Ismeretalapú modellezés

Tudnivalók

- Honlap: people.inf.elte.hu/gt/ism/ism.html
- Szakirodalom
  - Futó I. (szerk.): *Mesterséges intelligencia,* AULA Kiadó, 1999.
  - Russel, S.J., Norwig, P.: *Mesterséges intelligencia modern megközelítésben.* Panem Bp., 2005.
  - Sántáné-Tóth E.: *Tudásalapú technológia, szakértő rendszerek,* Dunaújv. Főisk., 2000. 2007.
  - Sántáné-Tóth E., Biró M., Gábor A., Kő A., Lovrics L., (szerk.: Sántáné-Tóth E.): *Döntéstámogató rendszerek,* PANEM Könyvkiadó, Budapest, 2008.
- Számonkérés: vizsga

I. Bevezetés

Az előadás összeállításánál erősen támaszkodtam Sántáné-Tóth Edit korábbi előadásaira

Intelligens rendszerek

- Megoldandó feladatai: nehezek
  - A feladat problématere hatalmas, ezért a megoldás megtalálása intuíciót, kreativitást igényel.
- A megoldó szoftver viselkedése: intelligens
  - Mintha egy ember tevékenykedne (Turing teszt).
  - Működése közben tanul.
- Megoldási módszerei: sajátosak
  - Átgondolt reprezentáció a feladat modellezéséhez.
  - Heurisztikával megerősített hatékony algoritmusok.
  - Kereső rendszer algoritmus séma

IS ⊇ KBS ⊇ ES

- IS – Intelligens rendszerek
  - Intelligensen viselkedő számítógépes programok intellektuálisan bonyolult feladatok megoldására.
- KBS – Ismeretalapú rendszer (IR)
  - Olyan alkalmazás, amely a benne felhalmozott nagy terjedelmű ismeret hatékony, meghatározott technológia mentén megvalósuló kezelését végzi.
- ES – Szakértő rendszer (SZR)
  - Olyan szoftver,amely a tárgyköri szakértő ismereteinek felhasználásával magas szintű teljesítményt nyújt egy szűk problémakör kezelésében.

IR fejlesztése

- Az IR készítése során a szokásos szoftverfejlesztési teendőkön túl különös hangsúlyt kap:
  - a feladattal kapcsolatos ismereteket megszerezése
  - az ismeretek hatékony ábrázolása
  - az ismeretekre támaszkodó megoldó algoritmus tervezése és implementálása
  - valamint a fenti teendőket támogató módszereknek és eszközöknek (technológiának) a használata



Az IR intelligens elemei

- A megoldandó feladatainál
  - nagy terjedelmű szakterületi ismeretre van szükség. A *kombinatorikus robbanás* elkerülése ezen ismeretek megfelelő ábrázolásán és visszakereshetőségén múlik.
- Megoldási módszer
  - Speciális *kereső rendszer*, ahol a szakterületi ismeretek explicit formában, a kereső rendszertől elkülönülő komponensben, az ismeretbázisban foglalnak helyet.
  - Szimbólumok formájában tárolja az ismereteket; szimbolikus következtetésekkel dolgozik, és képes megmagyarázni a következtető lépéseit.

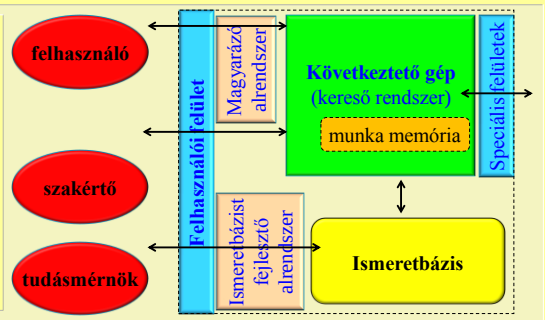
Kereső rendszer (KR)

Procedure *KR*
1. ADAT := *kezdeti érték*
2. ***while*** ¬*terminálási feltétel*(ADAT) ***loop***
3. ***select*** SZ ***from*** *alkalmazható szabályok*
4. ADAT := SZ(ADAT)
5. ***endloop***

end

ADAT	***SZ***	***SELECT***
globális munkaterület	*produkciós szabályok*	*vezérlési stratégia*

IR szerkezete



Ismeretreprezentáció szintjei

Tudati szint

Szimbólikus szint

Technikai szint

- Az emberi tudás, azaz a világ dolgainak tudati megjelenésének explicit, implicit és rejtett elemei.
- A megszerzett ismeret, az explicit (esetleg felszínre hozott implicit) tudás leírt, esetleg valamilyen reprezentációs módszerrel formalizált része.
- A szimbolikus leírás számítógépen rögzített (adatstruktúrákkal és algoritmusokkal ábrázolt) formája.

Ismeretek osztályozása

Felhasználás szerint:

- ***Deklaratív ismeretek*** – Tények, állítások.
- ***Procedurális ismeretek*** – Új ismeretek kikövetkeztetési lépéseit írják le.
- ***Vezérlési ismeretek*** – A megfelelő következtetési lépést választja ki. metaismeretek (heurisztikák)

Leírási mód szerint: deklaratív vagy procedurális

Egy ismeret felhasználása szerint lehet procedurális, de ettől még a leírása megvalósulhat deklaratív eszközökkel is.

„Nemcsak az általunk közölt ismeret, hanem annak kifejezésmódja is információt hordoz.”

Ha x és y egyaránt zérus, akkor a szorzatuk is az.

Ha x zérus de $x*y$ nem, akkor y sem zérus.

$x=0 \wedge y=0 \rightarrow x*y=0$

$x=0 \wedge x*y\neq 0 \rightarrow y\neq 0$

$Z(x) \wedge Z(y) \rightarrow Z(m(x,y))$

$Z(x) \wedge \neg Z(m(x,y)) \rightarrow \neg Z(y)$

$A \wedge B \rightarrow C$

$A \wedge \neg C \rightarrow \neg B$

$\neg A \vee \neg B \vee C$

$\neg A \vee C \vee \neg B$

Ismeretreprezentációs módszerek

- Procedurális reprezentáció
- Logikai alapú reprezentációk
  - Logikai nyelvek
  - Szabályalapú reprezentációk
- Strukturált reprezentáció
  - Szemantikus hálók és keretek (frame)
  - Döntési fák
- Hibrid reprezentációk

Miről lesz szó a félév során

- Szabályalapú technikák
  - Szabályalapú rendszerek
  - Szabály alapú logikai következtetés
    - klasszikus és nem-monoton
- Keretalapú technikák
  - Szemantikus hálók és frame-k
  - Leíró logika
- Induktív technikák
  - Döntési fák
  - Logikai leírás tanulása
- Esetalapú technikák

Ontológia
Technológia
Leírás formája
Bizonytalanság kezelés

Ismeretmodellezés az IR tervezésében

- Célkitűzéseinkkel harmonizáló ismeret megszerzése,
  - a szakterület (a tárgytartomány) lényeges elemeinek feltárása (knowledge acquisition)
  - Az ismeretek ábrázolására támaszkodó szimbolikus következtetési lépések meghatározása.
- Ismeretek ábrázolása, azaz megfelelő formában történő rögzítése ismeretszervezéstől a számítógépes reprezentációig (knowledge engineering). Lényeges eleme ennek az *ismeret visszakereshetősége*.
- A következtetést vezérlő ismeretek kinyerése.

Ismeretszerzés módszerei

- Szakértővel folytatott közvetlen módszerek
  - Direkt módszerek:
    - interjú (esetgyűjtés, fókuszálás, fogalmak rendezése),
    - megfigyelés (pl. protokollelemzés),
  - Indirekt módszerek
    - különleges feladatok (fogalmak távolsága, osztályozása)
- Szakértőtől független, közvetett vagy automatikus módszerek (induktív tanuló módszerek, adatbányászat)

IR tervezésénél érintett tevékenységek

- Általános ontológia: Megfogalmazza, hogy mit tekintsünk önálló objektumnak, mit egy objektum tartalmának, hogyan ábrázoljuk az eseményeket, a folyamatokat, a mértékeket, mit kezdjünk az idő és a változás fogalmaival, hogyan kategorizáljuk mindezeket?
- Szakterületi ontológia: Az alkalmazási területre jellemző fogalmakat, kifejezéseket és a köztük fennálló kapcsolatokat írja le, amely egyben egy kommunikációs szövegkörnyezetet is létrehoz.

- Ismeretbázis felépítése az alkalmazási terület általános ismereteinek a szakterületi ontológiára támaszkodva.
- Szakterület-független problémamegoldó módszerek kiválasztása
- A szakterület ontológiájában szereplő fogalmaknak és az erre támaszkodó ismeretbázisnak az összerendelése a problémamegoldó módszerek be- és kimeneti igényeivel.

Ismeretalapú technológia jellegzetes ismérvei

1. Az IR valamilyen iteratív szoftverfejlesztési modell (inkrementális, spirál) alapján készül.
2. A rendszerben az ismeretbázis külön álló komponenst képez.
3. Működésére természetes nyelvi ember-gép párbeszéd jellemző.
4. Folyamatos tájékoztatást ad a végrehajtásról természetes nyelvi magyarázatadással.
5. Hagyományos és más MI technikákkal történő integrálás.

SZR technológia jellegzetes ismérvei

Az ismeretalapú technológia ismérvein túlmenően

6. A tárgyköri ismeretszerzés különleges módszerekre és eszközökre támaszkodik
7. A hagyományostól eltérő módszerek szükségesek az ismeretbázis, valamint a teljes szakértő rendszer elemzéséhez (validáció és verifikáció)
8. Piaci megjelentetése speciális marketing tevékenységet igényel

Történeti áttekintés

- Az IR fejlődése együtt halad az MI fejlődésével
  - Az első SZR-ek a hetvenes években (ez az MI 2. korszaka) jelennek meg, amikor megkezdődik a speciális ismeret reprezentációs és következtetési technikák kutatása
  - A SZR-ek révén kilép az MI a szoftverpiacra a nyolcvanas években (MI 3. korszak), előtérbe kerül a technológia (módszer és azt támogató eszköz) kidolgozásának igénye.
- Az IR-ek fejlesztésének problémái rokonok a szoftverfejlesztés általános problémáival
  - szabványosított módszertan hiánya a fejlesztésnél
  - megfelelő verifikációs és validációs módszerek hiánya

Klasszikus SZR
DENDRAL (DENDritic ALgorithm)

Fejlesztő: Stanford University
Idő: 1960-as évek közepe
Feladat: a NASA Mars program műholdja számára elemző támogatás ismeretlen szerves vegyületek molekula-szerkezetét határozza meg tömegspektográfiai és mágnesesrezonancia-mérési adatokból (pl. a $C_{20}H_{43}N$ molekula 43 millió lehetséges szerkezeti képletéből kiválasztja a jót.)

Klasszikus SZR
MYCIN

MYCIN / NEOMYCIN
Fejlesztő: Stanford University
Projekt indítás: 1972-74
Feladat: A vér bakteriális eredetű fertőzései és az agyhártyagyulladás betegség diagnosztizálásában, valamint a terápia meghatározásában segít.

EMYCIN (Empty MYCIN); oktatás: ***GUIDON***

Egy MYCIN szabály és hibái

IF 1. A fertőzés agyhártyagyulladás AND
2. Az agyhártyagyulladás bakteriális eredetű AND
3. Nincs még bakteriológiai lelet AND
4. A beteg legalább 17 éves AND
5. A beteg alkoholista
THEN Erősen valószínű (0.8), hogy az agyhártyagyulladást a 'diplococcus-pneumoniae' baktérium okozta.

Szakterületi ismeretek: 1. 2. 3.
Metaismeretek: 4. 5.

Keveredne a különféle szintű ismeretek

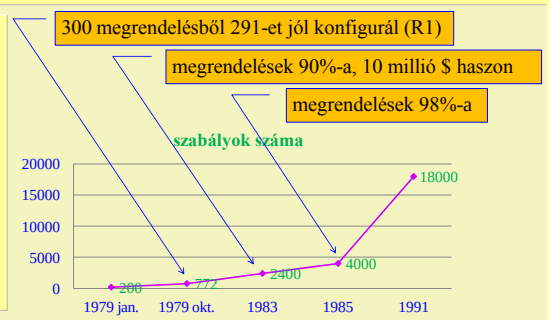
Klasszikus SZR PROSPECTOR

Fejlesztő: Stanford Research Institute
Idő: 1974-83
Feladat: A geológusokat segíti ásványlelőhelyek felkutatásában. (Alkalmazásával 1980-ban találtak egy 150 millió dolláros molibdén lelőhelyet ott, ahol a geológusok feladták a kutatást).

Klasszikus SZR R1/XCON (eXpert CONfigurer)

Fejlesztő: DEC +Carnegie-Mellon University
Projekt indítás: 1978
Feladat: VAX (később PDP11 is) számítógép-rendszer konfigurálását támogatja: egy megrendelés alapján megállapítja (anyagtakarékosan és megbízhatóan), hogy milyen komponensekkel kell kiegészíteni a vevő által adott tételeket.

XCON fejlődése



IR fejlesztését támogató eszközök megjelenése a nyolcvanas években

1. szakasz (1983-85): ism.alapú techn. kutatása

- általános célú (ún. horizontális) Lisp-alapú kertrendszerek (shell)
- 4 keretrendszer és „nagy" fejlesztő/értékesítő vállalat
  - **KEE** (Knowledge Engineer Env.) *Intellicorp*
  - **ART** (Autom. Reasoning Tool) *Inference Corp.*
  - **S.1** (System 1) *Teknowledge*
  - **KC** (Knowledge Craft) *Carnegie Group Inc.*

IR fejlesztését támogató eszközök megjelenése a nyolcvanas években

2. szakasz (1986-88): az új technológia bizonyít a piacon

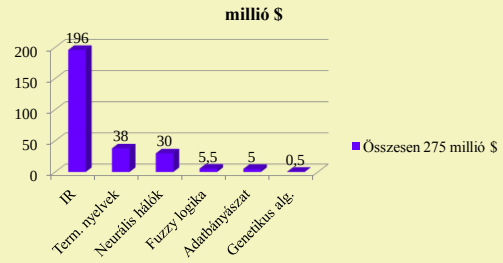
- technológia-vezérelt és prototípus-centrikus eszközök
- eszközhöz kerestek feladatokat
- horizontális eszközök közép-és kisgépekre
- PC-s eszközök megjelenése a piacon
- alkalmazási prototípusok készítése
- viszonteladó cégek jelennek meg
- kitör a shell-háború ("feature war")
- a piac telítődik, időleges befagy ("AI winter")

IR fejlesztését támogató eszközök megjelenése a nyolcvanas években

3. szakasz (1989-től): hagyományos és ismeretalapú technológia integrálása - középgépeken

- feladathoz keresnek "ahhoz illő" eszközt.
- alkalmazás-orientált, ún. vertikális eszközök
- technológia-váltás az eszközök terén:
  - szabályalapú (PC-s) eszközök mellett keretalapú eszközök is
- integrált alkalmazásokat kezdenek kidolgozni
- alkalmazásfejlesztési módszertan igénye (pl. KADS)
- éretlen az ismeretalapú technológia:
  - szabványosítás ismeretbázis-elemzés, V&V hiánya

Intelligens szoftverfejlesztő eszközök piaci forgalma Észak-Amerikában (1995)



Hazai SZR: UTI (Urinary Tract Infection)

Húgyúti fertőzések kezelése (demonstráció)

Fejlesztők: SZKI progr. tervezői, Péterfy S. Kórház orvosai
Implementálás: béta-változat Mprolog-ban1989-re
Cél: gyógyító erő + mellékhatások + gyógyszerár optimalizálása
Ismeretbázis: 10 diagnózis 21 baktérium
42 antibiotikum 250 szabály
A klinikai tesztelés eredménye:(100 beteg-adat)
(béta tesztelő: 3 kórház, eltérő spec. Igényekkel)
74 %-ban kiváló
20 %-ban jó
6 %-ban elfogadható

Hazai fejlesztésű ismeretalapú termékek

- **Prolog nyelvek:**
  - Mprolog (Modular Prolog) – *IQSOFT (SZKI)*
    - 1988-ig 25 országban, összesen 1500 installálás
  - CS-Prolog (Comm. Seq. Prolog) - *MultiLogic (ALL)*
    - 1996-ig 14 országban 150 installálás
- **Általános célú, horizontális eszközök:**
  - ALL EX PLUS: CS-Prolog alapú szimuláció - *ML*
  - GENESYS: szabályalapú keretrendszer - *SZÁMALK*
  - MProlog Shell - *IQSOFT*

Hazai fejlesztésű ismeretalapú termékek

- **Első vertikális eszközök (1990-ig):**
  - **CAPE:** számítógépes protokoll-elemzés (K*FKI*)
  - **KAS NES:** eset-vezérelt alkalmazások (*SZTAKI)*
  - **REALEX:** mérnöki valósidejű alkalmazások (*BME)*
  - **DINE:** intell. csoportos döntéstámogatás (*SZTAKI)*
  - **MetabolExpert:** kémiai/orv/biol. előjelz. (*CompuDrug)*
  - **PANGEA:** a mérnöki tervezés automatizálása (*BME)*
  - **ZEXPERT:** banki alkalmazások (*IQSOFT).*

Hazai IR projektek száma(1985-1991)

Orvos-egészségügy	16
Kémia	10
Számítástechnika	6
Közgazdaság, pénzügy	8
Építőipar	7
Energetika	7
Egyéb ipari terület	11