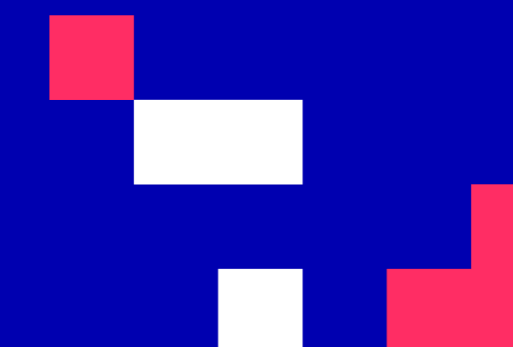


Русенски университет

ИНТЕЛИГЕНТНИ КОМПЮТЪРНИ СИСТЕМИ

Светлана Стефанова

Септември, 2022



ЛЕКЦИЯ 5**БАЗАТА ЗНАНИЯ В ЕКСПЕРТНИТЕ СИСТЕМИ****СЪДЪРЖАНИЕ**

1. Въведение
2. Проектиране на база знания
3. Модели за представяне на знания



СЪДЪРЖАНИЕ 1

Предметна област

- **Дефиниция** - съвкупност от обекти, събития, отношения и връзки между тях, които съставят необходимата БЗ, достатъчна за решаване на задачите от тази област на ниво добър експерт.
- **Включва** реални и абстрактни обекти, намиращи се в някакви отношения и имащи определени свойства, които се изразяват чрез съждения.

Анализът и изучаването на предметната област позволява да се изгради БЗ.



СЪДЪРЖАНИЕ 1

Въпроси, свързани с БЗ

- **Представяне** - как знанията за разглежданата област да се представят в компютърната програма?
- **Използване** - как знанията да се използват за решение на задачите на експертно ниво?

В някои ЕС се използват БЗ, които допускат само четене.

При съвременните ЕС в БЗ трябва да може да се изтрива и записва.

СЪДЪРЖАНИЕ 1

Технолог на знанията

По време на конструиране на БЗ експертът се подпомага от някой, който разбира синтаксиса, интерпретатора на правилата, процеса на конструиране на БЗ и практическата психология на взаимодействието с него. Този специалист се нарича **технолог на знанията**.



СЪДЪРЖАНИЕ 1

БЗ в традиционни програми vs ИКС

- **Традиционни програми** - работят с информация, организирана във вид на БД, при които е възможно извличане само на информация, зададена в явен вид.
- **ИКС** – работят с БЗ, където са възможни разсъждения/изводи, като в резултат може да бъде генерирана нова информация, която не присъства в явен вид в базата.



СЪДЪРЖАНИЕ 1

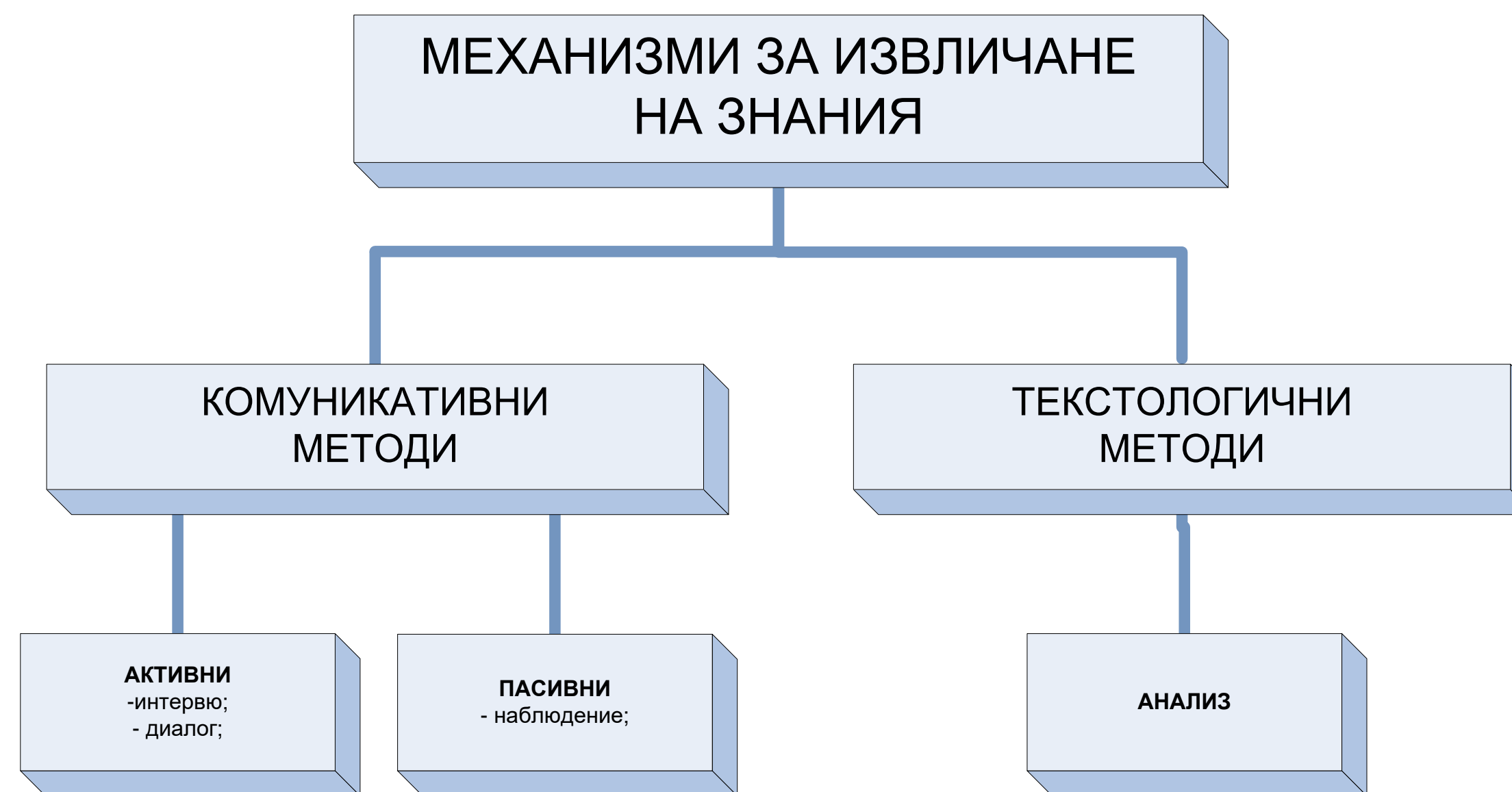
Специфики на експертния опит

- **Формулировка** - експертният опит невинаги е ясно кодиран и съществува в много форми. Задачата на технолога на знания включва обединяване на това, което е известно за задачата, с неговото трансформиране в системата.
- **Сигурност** - експертният опит до голяма степен е евристичен и не добре формализиран, т.е. знанията не винаги са сигурни.
- **Достъпност** - тъй като не са добре формализирани, знанията на експерта не винаги са леснодостъпни.
- **Пълнота** - наличието на непълнота, тъй като фактите и евристиките се променят с нарастване на опитността.
- **Терминология** - поради многообразието на източници на експертен опит експертът използва терминология, която варира (от широки понятия до точно дефинирани теоретични термини).

СЪДЪРЖАНИЕ 2

Проектиране на БЗ - етапи

- **Придобиване** - чрез различни техники и подходи и от различни източници. Едни от най-надеждните източници са експертите от дадената предметна област.
- **Анализ** - включва подредба, оценка и филтрация, които се осъществяват основно от технолога на знания.
- **Представяне** – използват се различни схеми и подходи за формално изобразяване на знанията върху носител с оглед на тяхното по-нататъшно систематизиране, анализ, допълване и приближаване към някаква машинна форма (близка до конструкциите на език за логическо програмиране).

СЪДЪРЖАНИЕ 2**1 етап: придобиване на знания**

СЪДЪРЖАНИЕ 2

Придобиване на знания - механизми

- **Текстологични** - извличане на знания от документи и специализирана литература.
- **Комуникативни** - методи за контакт на технолога на знанията с експерта:
 - **Активни** - инициативата е в ръцете на технолога на знания, който активно контактува с експерта по различни начини (беседи, диалози, игри и т.н.);
 - **Пасивни** - водещата роля е на експерта, а технологът на знанията само протоколира разсъжденията му по време на реална работа или записва това, което експертът счита за необходимо да разкаже по времето на отделни сесии.

Обикновено се комбинират различни методи.

СЪДЪРЖАНИЕ 2

Придобиване на знания - автоматизация

Проблемите, свързани с интервюирането на експертите, са причина за търсене на методи и средства за автоматизация на процеса.

Автоматизация се търси в две направления:

- създаване на програмни средства за **автоматично извличане** на знания в процес на диалог между програмната система и експерта.
- създаване на програмни средства за **машинно обучение и самообучение** - системата се обучава на базата на примери, аналогии, анализ на експериментални данни и др.

СЪДЪРЖАНИЕ 2

2 етап: анализ на знанията

Включва:

- Проверки;
- Съгласуване на неравномерни градивни елементи –
комбиниране на големи описателни доклади с точни твърдения за
факти.



СЪДЪРЖАНИЕ 2

Анализ на знанията – видове

- **Логически** – проверка за цялостност и непротиворечивост.
- **Структурен** – проверка дали всяко правило се подчинява на възприетата структура;
- **Статичен** – проверка за слаби места в БЗ (напр., едно правило да включва друго или да противоречи на друго).



СЪДЪРЖАНИЕ 2

3 етап: представяне на знания

Типово разделяне:

- базови знания за предметната област;
- специфични знания за задачата.

Унифицирани структури от данни – едно хомогенно представяне на знанията улеснява конструктора на системата при разработване на модулите за придобиване на знания и за обяснение.

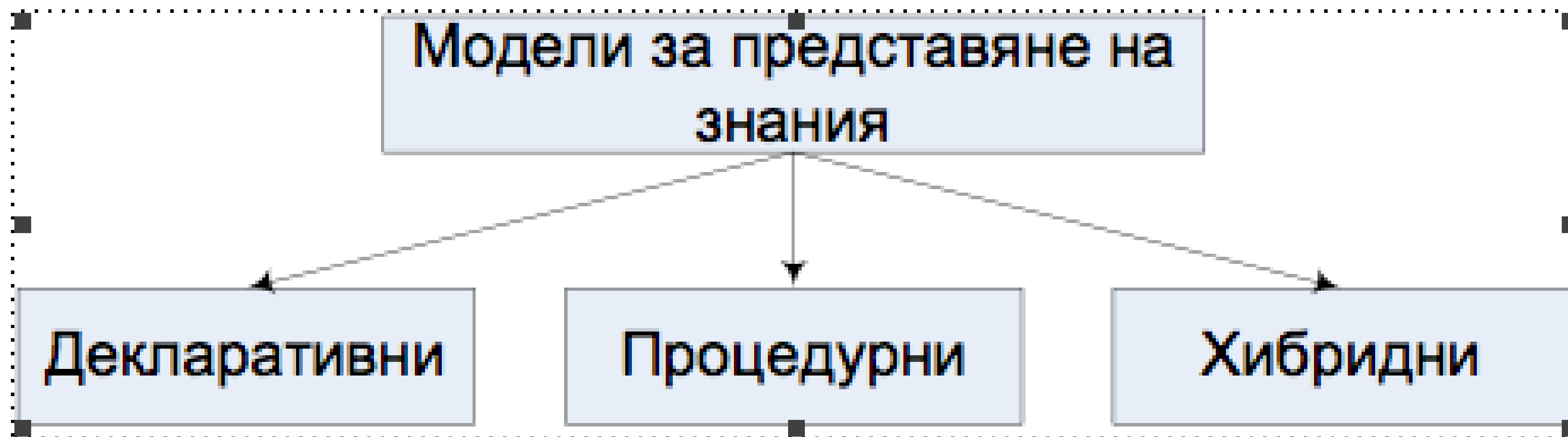
СЪДЪРЖАНИЕ 2

Специфика на знанията

- **Непълнота** - в медицинската диагностика, например, лекарят обикновено трябва да действа преди да са направени всички входящи тестове.
- **Способност да се обработва излишък** - различни знания могат да водят до еднакви интерпретации или до взаимно изключващи се интерпретации.
- **Несигурност** - източниците на информация не са безгрешни.
- **Шумове** - данните могат да съдържат шум по редица причини: електронен шум, неправилно прочитане на индикации, грешки при преписване и др.

СЪДЪРЖАНИЕ 3

Модели за представяне на знания



СЪДЪРЖАНИЕ 3

Декларативно представяне на знания

- **Декларативно представяне** - просто твърдение, че фактът е верен.
- Знанията се задават **описателно** (декларативно), като отделните единици знания се описват независимо една от друга:
 - позволява **модулност** при организирането на БЗ и облекчава актуализацията ѝ;
 - липсата на йерархия при представянето на знанията води до затруднения при използването им.
- **Задачи, за които са подходящи декларативните модели:**
 - с неструктурирана предметна област;
 - без ясни взаимовръзки и отношения между обектите;
 - боравещи с неточни знания;
 - знания с предимно емпиричен характер.
- Удобен за експерти и потребители.



СЪДЪРЖАНИЕ 3

Процедурно представяне на знания

- **Процедурно представяне** - множество от инструкции без явно твърдение.
- По-ефективни за употреба, но по-трудни за поддържане.
- Знанията за предметната област се представят **явно**. Възможност за развита вътрешна структура на БЗ, което подобрява ефективността при използването ѝ, но изисква стриктно спазване на определени правила и детайлизиране на знанията.

Задачи, за които са подходящи процедурните модели:

- силно структурирана предметна област;
- съществуващи зависимости и релации между обектите;
- знанията са предимно връзки между обектите и мета-знанията.

Удобно за технолозите на знания.



СЪДЪРЖАНИЕ 3

Смесено представяне на знания

- Знанията се представят чрез два или повече от известните методи едновременно.
- Декларативното и процедурното представяне са алтернативни стратегии, стигащи до един и същ резултат. Всяко процедурно представяне може да се превърне в декларативно и обратно.



СЪДЪРЖАНИЕ 3

Изискванията към моделите за представяне на знания

- **естественост на представянето;**
- **ясна семантика** – всеки израз да се интерпретира еднозначно;
- **коректност на правилата за извод** – ако знанията, зададени в БЗ, са верни, то всички изведени от тях нови знания също да са верни;
- **модулност;**
- **ефективност по отношение на памет и време.**



СЪДЪРЖАНИЕ 3

Основни модели за представяне на знания

- Семантични мрежи;
- Тройка “ОБЕКТ-АТРИБУТ-СТОЙНОСТ” (ОАС);
- Фреймове;
- Правила.



СЪДЪРЖАНИЕ 3

Семантични мрежи

- Множество от обекти, наречени **възли**, свързани помежду си с **дъги** (**връзки**);
- Могат да се разглеждат като **граф**.
- Обикновено и връзките, и възлите имат **етикети**, т.е. са наименовани, като ограничения върху имената не се налагат.



СЪДЪРЖАНИЕ 3

Семантични мрежи - възли

- **Обекти:**
 - Физически - предметите, които ни заобикалят;
 - Концептуални - могат да бъдат събития или абстрактни категории. Например: “България” и “2” могат да са възли от географска БЗ.
- **Описатели** - дават допълнителна информация за обектите. Например, “красива” и “малка” съдържат информация за държавата България.



СЪДЪРЖАНИЕ 3

Семантични мрежи - дъги (връзки)

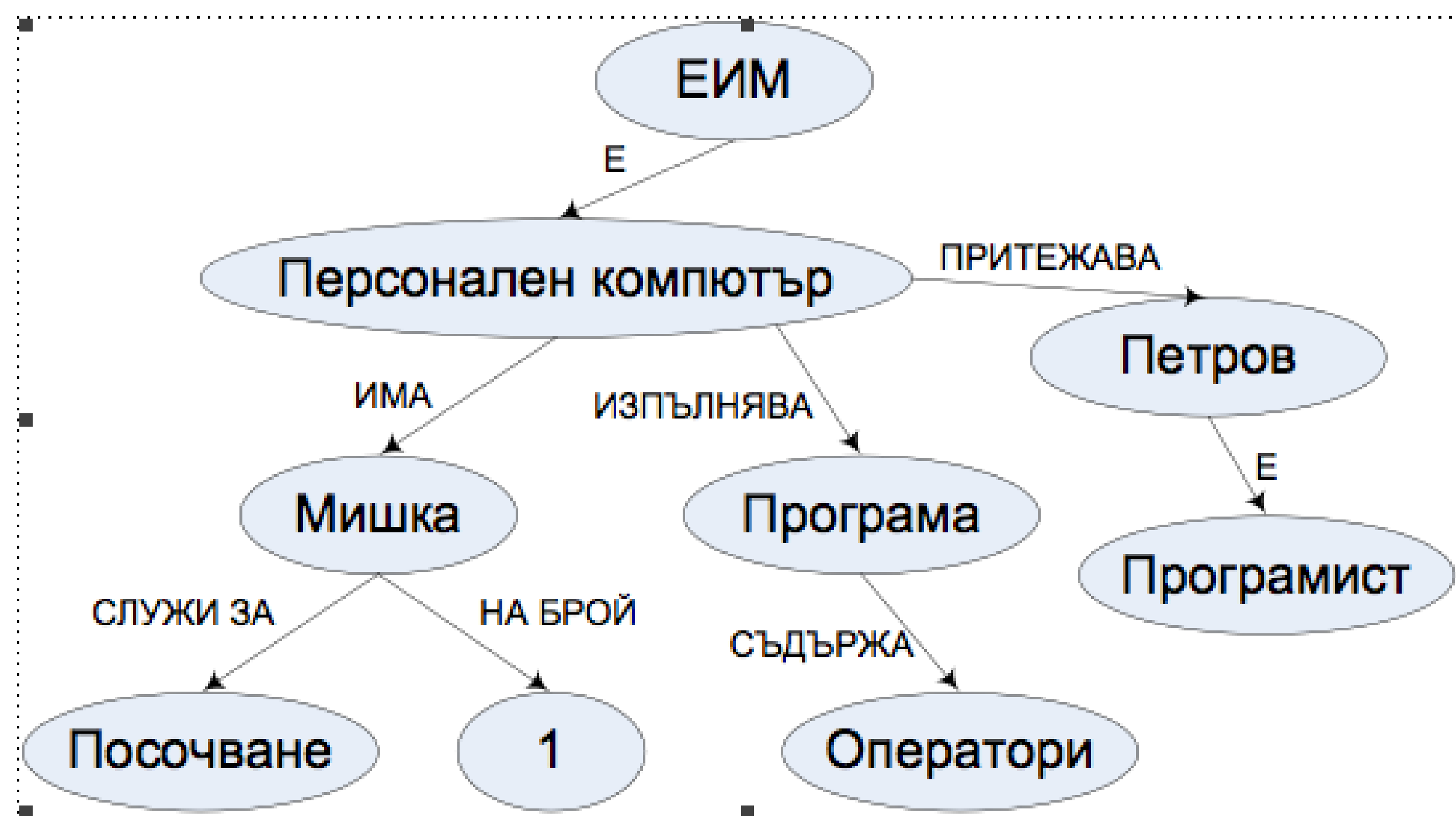
Често срещани връзки:

- **“Е”** - представя взаимоотношение от вида “клас-екземпляр”.
Пример, “Вълкът **е** хищник”;
- **“ИМА”** - идентифицира възлите, свойства на други възли, и показва взаимоотношение между част и подчаст. Пример, “Вълкът **има** лапи”;
- **дефиниционни** – дефинира някакво твърдение. Пример, “Вълците **обичат** месо”.



СЪДЪРЖАНИЕ 3

Семантични мрежи - пример



СЪДЪРЖАНИЕ 3

Семантични мрежи - синтактични подходи при представяне

- обект i – връзка ij – обект j - естествен за езика *LISP*;
- връзка ij (обект i , обект j) - естествен за езика за логическо програмиране *PROLOG*.
- Възможна е комбинация от двата подхода.



СЪДЪРЖАНИЕ 3

Семантични мрежи - предимства

- Простота;
- Нагледност;
- **Гъвкавост** – лесно могат да се дефинират нови възли и връзки;
- **Наследяване** - екземпляр от даден клас притежава свойствата на по-общите класове, на които е член. Приложение е на взаимоотношението “Е”.



СЪДЪРЖАНИЕ 3

Семантични мрежи - недостатъци

- **Трудност при представяне на произволни n -арни релации.**
Метод на Саймънс - n -арната релация се трансформира в система от бинарни релации чрез въвеждане на нов обект $n+1$ в мрежата, който представя цялата релация. След това се въвеждат нови бинарни релации за описване на връзките между този нов обект и всеки от оригиналните аргументи.
- **Трудност при представяне на знания, съдържащи квантори** – за целта могат да се използват разделени семантични мрежи (йерархична съвкупност от подмрежи, всяка от които съответства на областта на действие на една променлива). Някои специални дъги в мрежата сочат не към възли, а към цели семантични мрежи, които са подмрежи на дадената.
- **Проблем при извършване на операции със семантични мрежи.**
- **Проблем на двамата Смит (различни обекти, имащи еднакво име).**



СЪДЪРЖАНИЕ 3

Тройка “ОБЕКТ-АТРИБУТ-СТОЙНОСТ” (ОАС)

Използва се в експертната система MYCIN.

- **Обекти** - могат да бъдат физически (врата, транзистор) или концептуални (банков заем, продажба).
- **Атрибути** - основни характеристики или свойства, свързани с обектите. Например, големина, форма, цвят при физически обект или стойност на лихвата, срок на погасяване, размер при концептуален обект “банков заем”.
- **Стойност** - указва специфичната природа на атрибута в дадена ситуация. Например, цветът на ябълката може да е червен, а стойността на лихвата на банков заем да е 10%.



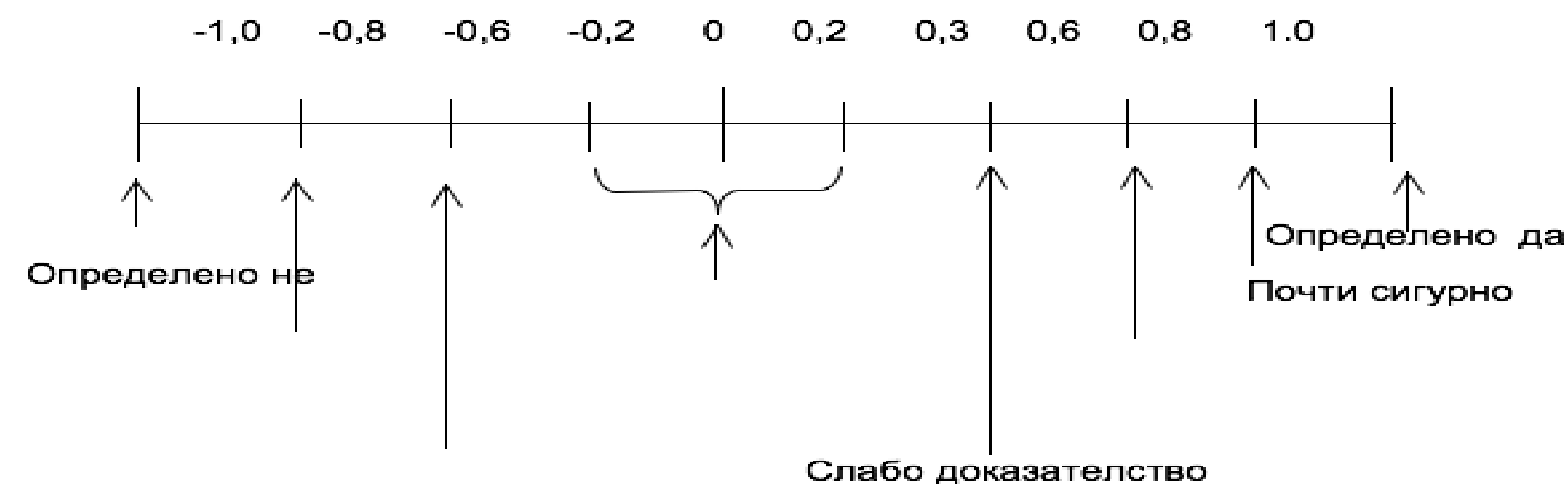
СЪДЪРЖАНИЕ 3**Тройка ОАС - примери**

СЪДЪРЖАНИЕ 3

Тройка ОАС - работа с несигурни факти

Ако фактът е неопределен, ОАС може да се модифицира с число, наречено **фактор/коефициент на достоверност/сигурност**, свързващ се с истинността на едно твърдение.

Факторът на достоверност **не е вероятност**, а неформална **мярка за доверие** или **сигурност** към дадено доказателство. Показва степента, до която е сигурно, че доказателството е вярно.



СЪДЪРЖАНИЕ 3

Тройка ОАС – предимства и недостатъци

Предимства:

- Обектите могат да се подреждат във вид на граф, което означава **пораждане и наследяване на свойства**;
- Възможност за **боравене с несигурни данни**;
- Възможност за използване на **ОО подход**.

Недостатъци:

- Знанията се представят **фрагментарно**;
- **Връзките са в неявен вид**;
- Съществуват **затруднения при проследяване на връзките между обектите** в голяма БЗ;

СЪДЪРЖАНИЕ 3

Фреймове

- **Протофрейм** - твърда структура на информационната единица.
- **Фрейм-екземпляр** – получава се от протофрейма, като на слотовете се присвояват конкретни имена и стойности.

Име *на*
фрейма:

Име слот 1:

Стойност на слот 1

Име слот 2:

Стойност на слот 2

.....

.....

Име слот N:

Стойност на слот N

□

СЪДЪРЖАНИЕ 3

Фреймове - стойности на слотове

- числа или математични съотношения;
- текст на естествен език или програми;
- множество от правила или процедури за получаване на стойности;
- стойности по премълчаване;
- цитиране на слотове от други фреймове;
- указатели към други фреймове;
- указател към набор от слотове на фреймове от по-ниско ниво.

СЪДЪРЖАНИЕ 3

Примерен фрейм

No	Описание на стока / услуга	Каталожен No	Гаранция месеци	Брой	Ед. цена без ДДС	Обща сума без ДДС	Валута
1	Дънна платка - GB H110M-DS2 / LGA1151	#Q0066010027505	36	1	88,07	88,07	BGN
2	Процесор - I3-7100 3.9GHZ/3MB/LGA1151/BOX	#Q0066010027506	36	1	261,12	261,12	BGN
3	Памет - 8G DDR4 2400 ADATA	#Q0066010027507	60	1	95,62	95,62	BGN
4	Твърд диск - ADATA SSD SU650 240GB 3D NAND	#Q0066010027508	36	1	65,31	65,31	BGN
5	Кутия - CM MASTERBOX LITE 3	#Q0066010027509	24	1	48,31	48,31	BGN
6	Захранване - PSU FORTRON FSP350-50AHBCC	#Q0066010027510	24	1	50,49	50,49	BGN
7	Програмен продукт с лицензен стикер - Microsoft Windows Pro 10 64Bit Eng Intl 1pk DSP DVD	FQC-08929	0	1	259,59	259,59	BGN
Версия - Windows 10 Тип на лиценза - OEM (Licence and media (DVD)) Език - English Bit - 64 Други - For sale with computer configuration							



СЪДЪРЖАНИЕ 3

Фреймове – предимства и недостатъци

Предимства :

- възможност за **декларативно и процедурно представяне** на знанията;
- **наследяване**;
- включване на **стойности по премълчаване**;
- възможност за **връзки** между фреймовете.

Недостатъци:

- по-трудна **разбираемост**;
- по-високи изисквания към **изчислителните ресурси**.

СЪДЪРЖАНИЕ 3

Правила

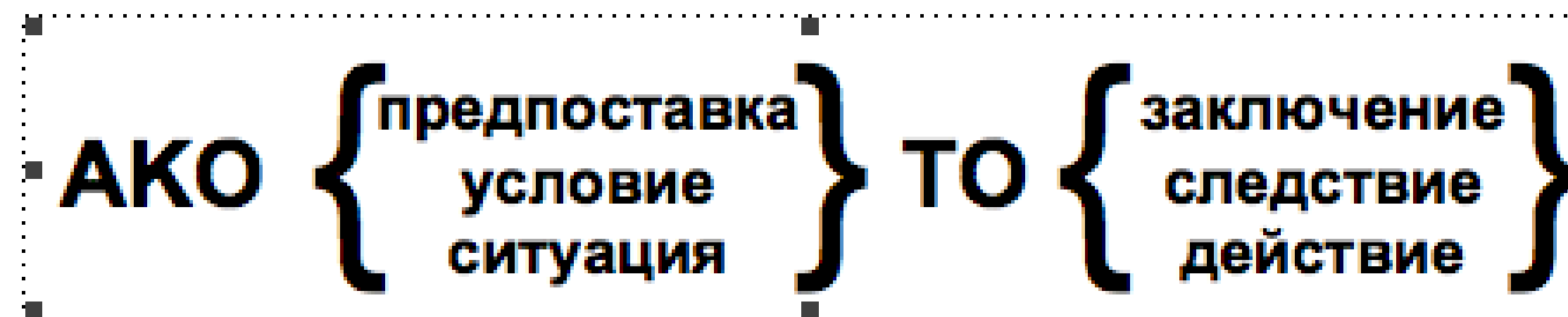
Предикат - условната част на правилото или предпоставката.

Правилата могат да бъдат:

- **прости** - съдържат едно елементарно твърдение;
- **сложни** - съдържат повече от едно условие, свързвани с логически релации от типа И, ИЛИ, НЕ.

Правилата от разгледания тип се наричат **продукционни**, както и системите, базираци се на тях - продукционни.

Факторите на достоверност могат да се свържат и с правилата, както се свързват с фактите.



СЪДЪРЖАНИЕ 3

Правила – предимства и недостатъци

Предимства:

- Представянето е в **описателен вид**;
- Възможност за **боравене с несигурни данни**;
- Възможност за **обединяване на факти** в едно правило.

Недостатъци:

- Знанията **трудно се формализират** и представят;
- **Обемът** на БЗ е значително по-голям.

Благодаря ВИ.

