

Frames or *schemata* are decomposable structures or *chunks* used to represent concepts.

Production systems are algorithms represented by sets of rules (*condition-action* pairs), which were used to model human reasoning.

Semantic Network



Question 1. *There are three types of edges in this graph. One is* `IS-A` *relation (or attribute). Suggest what should be the other two relations?*

Object Attribute Value (OAV) Triplets

Definition 4 (OAV-triplet)**.** is an association of *object* $o \in O$ with *value* $v \in V$ of *attribute* $a \in A$:

$$\text{object} \xrightarrow{a} \text{value}$$

An attribute can be seen as a mapping $a : O \to V$ from objects to values.

Example 5.

*o*bject	*a*ttribute	*v*alue
shark	dangerous	true
shark	locomotion	swims
shark	category	fish
canary	category	bird
canary	dangerous	false

Classification of Objects by Attributes

*o*bject	*a*ttribute	*v*alue
shark	dangerous	true
shark	locomotion	swims
shark	category	fish
canary	category	bird
canary	dangerous	false

Question 2. *What makes one object different from another?*

- Values of attributes can be used to classify objects:

$$\text{Fish} := \{\text{animals} : \text{swim and have gills}\}$$

- Objects are instances of classes (e.g. a shark is an instance of fish).
- Classes of objects can also be organised into hierarchies.

Frames

Definition 6 (Frame)**.** is a set of attributes A used to describe a class of objects. Each object is an instance of a frame (class).

Example 7.

animal	shark	salmon	canary
locomotion	swims	swims	flies
has	gills	gills	wings
colour	gray	gray	yellow
dangerous	true	false	false

Production Systems

Definition 8 (Production system (rule-based system))**.** A computer program based on a set of IF-THEN rules (*production rules*).

Example 9.

```
IF   saturday OR sunday                    THEN   go to cinema
IF   NOT (saturday OR sunday)              THEN   go to work
IF   go to cinema                          THEN   go outside
IF   go to work AND NOT at work            THEN   go outside
IF   NOT (can go outside)                  THEN   stay home
IF   good weather                          THEN   can go outside
IF   raining                               THEN   have an umbrella
IF   raining AND have an umbrella          THEN   can go outside
```

- The order of rules is not specified.
- The condition defines if a rule applies or not.
- Production systems have become popular in AI programming, such as expert systems, agents and cognitive models.

Summary of Knowledge Representation Approaches

Classification due to Mylopoulos and Levesque (1984):

Networks are used to represent objects and relations (associations) between them (e.g. semantic networks, ontologies).

Logical calculus is used, such as first-order predicate calculus, modal, temporal logics.

Structured data is used to represent classes of objects and relations between them (e.g. frames).

Procedural or algorithmic representations can encode how to solve particular problems (e.g. rule-based systems).

Reading

- Each group should prepare presentation on one of the following four papers:
  1. For knowledge acquisition in accounting and finance read Wagner, Otto, and Chung (2002)
  2. For probabilities in knowledge acquisition read Doyle (1988)
  3. For overview of knowledge representation read McCalla and Cercone (1983)
  4. For knowledge representation issues read Stanojevic and Vranes (2007)

References

Doyle, J. R. (1988). Probability problems in knowledge acquisition for expert systems. *Knowledge-Based Systems*, *1*(2), 114–120.

McCalla, G., & Cercone, N. (1983). Guest editors' introduction: Approaches to knowledge representation. *Computer*, *16*(10), 12–18.

Mylopoulos, J., & Levesque, H. J. (1984). An overview of knowledge representation. In M. L. Brodie, J. Mylopoulos, & J. W. Schmidt (Eds.), *On conceptual modelling: Perspectives from artificial intelligence, databases, and programming languages* (pp. 3–17). New York: Springer.

Stanojevic, M., & Vranes, S. (2007). Knowledge representation with SOUL. *Expert Systems with Applications*, *33*, 122–134.

Wagner, W. P., Otto, J., & Chung, Q. B. (2002). Knowledge acquisition for expert systems in accounting and financial problem domains. *Knowledge-Based Systems*, *15*(8), 439–447.