MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

University of Ruse

INTELLIGENT COMPUTER SYSTEMS

Svetlana Stefanova

September, 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

LECTURE 5

KNOWLEDGE BASE IN EXPERT SYSTEMS

1. Introduction
2. Knowledge base design
3. Models for knowledge representation

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Subject area

- **Definition** - a set of objects, events, relationships and connections between them that make up the necessary KB, sufficient enough to solve the tasks in this field at the level of a good expert.
- **Includes** real and abstract objects with some relation and having certain properties that are expressed through propositions.

The analysis and study of the subject area allows building a KB.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Questions, related to KB

- **Representation** - how to present the knowledge of the subject area in the computer program?
- **Use** - how to use the knowledge to solve tasks at an expert level?

Some ES use KB that only allow reading.

With modern ES, it should be possible to delete and write in the KB.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Knowledge technologist

During KB construction, the expert is assisted by someone who understands the syntax, the rule interpreter, the KB construction process, and the practical psychology of interacting with it. This specialist is called a **knowledge technologist**.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

KB in traditional programs vs ICS

- **Traditional programs** - work with information, organized in the form of a DB, in which it is possible to extract only information that is specified in an explicit form.
- **ICS** – work with KB where reasoning/conclusions are possible, as a result of which new information can be generated that is not clearly present in the base.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Specifics of expertise

- **Wording/Formulation** - expertise is not always clearly codified and exists in many forms. The task of the knowledge technologist involves bringing together what is known about the task with its transformation into the system.
- **Certainty** - expertise is largely heuristic and not well formalized, i.e. knowledge is not always certain.
- **Accessibility** - because it is not well formalized, the expert's knowledge is not always easily accessible.
- **Completeness** - the presence of incompleteness, as facts and heuristics change with increasing experience.
- **Terminology** - due to the variety of sources of expertise, the expert uses terminology that varies (from broad concepts to precisely defined theoretical terms).

Co-financed by the European Union
Connecting Europe Facility

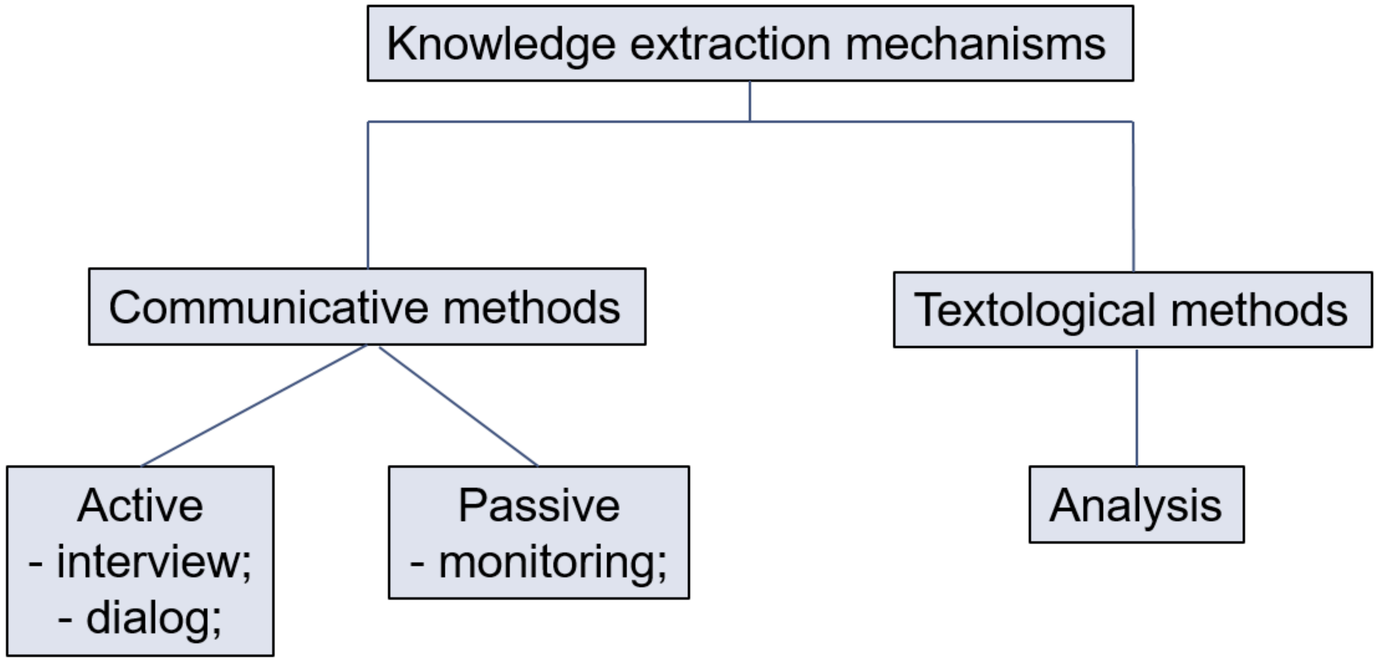
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

KB designing - stages

- **Acquisition** - through different techniques and approaches and from different sources. Some of the most reliable sources are experts from the given subject area.
- **Analysis** - includes arrangement, evaluation and filtration, which are mainly carried out by the knowledge technologist.
- **Representation** - various schemes and approaches are used to formally represent knowledge on a carrier with aim for their further systematization, analysis, supplementation and approximation to some machine form (close to the constructs of a logical programming language).

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

1st stage: knowledge acquisition



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Knowledge acquisition - mechanisms

- **Textological** - extracting knowledge from documents and specialized literature.
- **Communicative** - methods of contact of the knowledge technologist with the expert:

➢ **Active -** the initiative is in the hands of the knowledge technologist, who actively contacts the expert in various ways (conversations, dialogues, games, etc.);

➢ **Passive** - the leading role belongs to the expert, and the knowledge technologist only protocols his reasoning during real work or records what the expert considers necessary to talk about during separate sessions.

Different methods are usually combined.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Knowledge acquisition - automation

The problems associated with interviewing the experts are the reason for the search for methods and means of automating the process.

Automation is sought in two directions:

- creation of software tools for **automatic extraction** of knowledge in the process of dialogue between the software system and the expert.
- creation of software tools for **machine learning and self-learning** - the system is trained on the basis of examples, analogies, analysis of experimental data, etc.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

2nd stage: knowledge analysis

Includes:

- **Verifications;**
- **Coordinating uneven constructive elements –** combining large narrative reports with precise statements about facts.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Knowledge analysis – types

- **Logical** - verification for completeness and consistency.
- **Structural** – verification whether each rule subjects to the adopted structure;
- **Static** - verification for weaknesses in the KB (eg, one rule includes another or contradicts another).

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3rd stage: knowledge representation

Division by type:

- basic knowledge of the subject area;
- specific knowledge about the task.

Unified data structures – a homogeneous knowledge representation facilitates the system designer in developing the knowledge acquisition and explanation modules.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

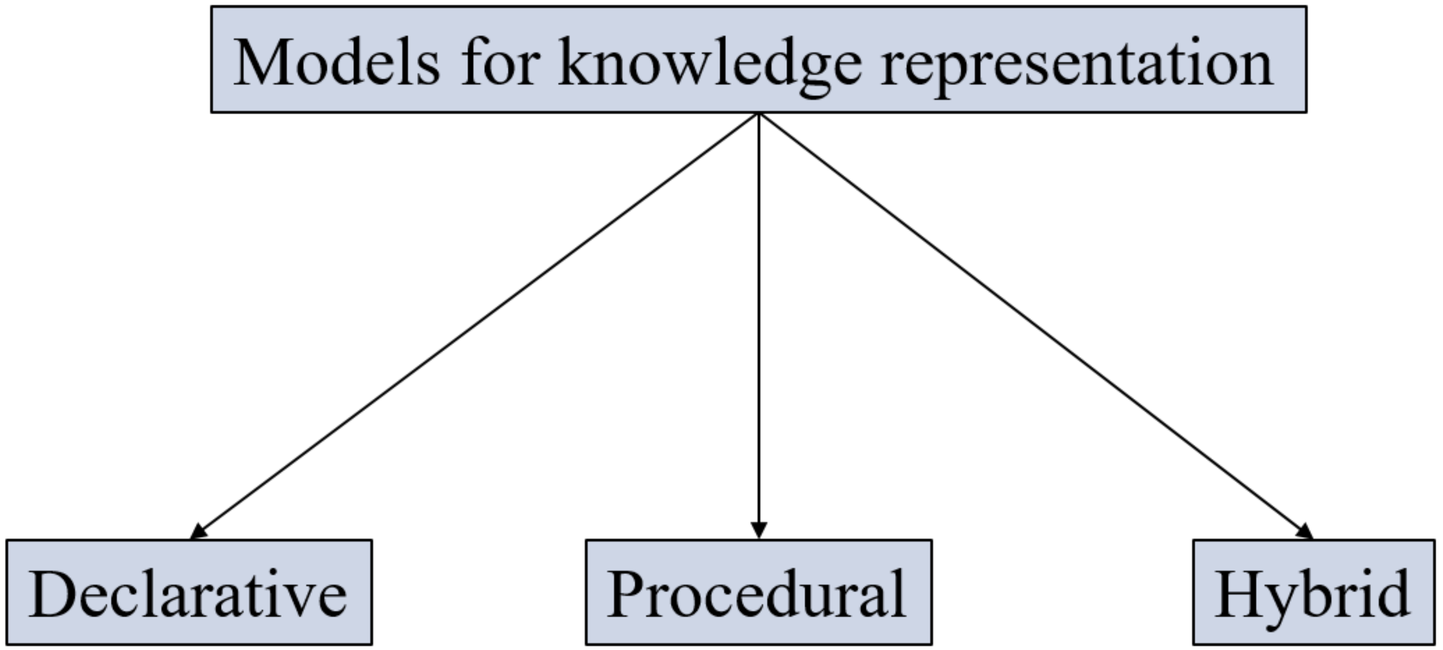
Knowledge specifics

- **Incompleteness** - in medical diagnostics, for example, the doctor usually has to act before all the incoming tests have been done.
- **Ability to process redundancy** - different knowledge can lead to the same interpretations or to mutually exclusive interpretations.
- **Uncertainty** - sources of information are not infallible.
- **Noise** - data can contain noise for a number of reasons: electronic noise, misreading of indications, transcription errors, etc.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Models for knowledge representation



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Declarative representation of knowledge

- **Declarative representation** – a simple statement that the fact is true.
- Knowledge is given as **descriptive** (declarative), where the individual units of knowledge are described independently of one another:
  - ➢ allows for **modularity** when organizing the KB, which facilitates its update;
  - ➢ the lack of hierarchy in the presentation of knowledge leads to difficulties in using it.

- **Declarative models are suitable in tasks where**:
  - ➢ the subject area is unstructured;
  - ➢ there are no clear links and relationships between the objects;
  - ➢ inaccurate knowledge is handled;
  - ➢ the knowledge of the subject area is predominantly of an empirical character.
- Convenient for experts and users.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Procedural representation of knowledge

- **Procedural representation** – a set of instructions without an explicit statement.
- More effective to use, but harder to maintain.
- Knowledge of the subject area is presented **clearly**. Availability for an advanced internal structure of the KB, which improves the efficiency of its use, but it requires strict adherence to certain rules and detailing of the knowledge.

Procedural models are suitable in tasks where:

- the subject area is highly structured;
- there are dependencies and relations between objects;
- the knowledge is mainly relations between objects and meta-knowledge.

Convenient for the knowledge technologists.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Hybrid representation of knowledge

- Knowledge is represented by two or more of the known methods at the same time.
- Declarative and procedural representation are alternative strategies that reach the same result. Any procedural representation can become a declarative and vice versa.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Requirements for the knowledge representation models

- **naturalness of presentation**;
- **clear semantics** – each sentence to be plainly interpreted;
- **correctness of the rules of inference** - if the knowledge given in the KB is true, then all the new knowledge, derived from it, is also true;
- **modularity**;
- **efficiency in terms of memory and time requirements**.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Basic models for knowledge representation

- Semantic networks;
- The triad "OBJECT-ATTRIBUTE-VALUE" (OAV);
- Frames;
- Rules.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks

- A set of objects, called **nodes,** connected to each other by **arcs**;
- Can be considered as a **graph**.
- Usually, both arcs and nodes have **labels**, i.e. are named and no limitations on node names and arcs are required.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks - nodes

- **Objects:**
  - Physical – the objects that surround us;
  - Conceptual - can be actions, events, or abstract categories. Example: “Bulgaria" and “2” may be nodes from a geographic KB.
- **Descriptors** - provide additional information about the objects. Example, “beautiful" and “small“ contain information about the country Bulgaria.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks – arcs (links)

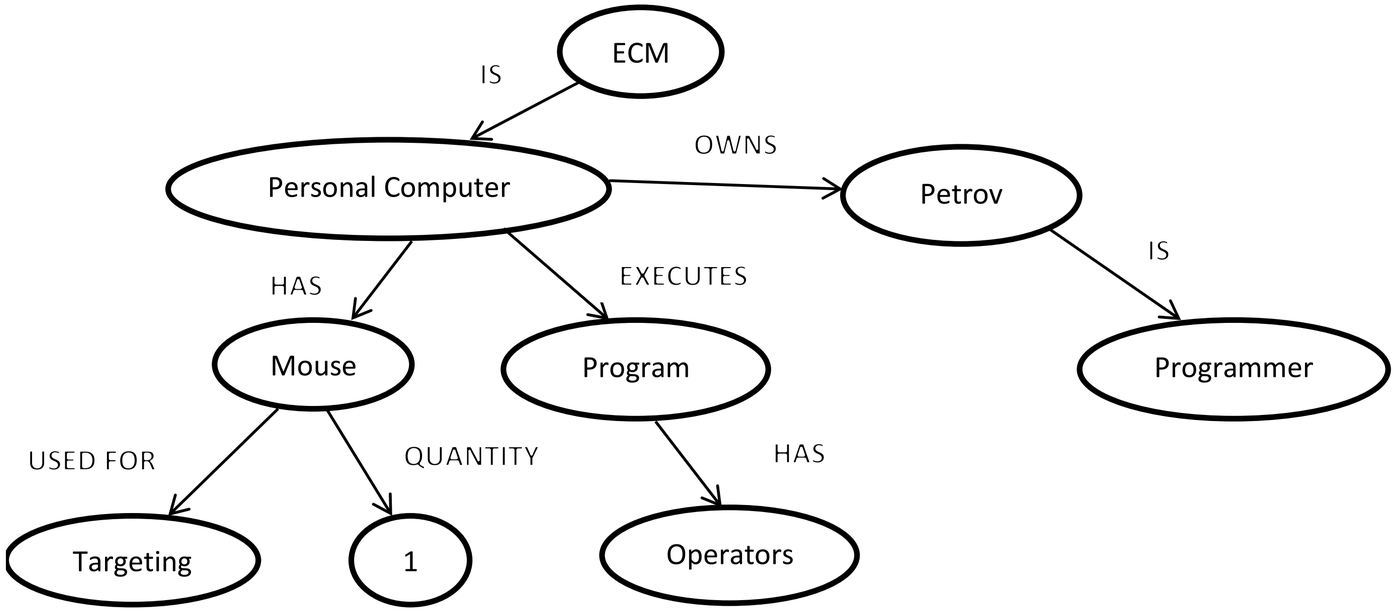
Common arcs are :

- **“IS”** - represents a "class-type" relationship. Example, "The wolf **is** a predator”;
- **“HAS”** – identifies the nodes that are properties of other nodes and it shows a relationship between a part and a subpart. Example, "The wolf **has** paws”;
- **definitive** – defines some kind of a statement. Example, “Wolves **love** meat”.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks - example



Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks – syntactic approaches for representation

- **object i – link ij – object j** - more natural to the language ***LISP***;
- **link ij (object i, object j)** - more natural for use in the ***PROLOG*** logical programming language.
- A combination of the two approaches is possible.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks - advantages

- **Simplicity**;
- **Visibility**;
- **Flexibility** - new nodes and arcs can easily be defined;
- **Inheritance** – a specimen of a given class has the properties of the more general classes that it is a member of. It is an application of the "**IS**" relationship.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Semantic networks - disadvantages

- **Difficulty in presenting n-array relationships**. ***Simon's method*** – a given *n*-array relationship is transformed into a system of binary relatives by introducing a new object *n+1* into the network that represents the whole relation. New binary relations are then introduced to describe the links between the new object and each of the original arguments.
- **Difficulty in presenting knowledge with quantum** – separated semantic networks can be used (an hierarchical set of subnets, each of which corresponds to the scope of a variable). Some special arcs in the network do not point to nodes, but to entire semantic networks that are sub-networks of the given one.
- **Problem when executing operations with semantic networks**.
- **Problem of the two Smiths (different objects with the same name)**.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Triad "OBJECT-ATTRIBUTE-VALUE“ (OAV)

Used in the MYCIN expert system.

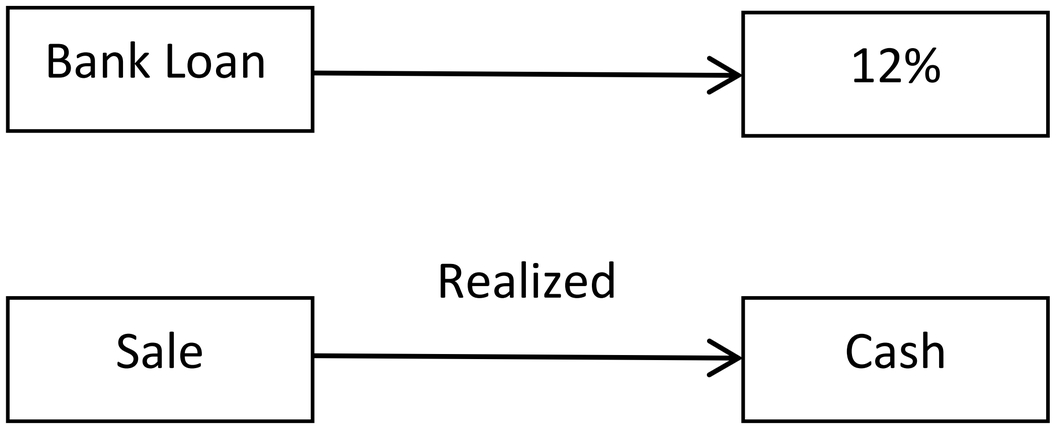
- **Objects -** can be physical (door, transistor) or conceptual (bank loan, sale).
- **Attributes** - main features or properties associated with the objects. For example, size, shape, and color for physical objects or values of interest, payment term, size for the conceptual object “bank loan”.
- **Value** - indicates the specific nature of the attribute in the given situation. For example, the apple's color may be red, and the bank loan rate may be 10%.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Triad OAV - examples



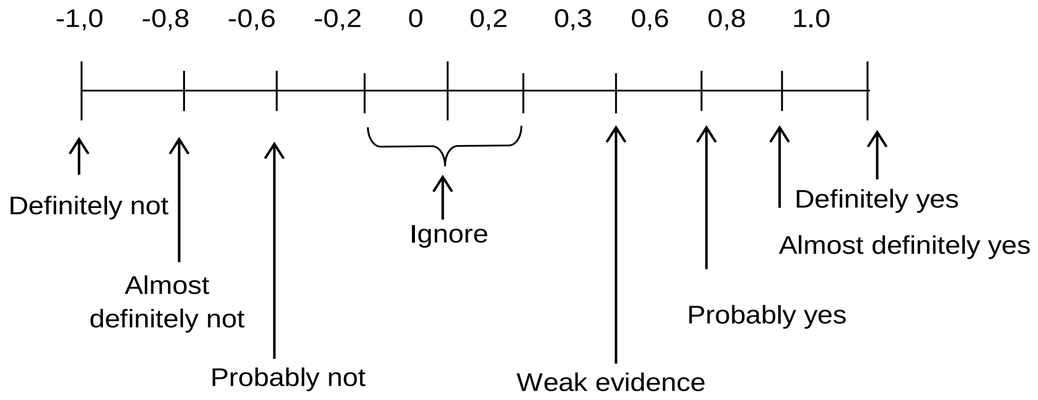
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Triad OAV – working with uncertain facts

If the fact is uncertain, OAV could be modified with a number, called a **factor/coefficient of reliability/confidence,** that correlates to the truth of a statement.

The factor of reliability **is not a probability**, but an informal **measure of trust** or **security** towards a given evidence. It shows the extent to which it is certain that the evidence is true.



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Triad OAV – advantages and disadvantages

Advantages:

- Objects can be arranged in the form of a graph, which means **generating and inheriting properties**;
- Ability to **handle uncertain data**;
- Ability to use an **OO (object-oriented) approach**.

Disadvantages:

- Knowledge is presented **fragmentarily;**
- **Links are in an implicit form**;
- There are **difficulties in tracking the links between objects** in a large KB;

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Frames

- **Protoframe** - a solid structure of the information unit.
- **Frame-exemplar** – obtained from the protoframe by assigning specific names and values to the slots.

Name of the frame:

Name slot 1:	*Value of slot 1*
Name slot 2:	*Value of slot 2*
……	……
Name slot N:	*Value of slot N*

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Frames – values of the slots

- numbers or mathematical ratios;
- text in natural language or programs;
- a set of rules or procedures for obtaining values;
- default values;
- citing of slots from other frames;
- pointers to other frames;
- a pointer to a set of slots of lower level frames.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Example frame

No	Description of goods / service	Catalogue No	Warranty months	Number	Unit price w/o VAT	Total price with VAT	Currency
1	Motherboard - GB H110M-DS2 / LGA1151	#Q0066010027505	36	1	88,07	88,07	BGN
2	Processor - I3-7100 3.9GHZ/3MB/LGA1151/BOX	#Q0066010027506	36	1	261,12	261,12	BGN
3	Memory - 8G DDR4 2400 ADATA	#Q0066010027507	60	1	95,62	95,62	BGN
4	Hard Drive - ADATA SSD SU650 240GB 3D NAND	#Q0066010027508	36	1	65,31	65,31	BGN
5	Box - CM MASTERBOX LITE 3	#Q0066010027509	24	1	48,31	48,31	BGN
6	Power Supply - PSU FORTRON FSP350-50AHBCC	#Q0066010027510	24	1	50,49	50,49	BGN
7	Software product with license sticker - Microsoft Windows Pro 10 64Bit Eng Intl 1pk DSP DVD	FQC-08929	0	1	259,59	259,59	BGN
Version - Windows 10 Type of license - OEM (License and media (DVD)) Language - English Bit - 64 Other - For sale with computer configuration							

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Frames – advantages and disadvantages

Advantages :

- possibility of a **declarative and procedural presentation** of knowledge;
- **inheritance**;
- inclusion of **default values**;
- possibility **for interconnections** between the frames.

Disadvantages:

- more difficult to **understand**;
- higher requirements for **computing resources**.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Rules

Predicate - the conditional part of the rule or prerequisite.

Rules can be:

- **simple** - contain one simple statement;
- **complex** - contain more than one condition, connected with logical relations of the types AND, OR, NO.

The rules of the given type are called **production**, as well as the systems based on them.

Credibility factors can relate to rules just as they relate to facts.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Rules – advantages and disadvantages

Advantages:

- Representation is in a **descriptive form**;
- Ability to **handle insecure data**;
- Ability to **merge facts** into one rule.

Disadvantages:

- Knowledge is **difficult to formalize** and represent;
- **The volume** of the KB is significantly higher.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423