

## Razumljivost baz znanja kot dejavnik učinkovitosti ekspertnih sistemov za podporo odločanju

MOJCA VIZJAK PAVŠIČ  
JANEK MUSEK  
VLADISLAV RAJKOVIČ

### POVZETEK

V pričujoči raziskavi smo na osnovi empiričnega eksperimenta ugotavljali razlike v razumljivosti štirih reprezentacij znanja v okviru lupine ekspertnega sistema DEX (tabele izpeljanih pravil, odločitvena drevesa, trodimenzionalni grafi in uteži pomembnosti posameznih kriterijev) ter skušali pojasniti izvire variabilnosti rezultatov. Lupina ekspertnega sistema DEX, ki je bil razvit v Laboratoriju za umetno inteligenco na Institutu Jožef Stefan v Ljubljani, temelji na metodah kibernetike in umetne inteligence, namenjena pa je iskanju pravilnih odločitev, zlasti pri reševanju kompleksnih, večparametrskih problemov. Kot so pokazali rezultati multivariantnih analiz zbranih podatkov, obstajajo med obravnavanimi reprezentacijami znanja ekspertnega sistema DEX pomembne razlike v razumljivosti. Najvišje indekse razumljivosti ( $I_R$ ) dosegajo odločitvena drevesa, ki jim sledijo tabele izpeljanih pravil, medtem ko dosegajo trodimenzionalni grafi in uteži pomembnosti posameznih kriterijev precej nižje  $I_R$ . Prednosti odločitvenih dreves pred ostalimi obravnavanimi reprezentacijami znanja so v strukturi, ki omogoča hitro prepoznavanje vsebovanih hierarhij ter direktno, kognitivno ekonomično preiskovanje od korena drevesa do ustreznega lista, pri čemer ni potrebno pregledovanje celotnega prikaza podatkov, kar je še posebej pomembno pri bolj kompleksnih reprezentacijah. Kot je razvidno iz rezultatov študije, so odločitvena drevesa edina od obravnavanih reprezentacij znanja, katerih razumljivost bistveno ne upada v pogojih višjih stopenj kompleksnosti. Prav tako njihova razumljivost le malo variira glede na semantično vsebino konceptov kot tudi glede na tip vsebovanih informacij. Izrazito prednost odločitvenih dreves pred drugimi predstavitvami logičnih pravil sistema DEX so pokazale tudi ocene na 7-stopenjskih lestvicah za evaluacijo njihove razumljivosti, preglednosti in težavnosti.

### ABSTRACT

#### THE COMPREHENSION OF KNOWLEDGE BASES AS A FACTOR OF EFFECTIVENESS OF EXPERT SYSTEMS FOR DECISION SUPPORT

The present study reports on the results of the experiment to test the comprehensibility of four knowledge representation modes in the expert system shell DEX (tables of aggregate rules, decision trees, three-dimensional graphs, regression weights of individual criteria) and explains the sources of variability of the results. The shell of the DEX expert system, developed in the Artificial Intelligence

*Laboratory at the Jožef Stefan Institute in Ljubljana, is based on the methods of cybernetics and artificial intelligence and is intended for the search for correct decisions, above all in solving complex, multiparameter problems. As shown by the results of multivariate analyses of the collected data there are significant differences in comprehensibility among the studied knowledge representation modes of the DEX expert system. Decision trees, followed by aggregate rules have the highest comprehension indices ( $I_C$ ), while three-dimensional graphs and regression weights of individual criteria achieve considerably lower  $I_C$  values. An advantage of decision trees over other studied knowledge representation modes is in their structure, which enables quick recognition of the inherent hierarchies and direct, cognitively economical study, from the root of the decision tree to the corresponding leaf, whereby it is not necessary to review the complete presentation of data, which is especially important in more complex representations. As the results of this study showed, decision trees are the only representation mode studied in which comprehensibility does not decrease in conditions of higher levels of complexity. Their comprehensibility also varies very little with respect to the semantic content of concepts or the type of information contained. A distinctive advantage of decision trees over other presentation modes of logical rules of the DEX expert system was also demonstrated by scores on 7-point scales for the evaluation of their comprehensibility, clarity and difficulty.*

## UVOD

Eno izmed najpomembnejših vprašanj pri izdelavi računalniških ekspertnih sistemov je, kako v računalniku predstaviti znanje o sami problemski domeni, tako da je to znanje mogoče učinkovito uporabljati pri reševanju problemov, da ga je mogoče z lahkoto dopolnjevati in spreminjati, da je v obliki, ki je razumljiva strokovnjaku ekspertu za problemsko domeno in da ga je mogoče verificirati (Rajkovič, 1987; Gaines, 1989; Bratko, 1990; Rajkovič, Bohanec, 1991; Puppe, 1993). Kvaliteta ekspertnega sistema je v glavnem odvisna od obsega in kvalitete njegove baze znanja.

Ekspertni sistemi za podporo odločanju strokovnjaku omogočajo, da v dialogu z računalnikom sistematično izvede potrebne korake v procesu reševanja problema, pri čemer zastavlja računalnik vprašanja o vrednostih, prioritetah in pomembnosti posameznih parametrov, ki so relevantni za določeno problemsko področje (Rajkovič, 1987). Proces odločanja, ki ga običajno podpiramo z ekspertnimi sistemi, se praviloma opisuje kot kvalitativen, ker se ti programi opirajo bolj na človeško presojo kot na obsežne kvantitativne oziroma numerične analize.

Razumljivost, razlaga rezultatov in transparentnost sistemov so temeljni pojmi umetne inteligence (Michie, Bratko, 1986; Kodratoff, 1987; Bratko, Kodratoff, 1989). To še posebej velja za ekspertne sisteme, ki so že v zasnovi namenjeni reševanju problemov na t.i. "mehkih" problemskih področjih (Bohanec, 1990). To so področja, ki niso dosegla zadostne stopnje formalizacije, potrebne za natančne algoritmične rešitve. Problemi na teh področjih se rešujejo z uporabo strokovnega znanja, intuicije in izkušenj, pogosto pa potekajo tudi na osnovi nenatančnih in nezanesljivih podatkov. Nek sistem, ki rešuje ali pomaga reševati take probleme, mora biti zmožen prikazati znanje, ki ga je uporabil pri reševanju problema, in potek sklepanja, ki ga je pripeljalo do rešitve, v taki obliki, da ga uporabnik lahko razume, preveri, po potrebi spremeni ali posreduje drugim ljudem (Bohanec *et al.*, 1992; Vizjak Pavšič, 1994; Bohanec, Bratko, 1994).

Zato je poleg klasifikacijske točnosti eno najpomembnejših meril kvalitete zgra-

jenih odločitvenih pravil v ekspertnem sistemu tudi njihova razumljivost za človeka (Bergadano *et al.*, 1988). Razumljivost je pogosto tako pomembna, da se izmed več predstavitev nekega znanja v določenih primerih namenoma zavržejo predstavitve, ki imajo veliko klasifikacijsko točnost, vendar niso dovolj razumljive (Bohanec, 1990). Tipičen primer je "naivni" Bayesov klasifikator: na mnogih problemskih področjih presega klasifikacijsko točnost, doseženo z nekim drugim algoritmom avtomatskega učenja, vendar se kljub temu kasneje v praksi uporabijo rezultati avtomatskega učenja, ki so bolj razumljivi in s tem lažje preverljivi.

Lupina ekspertnega sistema DEX (Decision Expert), ki temelji na metodah kibernetike in umetne inteligence, je namenjena iskanju pravilnih odločitev, zlasti pri reševanju kompleksnih, večparametrskih problemov (Rajkovič, 1987; Bohanec *et al.*, 1988). Sistem DEX, ki je bil razvit v Laboratoriju za umetno inteligenco na Institutu Jožef Stefan v Ljubljani, omogoča ekspertu oziroma skupini ekspertov, da v sodelovanju z inženirjem znanja oblikujejo računalniško bazo znanja za odločanje. To bazo se nato uporabi za vrednotenje variant, za razlago vrednotenja in primerjalno analizo posameznih variant (Bohanec, Rajkovič, 1990; Bohanec *et al.*, 1995). Bazo odločitvenega znanja v sistemu DEX sestavlja ocenitveno znanje skupaj s kriteriji ocenjevanja, ki so organizirani v drevo. Poleg kriterijev vsebuje baza znanja tudi ekspertna pravila kombiniranja delnih ocen po kriterijih v skupne nadredne ocene vse do končne globalne ocene ustreznosti variant (Rajkovič, Bohanec, 1991; Rajkovič *et al.*, 1991). Sistem DEX je bil uporabljen v več kot 50 kompleksnih odločitvenih situacijah v okviru različnih industrijskih, vladnih in raziskovalnih ustanov v Sloveniji in v tujini.

Pomembna novost v okviru sistema DEX je možnost različnih predstavitev funkcije koristnosti, ki predstavljajo ključni del baze znanja, in sicer s tabelo osnovnih logičnih pravil, s tabelo izpeljanih pravil, z ekstrakcijo najbolj pomembnih pravil, z odločitvenim drevesom, grafično in z utežmi pomembnosti posameznih parametrov. Gre za prevode iz enega jezika v drugega, pri čemer se spreminja sintaksa, vsebina pa ostaja nespremenjena. Izbor prave, to je vsebinsko ustrezne in razumljive predstavitve pa je odvisen od problema ter od (pred)znanja in izkušenj uporabnikov - odločevalcev, ki sodelujejo v procesu izgradnje baze znanja oz. njene verifikacije. Pa tudi potem, ko se izbere določen jezik predstavitve, je mogoče nek koncept, se pravi neko funkcijo koristnosti, predstaviti na več načinov. In dalje: celo enake predstavitve je mogoče prikazati na različno razumljive načine. Razumljivost nekega odločitvenega drevesa je lahko na primer večja, če se prikaže grafično, kot pa če se prikaže v linearizirani obliki.

Proces izgradnje baze znanja vključuje identifikacijo problema, identifikacijo virov znanja, konceptualizacijo, formalizacijo in implementacijo. Ta kompleksen proces poteka praviloma iterativno, pri čemer se na vsakem koraku preverja veljavnost, popolnost in konsistentnost dejstev, pravil in konceptov, vnešenih v sistem. Proces izgradnje baze znanja zahteva tesno sodelovanje med inženirjem znanja in ekspertom za določeno problemsko področje (Hawkins, 1983; Evans, 1988; Waern, 1989). Interakcija med predstavniki dveh popolnoma različnih strok, med katerimi naj bi bila komunikacija obojestransko razumljiva tudi na globljih ravneh analize problema, pogosto povzroča velike težave pri vnašanju znanja v sistem (Neale, 1988; Giarratano, Riley, 1994). Verifikacija in validacija znanja, vnešenega v sistem, je odvisna med drugim od tega, kako dobro ekspert razume in interpretira uporabljene reprezentacijske sheme. Kot so pokazale izkušnje, lahko dobro načrtovane oblike reprezentacije znanja izboljšajo učinkovitost transfera informacij, s tem da vplivajo na motivacijo eksperta, da so jasne in kognitivno ekonomične in da stimulirajo analitično razmišljanje. Vendar je dejanska uporabnost in učinkovitost različnih načinov reprezentacije znanja, ki se uporabljajo v okviru ekspertnih sistemov, kot opozarja več avtorjev (Bergadano *et al.*, 1988; Monta-

zemi, 1991; Subramanian *et al.*, 1992; Lalomia *et al.*, 1992), relativno slabo raziskana.

V okviru pričujoče raziskave smo skušali na podlagi empiričnega eksperimenta identificirati tiste predstavitve znanja ekspertnega sistema DEX, ki jih ob zahtevani točnosti predstavitev človek najlaže razbere in razume, ter pojasniti izvore variabilnosti rezultatov. S stališča kognitivne psihologije lahko obravnavamo razumevanje posameznih reprezentacij znanja kot kompleksen proces reševanja problemov, na podlagi katerega sklepa recipient o odnosih med elementi, ki jih vsebuje predstavitev (Baron, 1988; Fletcher, Bloom, 1988; Musek, 1990, 1993; Trstenjak, 1980, 1981; van den Broek, 1990). Teorija človekovega sistema predelovanja informacij predstavlja trdne teoretične osnove za nadaljnje raziskave človekovih miselnih procesov. Newell in Simon (1972) opredeljujeta tri glavne dimenzije, vzdolž katerih variira proces reševanja problemov, to so: dimenzija naloge, dimenzija individualnih razlik in dimenzija delovanje - učenje - razvoj.

Naloga (problem, besedilo, grafična predstavitev podatkov) lahko variira glede na tri širše kategorije spremljivk: glede na vsebino (abstraktna vs. konkretna; dvoumna vs. enopomenska; recipientu poznana vs. nepoznana) oz. tip vsebovanih informacij (npr. numerični vs. verbalni), glede na kompleksnost (na področju odločanja se kompleksnost praviloma opredeljuje na podlagi števila atributov in števila alternativ) ter glede na zahtevane kognitivne operacije (strategije branja oz. reševanja problema), ki so lahko bolj ali manj naučene, bolj ali manj avtomatizirane in v skladu s tem nalagajo delovnemu spominu večje oz. manjše obremenitve (Kintsch, van Dijk, 1978; van Dijk, Kintsch, 1983; Kintsch, 1988). Kot so pokazale empirične raziskave (Daneman, Carpenter, 1980; Musek, 1990; Baddeley, 1992), obstajajo velike individualne razlike v razumevanju in hitrosti branja ter reševanja problemov, ki korelirajo z individualnimi razlikami v percepciji, pomnjenju, inteligentnosti in v drugih kognitivnih procesih.

Pričujoča raziskava se je osredotočila na prvi dve dimenziji, to sta: dimenzija naloge in dimenzija individualnih razlik. Na podlagi empiričnega eksperimenta smo ugotavljali, kako razumevanje različnih, vendar informacijsko ekvivalentnih reprezentacij znanja ekspertnega sistema DEX variira kot funkcija dveh značilnosti problema, tj. kompleksnosti in semantične vsebine, oz. kot funkcija nekaterih osebnih karakteristik, predvsem tistih, ki jih je v praktični odločitveni situaciji možno opazovati, medtem ko variacije glede na bolj elementarne procese predelovanja informacij niso bile neposredno predmet tega eksperimenta. Različne oblike reprezentacij znanja smo proučili tudi z lestvicami razumljivosti, preglednosti in težavnosti.

## METODA

### *Eksperimentalni načrt*

Namen pričujoče raziskave je bil ugotoviti razlike v razumljivosti štirih reprezentacij znanja ekspertnega sistema DEX in pojasniti izvore variabilnosti rezultatov. V eksperimentalnem delu smo variirali način reprezentacije znanja (Tabela 1). Deset konceptov iz petih problemskih področij smo predstavili na štiri reprezentacijske načine, izmed katerih zahteva vsak obvladovanje drugačne strategije branja, to so: tabele izpeljanih pravil, odločitvena drevesa, trodimenzionalni grafi in uteži pomembnosti posameznih kriterijev (Slika 1). Funkcije koristnosti je v okviru sistema DEX namreč možno predstaviti na več načinov, pri čemer gre za prevode iste vsebine iz enega jezika v drugega. Iz različnih predstavitev mora odločevalec razbrati isto vsebino. Prve tri od naštetih reprezentacij so, v kolikor predstavljajo isto funkcijo koristnosti, informacijsko ekvivalentne, medtem ko vsebujejo uteži bistveno manj podrobnosti.

POGOJ	Kompleksnost	
	1	2
	Prob1..Prob5	Prob1..Prob5
Repr1: izpeljana pravila	65	65
Repr2: odločitvena drevesa	73	73
Repr3: trodimenzionalni grafi	68	68
Repr4: uteži	60	60

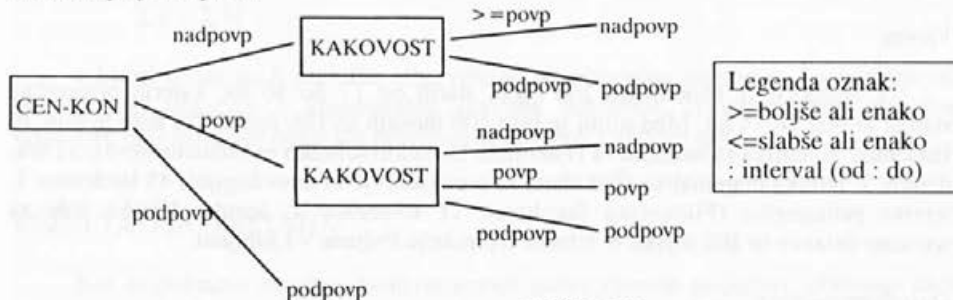
Op.: Način reprezentacije (Repr1..Repr4) je spremenljivka, ki variira med štirimi glede na izobrazbo izenačenimi skupinami poskusnih oseb. Problem (P1..P5) in stopnja kompleksnosti (1 -nižja, 2 - višja) pa sta spremenljivki, ki variirata znotraj skupin. V celicah je navedeno število poskusnih oseb v vsakem eksperimentalnem pogoju.

Tabela 1. Eksperimentalni načrt

CEN. KONKUR.	KAKOVOST PROIZ.	TRŽNE ZNAČ. PROIZ.
1. nadpovp	≥ povp	nadpovp
2. ≥ povp	nadpovp	nadpovp
3. povp	povp :	povp
4. podpovp	‡	podpovp
5. ‡	podpovp	podpovp

Legenda oznak:  
‡ poljubna vrednost  
≥ boljše ali enako  
≤ slabše ali enako  
: interval (od : do)

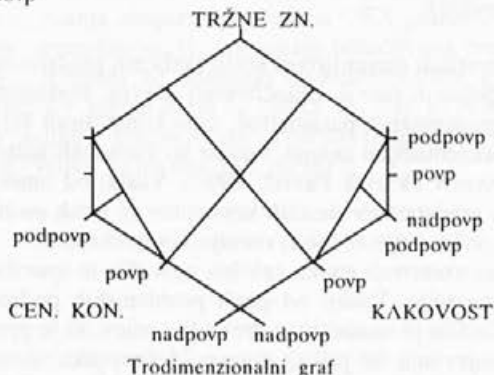
Tabela izpeljanih pravil



CENOVNA  
KONKURENČNOST 50%

KAKOVOST  
PROIZVODA 50%

Uteži pomembnosti  
posameznih kriterijev



Slika 1. Pravila za oceno tržnih značilnosti proizvoda, prikazana s tabelo izpeljanih pravil, z odločitvenim drevesom, s trodimenzionalnim grafom in z utežmi pomembnosti posameznih kriterijev.



Prvi od petih v eksperimentalne namene uporabljenih problemov (tj. nakup avtomobila) je demonstracijski, medtem ko ostali štirje (to so: problem vrednotenja proizvodnih programov s portfolio analizo, problem ocenjevanja športne nadarjenosti otrok, problem razdeljevanja posojil Stanovanjskega sklada Republike Slovenije in problem izbiranja kandidata za vodstveno delovno mesto) izhajajo iz realnih odločitvenih situacij, v katerih je bil uporabljen ekspertni sistem DEX. V okviru vsakega sta bila predstavljena dva koncepta, ki sta variirala v stopnji kompleksnosti. Prvi, manj kompleksen je bil opredeljen z dvema atributoma, medtem ko je bil bolj kompleksen opredeljen s tremi atributi. Višjih stopenj kompleksnosti nismo proučevali, ker jih baze znanja sistema DEX ne vsebujejo. Kot je pokazalo teoretično in aplikativno delo (Bohanec *et al.*, 1988), je namreč metoda večparametrskega odločanja, podprta z lupino ekspertnega sistema DEX izvedljiva ob upoštevanju nekaterih omejitev; med drugimi te, da število parametrov, ki jih ekspert kombinira, ne sme biti večje od štiri, ker človek ni sposoben konsistentno razmišljati o povezavi več postavk. Reševanje realnih odločitvenih problemov pa je pokazalo, da daje najboljše rezultate delo z dvema ali tremi parametri. To na videz ostro omejitev v procesu izgradnje baze znanja rešujemo s strukturiranjem parametrov v ustrezno majhne skupine.

Razumevanje in hitrost branja sta v veliki meri odvisna tudi od individualnih značilnosti recipienta, kot so znanje, motivacija in kognitivne sposobnosti. V okviru pričujoče študije smo variacije v rezultatih proučevali predvsem z vidika tistih karakteristik, ki jih je v praktični odločitveni situaciji možno opazovati, tj. predhodno znanje, izkušnje, pogostost uporabe računalnika v različne namene, stališča do računalniške tehnologije ipd.

#### *Vzorec*

V raziskavi je sodelovalo 266 oseb, starih od 17 do 30 let, katerih povprečna starost je bila 18,94 let. Med njimi je bilo 100 moških in 166 žensk. Od tega je bilo 76 študentov 1. letnika računalništva (Fakulteta za elektrotehniko in računalništvo), 32 študentov 2. letnika matematike (Fakulteta za naravoslovje in tehnologijo), 45 študentov 1. letnika pedagogike (Filozofska fakulteta), 11 študentov 1. letnika Visoke šole za socialne delavce in 102 dijaka 4. letnika Gimnazije Poljane v Ljubljani.

#### *Opis instrumentov*

Da bi testirali razumljivost štirih različnih predstavitev znanja ekspertnega sistema DEX, tj. izpeljanih pravil, odločitvenih dreves, trodimenzionalnih grafov in uteži pomembnosti posameznih parametrov, smo konstruirali štiri različne teste, ki so se razlikovali v reprezentacijah znanja, vendar so vsebovali natančno isti vprašalnik z multiplo izbiro odgovorov (Vizjak Pavšič, 1994). Vsako od omenjenih štirih reprezentacij smo uporabili za predstavitev desetih konceptov iz petih problemskih področij z nižjo (dva atributa) oz. višjo (trije atributi) ravnijo kompleksnosti.

Testi so vsebovali enaka splošna navodila in specifična navodila, glede na vsebovane reprezentacije. Vsako od petih problemskih področij smo predstavili z opisom problema. Sledilo je semantično drevo kriterijev, ki je predstavljalo v opisu opredeljene kriterije ocenjevanja ter prikaz domen (definijska območja) kriterijev. Na tej osnovi so namreč v računalniški bazi znanja opredeljena pravila (funkcije koristnosti) za združevanje delnih ocen po posameznih kriterijih v skupno združeno oceno. V okviru vsakega problema sta bila v testu predstavljena dva koncepta, od katerih je bil prvi manj in drugi bolj kompleksen. Manj kompleksni reprezentaciji je sledilo 5 vprašanj z multi-

plo izbiro odgovorov, bolj kompleksni pa 10 vprašanj z multiplo izbiro odgovorov.

Vsi testi so uvodoma, poleg navodil za reševanje, vsebovali vprašalnik o socio-demografskih značilnostih preskušanca (starost, spol, fakulteta oz. gimnazija, smer in letnik študija, učni uspeh, izobrazba staršev), o izkušnjah in pogostosti dela z računalnikom ter o stališčih do računalniške tehnologije; in na koncu 7-stopenjsko lestvico za evaluacijo razumljivosti, preglednosti in težavnosti vsebovane reprezentacije ter vprašanja, ki so se nanašala na težavnost obravnavanih problemov in načine njihove predstavitve.

### Postopek

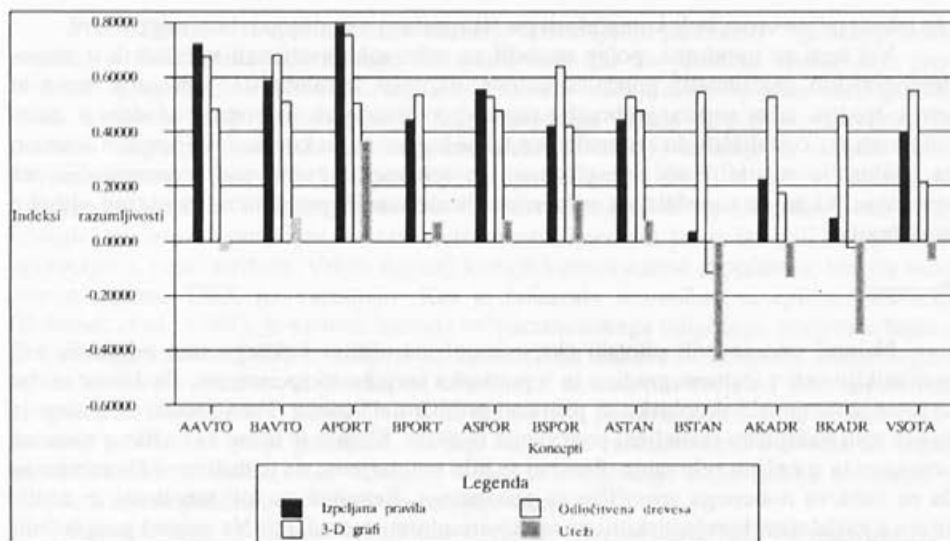
Najprej smo izvedli pilotski eksperiment, na osnovi katerega smo odpravili več pomanjkljivosti v testnem gradivu in v postopku izvedbe eksperimenta. Poskusne osebe so teste izpolnjevale skupinsko in pismeno približno 45 minut. Pred začetkom testiranja so bili testi naključno razdeljeni poskusnim osebam. Sledilo je ustno navodilo o namenu raziskave in o načinu reševanja. Posebej je bilo poudarjeno, da je testiranje anonimno in da ne zahteva nobenega specifičnega predznanja. Rezultati so bili tabelirani in analizirani z različnimi korelacijskimi in multivariantnimi metodami. Na osnovi povprečnih točkovnih vrednosti, doseženih glede na obravnavanih deset konceptov in glede na vsoto vseh pravih odgovorov v testu, smo za vse obravnavane reprezentacije znanja sistema DEX izračunali tudi indekse razumljivosti ( $I_R$ ) na podlagi naslednje enačbe:

$$I_R = \frac{\sum P - \sum N}{\sum P + \sum N}$$

v kateri označuje P pravilne odgovore in N nepravilne odgovore.  $I_R$  variirajo od [1, -1], kjer označuje 1 konsistentno pravilne odgovore, -1 konsistentno napačne odgovore in 0 označuje polovico pravih odgovorov v okviru testa.

### REZULTATI IN DISKUSIJA

Kot so pokazali rezultati multivariantnih analiz zbranih podatkov, obstajajo med obravnavanimi reprezentacijami znanja ekspertnega sistema DEX pomembne razlike v razumljivosti. Najvišje indekse razumljivosti ( $I_R$ ) dosegajo odločitvena drevesa, ki jim sledijo tabele izpeljanih pravil, medtem ko dosegajo trodimenzionalni grafi in uteži pomembnosti posameznih kriterijev precej nižje  $I_R$  (Slika 2). Kot je pokazala multivariantna analiza variance, so razlike v razumljivosti obravnavanih reprezentacij znanja statistično visoko signifikantne (Wilks' Lambda=.355,  $p < .000$ ).



Slika 2. Grafični prikaz indeksov razumljivosti reprezentacij znanja ekspertnega sistema DEX glede na različna problemska področja (AVTO - problem nakupa avtomobila, PORT - problem vrednotenja proizvodnih programov s portfolio analizo, SPOR - problem ocenjevanja športne nadarjenosti otrok, STAN - problem razdeljevanja posojil Stanovanjskega sklada Republike Slovenije in KADR - problem izbiranja kandidata za vodstveno delovno mesto) in glede na različne stopnje kompleksnosti (A - nižja, B - višja).

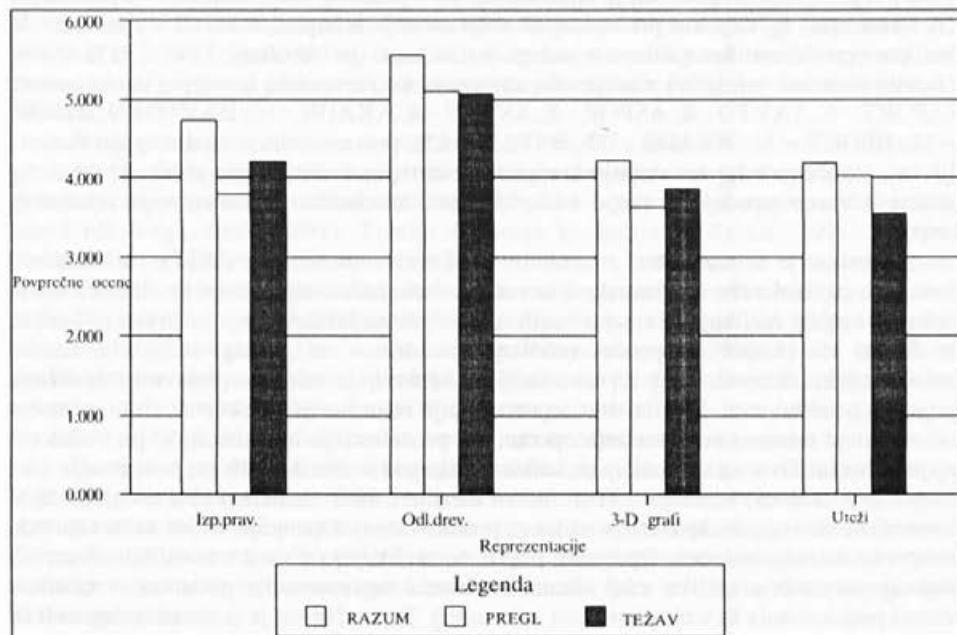
Rezultati pričujoče raziskave so konsistentni z rezultati eksperimentov Noseka in Rotha (1990) ter Subramaniana *et al.* (1992), ki so pokazali, da so odločitvena drevesa učinkovitejša osnova za interpretacijo logičnih pravil kot tabele, medtem ko drugih predstavitev niso proučevali, kakor tudi ne razumljivosti reprezentacij, v kolikor varirajo glede na semantično vsebino, glede na tip vsebovanih informacij ter glede na stopnjo kompleksnosti.

Prednosti *odločitvenih dreves* pred ostalimi obravnavanimi reprezentacijami znanja ( $I_R$  izračunani glede na vsote vseh pravih odgovorov v testu znašajo za drevesa 0.55, za tabele izpeljanih pravil 0.32, za trodimenzionalne grafe 0.26 in za uteži pomembnosti posameznih kriterijev -0.07) so v strukturi, ki omogoča hitro prepoznavanje vsebovanih hierarhij ter direktno, kognitivno ekonomično preiskovanje od korena drevesa do ustreznega lista, pri čemer ni potrebno pregledovanje celotnega prikaza podatkov, kar je še posebej pomembno pri bolj kompleksnih reprezentacijah. Kot je razvidno iz grafičnega prikaza  $I_R$ , so odločitvena drevesa edina od obravnavanih reprezentacij znanja, katerih razumljivost bistveno ne upada v pogojih višjih stopenj kompleksnosti. Prav tako njihova razumljivost le malo variira glede na semantično vsebino konceptov kot tudi glede na tip vsebovanih informacij.

Izrazito prednost odločitvenih dreves pred drugimi predstavitvami logičnih pravil sistema DEX so pokazale tudi ocene na 7-stopenjskih lestvicah za evaluacijo njihove razumljivosti, preglednosti in težavnosti (Slika 3). Kot najbolj razumljiva, pregledna in najmanj težavna so respondenti ocenili odločitvena drevesa, ki so dobila najvišjo oceno na lestvici razumljivosti (5.56), le nekoliko nižjo pa na lestvicah preglednosti (5.09) in



težavnosti (5.06). Odločitvenim drevesom sledijo izpeljana pravila, katerih povprečna ocena na lestvici razumljivosti je 4.74, na lestvici težavnosti pa 4.21, medtem ko so jih respondenti precej slabše ocenili na lestvici preglednosti (3.79). Pomembno nižje so na vseh treh lestvicah respondenti ocenili trodimenzionalne grafe in uteži pomembnosti posameznih kriterijev. Trodimenzionalne grafe so ocenili kot precej nerazumljive (4.22) in težke (3.85), na lestvici preglednosti pa so dobili najnižjo oceno izmed vseh štirih reprezentacij (3.63). Uteži pomembnosti posameznih kriterijev so respondenti ocenili kot najmanj razumljive (4.18) in najtežje (3.55) od vseh Dexovih reprezentacij, medtem ko so na lestvici preglednosti dobile precej visoko oceno (4.02), ki jih v tem pogledu uvršča takoj za odločitvena drevesa. Kot je pokazala multivariantna analiza variance, so obravnavane razlike statistično visoko signifikantne (Wilks' Lambda=.833,  $p < .000$ ).



Slika 3. Grafični prikaz povprečnih ocen posameznih reprezentacij sistema DEX na 7-stopenjskih lestvicah razumljivosti, preglednosti in težavnosti.

Kot je nakazala raziskava Vessey in Webra (1986), izhajajo izrazite prednosti odločitvenih dreves iz njihove strukture, ki kombinira značilnosti sekvenčnih reprezentacij, ki jih tvorijo zaporedja informacijskih elementov in značilnosti diagramskih reprezentacij, v katerih je informacija predstavljena prostorsko ter s tem podpirajo oboje, tako serijske kot tudi paralelne strategije procesiranja informacij. Čeprav potekata oba tipa predelovanja informacij hkrati in povezano med reševanjem različnih nalog, so rezultati empiričnih raziskav pokazali, da sekvenčne reprezentacije pospešujejo predvsem serijsko, analitično procesiranje in pomnjenje komponent v zaporedju, medtem ko diagramske bolj paralelno, sintetično obdelavo informacij ter pomnjenje vzorca, ki ga posamezne komponente formirajo. Rezultati raziskav so nadalje pokazali, da dodajanje podrobnosti pospeši serijsko procesiranje, medtem ko odstranitev podrobnosti vzpodbudi paralelno procesiranje, ki je potrebno, če naloga zahteva učenje

vzorca (Winn, 1989). Z vidika praktične odločitvene situacije, v kateri je treba logična pravila, vnešena v bazo znanja ekspertnega sistema, dopolnjevati in verificirati, to pomeni, da omogočajo odločitvena drevesa bodisi s perceptivnim poudarjanjem vzorca, ki ga tvorijo komponente, bodisi s poudarjanjem posameznih vej, ki potekajo od notranjih vozlov do listov (s pomočjo barv ali zgolj z usmerjanjem pozornosti), hitro in enostavno spreminjanje strategij predelovanja informacij.

Odločitvenim drevesom sledijo glede na dosežene  $I_R$  tabele izpeljanih pravil, ki so jih respondenti ocenili na 7-stopenjskih lestvicah kot precej razumljive in nezahtevne, vendar kot precej nepregledne. Kot je razvidno iz grafičnega prikaza  $I_R$ , razumljivost tabel izpeljanih pravil izrazito upada s stopnjo kompleksnosti reprezentacije, odvisna pa je tudi od semantične vsebine koncepta ter tipa vsebovanih informacij. Za vseh pet problemskih področij je značilno, da so v primerih nižje stopnje kompleksnosti (A - koncepti)  $I_R$  višji kot pri konceptih višje stopnje kompleksnosti (B - koncepti). V kolikor opredelimo kompleksnost naloge natančneje (po Abelson, Levi, 1985) kot  $m$  (število alternativnih izbir)  $\times n$  (število atributov) ter razvrstimo koncepte na tej osnovi (APORT - 6, AAVTO - 8, ASPOR - 8, ASTAN - 8, AKADR - 10; BAVTO - 9, BSPOR - 12, BPORT - 15, BKADR - 15, BSTAN - 15), postane odnos med stopnjo razumljivosti, izraženo v  $I_R$ , ter stopnjo kompleksnosti reprezentacije zelo jasno razviden:  $I_R$  skoraj linearno upadajo z rastjo kompleksnosti tabelarične reprezentacije izpeljanih pravil.

Rezultat je konsistenten z ugotovitvijo Larkina in Simona (1987), da se informacijsko ekvivalentne diagramske (drevesa) in sekvenčne reprezentacije (tabele) medsebojno močno razlikujejo v zahtevanih operacijah za predelovanje informacij. Larkin in Simon sta proučevala proces reševanja problemov na podlagi različnih, vendar informacijsko ekvivalentnih reprezentacij in ugotavljala relativno težavnost različnih pogojev predstavitve. Ko sta dve reprezentaciji informacijsko ekvivalentni, je njuna učinkovitost odvisna od potrebnih operacij za predelovanje informacij, ki jih vsaka od njiju zahteva. Dva seta operacij se lahko razlikujeta v zmožnostih prepoznavanja (re-kognicije) vzorcev, v sklepih, ki jih lahko direktno, hitro in nenaporno izvedeta, in v kontrolnih strategijah (še zlasti v nadzoru preiskovanja). Operacije, ki jih zahtevajo odločitvena drevesa in tabele izpeljanih pravil, se razlikujejo v vseh teh vidikih. Torej so najbolj pomembne razlike med obema načinoma reprezentacije podatkov v učinkovitosti preiskovanja in v eksplicitnosti informacij. Ta značilnost je jasno razvidna tudi iz ocen respondentov: medtem ko so drevesa ocenili kot razumljiva, pregledna in nezahtevna, so tabele izpeljanih pravil ocenili kot razumljive in nezahtevne, vendar slabo pregledne.

Kot je nadalje pokazala analiza rezultatov, je razumljivost tabel izpeljanih pravil v veliki meri odvisna tudi od semantične vsebine problemskega področja (poznana vs. neznan) ter od tipa informacij (abstraktni vs. konkretni izrazi, kratice vs. neokrajšane verbalne oznake, dvoumni vs. enopomenski izrazi, verbalne vs. numerične oznake), ki jih vsebuje koncept. Rezultati so skladni z izsledki raziskav, po katerih ljudje logična pravila lažje razumejo, v kolikor ima relevantno gradivo pomenljivo vsebino. Eksperimenti, v katerih so uporabili različne vrste silogizmov, pogojne stavke in naloge s kartami, so pokazali, da ljudje v vseh primerih mislijo bolj "logično", v kolikor je vsebina konkretna in poznana (Johnson-Laird, 1983; Ormerod *et al.*, 1986; Waern, 1989). Te izsledke je možno razložiti na dva načina. Po prvi razlagi ljudem ni treba misliti logično, ker preprosto uporabijo svoje že vskladiščeno znanje za razumevanje trditev. Pri tem konstruirajo modele realnosti ter elaborirajo relacije v teh modelih in ne razmišljajo na podlagi logičnih operacij (Johnson-Laird, 1983).

Po drugi razlagi teh izsledkov pa ima poznana vsebina svoj pomen in jo je zato

laže obdržati v delovnem spominu. Količina informacije, ki jo lahko človek obdrži v delovnem spominu, je omejena. Pomenljivo informacijo je lažje povezati v informacijske enote, zaradi česar jo je možno obdržati v spominu v večjih količinah in za daljša časovna obdobja kot abstraktno informacijo. Znan je npr. Katonov eksperiment pomnjenja zaporedja števil 581215192226. Skupina, ki ji je razložil notranjo strukturo števila (serija, ki se začne s 5, 8, 12 itd.), si je število zapomnila bistveno boljše. Soroden primer navaja Hirst (1988), ki je uporabil kot eksperimentalno gradivo črke: medtem ko si ljudje zelo težko zapomnijo zaporedje črk, kot je npr. AREPOSNAEMTEMEHT, pa težav večinoma nimajo več, če ga transformirajo v zaporedje THE MET MEANS OPERA, obrnjeno črkovanje prvotnega niza, razčlenjeno v pomenljiv stavek. Tako pomaga pomenljiva vsebina ljudem procesirati logične trditve bolj logično. Večina raziskav logičnega mišljenja je tudi pokazala, da ljudje težje procesirajo negacije kot trdilne stavke. Tudi ta izsledek je možno deloma razložiti kot problem mentalne reprezentacije: kako naj si predstavljamo nekaj, česar NI?, deloma pa z vidika zahtevanih operacij, saj zastavljajo negacije pri procesiranju jezika več operacijskih problemov kot trdilni stavki (Waern, 1989).

Procesiranje in pomnjenje abstraktnih v primerjavi s konkretnimi besedami je bilo predmet številnih empiričnih študij tudi v okviru raziskav mentalnih predstav, ki so pokazale, da je pomnjenje konkretnih besed veliko boljše kot pomnjenje abstraktnih besed (de Vega, Diaz, 1991). Teorija dvojnega kodiranja A. Paivia (1971) predpostavlja, da so konkretne besede običajno kodirane v verbalnem in imaginacijskem sistemu, medtem ko so abstraktno besede kodirane samo verbalno. Pomnjenje je namreč najboljše, v kolikor se uporabita obe vrsti kodov. To pojmovanje potrjujejo tudi eksperimenti G. Zhanga in H. A. Simona (po Johnson-Laird, 1988), ki sta kitajskim poskusnim osebam prezentirala sete dveh različnih vrst kitajskih pismenk, od katerih so nekatere imele imena, medtem ko jih druga vrsta ni imela. Osebe so si lahko zapomnile okoli šest pismenk, ki so imele imena, vendar niti treh pismenk brez imen.

Omenjene značilnosti človekovega predelovanja informacij se odražajo tudi v izsledkih pričujoče raziskave. Pri konceptih, katerih vsebina je bila respondentom dobro poznana in so vsebovali poznano terminologijo, so dosegali višje indekse razumljivosti. Poleg tega je jasno razviden učinek interakcije med variablama kompleksnosti reprezentacije in pomenljivosti vsebovanih izrazov, katerih vpliv se odraža v obremenjenosti delovnega spomina.

Medtem ko je večina respondentov (115) ocenila problem nakupa avtomobila (AVTO) kot najlažji, na drugo mesto pa se je po frekvencah odgovorov (37) uvrstil problem ocenjevanja športne nadarjenosti otrok (SPOR), je večina respondentov (112) ocenila kot najtežji problem razdeljevanja posojil Stanovanjskega sklada Republike Slovenije (STAN), na drugo mesto po težavnosti pa se je glede na frekvence odgovorov (41) uvrstil problem vrednotenja proizvodnih programov s portfolio analizo (PORT). Problem izbire kandidata za vodstveno delovno mesto (KADR) se po težavnosti uvršča med obe navedeni kategoriji problemov.

Zakaj? Probleme, ki so jih respondenti označili kot lažje, so reševali lažje, ker - kot so navajali - gre za aktualno, življenjsko temo, ker vsebino problema poznajo in jih zanima; in dalje, ker so podatki predstavljeni razumljivo in pregledno, ker predstavitev ni preveč kompleksna in obsežna. Nekateri respondenti pa so navedli, da jim je olajšala reševanje numerična opredelitev domen kriterijev.

Probleme, ki so jih respondenti označili kot težje, pa so reševali težje, ker - kot so navajali - gre za nepoznano vsebino problema, zaradi zapletenosti (kompliciranosti) izražanja, zaradi nepoznatih izrazov (pojmov), zaradi okrajšav (kratic); zaradi različno (neenotno) definiranih domen kriterijev, ker si je težko zapomniti definicijska območja

kriterijev; in dalje, ker predstavitve niso znali brati in zaradi izredne zgoščenosti podatkov. Le manjše število respondentov pa je kot razlog navedlo pomanjkanje motivacije za reševanje testa.

Kot je razvidno iz grafične predstavitve podatkov, dosežejo tabele izpeljanih pravil najnižjo stopnjo razumljivosti prav pri konceptu BSTAN, za katerega je značilna visoka stopnja kompleksnosti in s stališča respondentov nepoznana vsebina. Po drugi strani pa so za oba problema, ki so ju respondenti označili kot lahka, značilni visoki  $I_R$  tako v pogojih nizke (AAVTO, ASPOR) kot tudi v pogojih višje (BAVTO, BSPOR) stopnje kompleksnosti predstavitve. Ta značilnost, namreč majhna razlika v  $I_R$  med manj in bolj kompleksno predstavitvijo, je razvidna tudi pri problemu izbire kandidata za vodstveno delovno mesto (KADR), ki se uvršča med sematično lažje probleme. Za ta problem, ki je v testu zavzemal zadnje mesto, so sicer značilni relativno nizki  $I_R$ , vendar na podlagi porazdelitve frekvenc manjkajočih odgovorov lahko sklepamo, da ne zaradi težavnosti obeh predstavitev, temveč predvsem zaradi obsežnosti testa. Tisti respondenti, ki so test reševali počasneje, tega zadnjega dela niso uspeli rešiti, ali pa so, da bi pohiteli, odgovarjali površno in napačno glede na to, da je bil čas reševanja omejen.

Nasprotno pa je za oba problema, ki so ju respondenti označili kot težja, tj. problem razdeljevanja posojil Stanovanjskega sklada Republike Slovenije in problem vrednotenja proizvodnih programov s portfolio analizo, značilna izrazita razlika med  $I_R$ , v pogojih nižje stopnje kompleksnosti in v pogojih višje stopnje kompleksnosti, se pravi med  $I_R$  pri APORT in BPORT ter med  $I_R$  pri ASTAN in BSTAN. Torej je učinek interakcije med obema variablama, ki vplivata na obremenjenost delovnega spomina, tj. kompleksnost reprezentacije in pomenljivost vsebovanih izrazov, iz prezentiranih rezultatov jasno razviden. Da bi bralec lahko sklepal, mora obdržati ustrezno sled v spominu. Kot poudarja A. Trstenjak (1980, 1981), je hkratnost bistvena lastnost vsakega pravega spoznavnega procesa. Kjer ni te hkratnosti, ni razumevanja. V pogojih nižje stopnje kompleksnosti, ki delovni spomin manj obremenjuje, človek lahko obdrži v kratkoročnem spominu tudi manj razumljive in nejasne referente, medtem ko mu v pogojih višjih stopenj kompleksnosti kapaciteta kratkoročnega spomina več ne dopušča obdržati v spominu vseh ustreznih referentov, v kolikor so nerazumljivi, nejasni ali dvoumni.

Natančnejši vpogled v značilnosti problema SPOR nadalje pokaže, da so domene enotno opredeljene pri vseh kriterijih ter da so izrazi, s katerimi so označene možne vrednosti, razumljivi, nedvoumni in jasno medsebojno razločljivi: neprimeren, sprejemljiv, dober, odličen. Tudi kratice, ki jih vsebujejo reprezentacije logičnih pravil, so nedvoumne: nepr, spr, dober, odl. Na podoben, kognitivno ekonomičen način so opredeljene tudi domene kriterijev pri problemu KADR: nezadovoljiv, zadovoljiv, dober, prav dober, odličen oz. okrajšano: nezad, zad, dob, p.dob, odl.

Nasprotno, pa je za problem STAN značilna kognitivno obremenjujoča opredelitev domen kriterijev, ki so definirane neenotno, uporabljeni izrazi pa so precej nejasni in medsebojno premalo razločljivi, npr. pri kriteriju "stanovanje": (1) ne-pred, (2) pred, (3) bolj-pred, (4) naj-pred; ali pri kriteriju "socialno in zdravstveno stanje družine": (1) ustr, (2) pred, (3) bolj-pred, (4) naj-pred. Eden od respondentov je obravnavani problem povsem točno zaznal: "Nekateri pojmi so težje razumljivi in veliko je podobnih besed, ki jih težko razvrščam." Medtem ko so vrednosti pri problemu SPOR opredeljene s štirimi fonološko popolnoma različnimi izrazi, se v primeru STAN pri kriteriju "stanovanje" izraz "pred" ponovi v vsaki od možnih vrednosti, kar lahko povzroči učinek akustične podobnosti (Baddeley, 1992, 1994). Raziskave so namreč pokazale, da je neposreden priključ serije postavk slabši, v kolikor so si po zvoku podobne, kot če se razlikujejo. Torej je lažje ponavljati fonološko različne besede, kot fonološko podobno sosledje. Fonološko podobne postavke imajo manj razlikujočih značilnosti in so torej



bolj podvržene pozabljivosti. Številke npr. so v večini jezikov opredeljene s fonološko zelo različnimi izrazi: ena, dve, tri, štiri, pet...; ein, zwei, drei, vier, fünf...; in japonsko, ichi (1), ni (2), san (3), shi (4), go (5)...

V kolikor so respondenti (večinoma naravoslovne študijske usmeritve) obravnavane postavke v okviru problema STAN kodirali numerično in ne verbalno, so si -v skladu z navedbami - olajšali procesiranje informacij. Na podlagi izsledkov pričujoče raziskave ni možno zaključiti, ali je procesiranje obravnavanih reprezentacij lažje oz. težje, v kolikor so postavke kodirane bodisi numerično bodisi verbalno, kajti numerično kodiranje domen kriterijev vsebuje zgolj problem STAN, medtem ko vsebuje problem AVTO numerično kodiranje samo pri dveh parametrih, za vse druge primere pa je značilna verbalna opredelitev domen kriterijev. Da bi osvetlili odnos med razumljivostjo obravnavanih reprezentacij in načinom kodiranja domen kriterijev (verbalno vs. numerično) bolj natančno, bi bilo potrebno nadaljnje eksperimentalno raziskovalno delo.

Kot ugotavlja Huber (1980), lahko različni tipi informacij (verbalni vs. numerični), s katerimi so opredeljene postavke, pomembno vplivajo na način procesiranja prezentiranega problema. Numerično opredeljene postavke so reprezentirane v spominu bodisi kot internalni numerični sistem, ali pa so kodirane na tak način, da se v spominu ohrani količina oz. velikost, ki jo postavka predstavlja (tj. v analogni obliki). Za razliko od numerične, je v nekaterih pogojih odločanja oz. reševanja problemov potrebno verbalno informacijo na nek način transformirati, npr. s štetjem verbalnih stopenj med dvema verbalnima oznakama, kar poveča kognitivni napor pri reševanju.

Kot je na podlagi empiričnega eksperimenta pokazal Huber, je numerični način kodiranja olajšal procesiranje problema, v kolikor so respondenti uporabili naslednje strategije odločanja v nalogi izbiranja kandidata za vodstveno delovno mesto: MAX (odločevalec izdelava hierarhijo pomembnosti dimenzij in nato hierarhijo alternativ glede na najpomembnejšo dimenzijo in določi najustreznejšo alternativo - tisto, ki je pri vrhu ordinalne lestvice), MIN (postopek je enak kot pri MAX, le da pade izbor na tisto alternativo, ki je na dnu ordinalne lestvice), DIFF (odločevalec ocenjuje razliko med elementi dimenzij in alternativ) in CONCATENATE (odločevalec kombinira kodirane vrednosti elementov dimenzij odločanja za posamezne alternative in izbere alternativo z največjo kombinirano vrednostjo).

Verbalni način kodiranja pa je olajšal procesiranje problema, v kolikor sta bili kot strategiji odločanja uporabljeni EVALUATE (odločevalec na podlagi kodiranih vrednosti za posamezne alternative glede na določeno pomembno dimenzijo izdelava vrednostno lestvico in tako izbere najboljšo alternativo) in CRITERION (odločevalec določi neko vrednost dimenzije odločanja kot kriterij za sprejem alternative).

Kot so pokazali izsledki Huberjeve raziskave, je numerični način kodiranja postavk olajšal tisti tip kognitivnega procesiranja, ki zahteva direktno primerjanje in kombiniranje elementov (MAX, MIN, DIFF, CONCENATE), medtem ko je verbalni način kodiranja olajšal tisto vrsto kognitivnega procesiranja, ki zahteva evaluacijo enega elementa glede na nek internalni ali eksternalni standard (EVALUATE, CRITERION). Razlikovanje med tema tipoma kognitivnih procesov velja za temeljno v okviru teorije odločanja (Huber, 1980). Nadaljnje raziskovalno delo bi pokazalo, kakšen tip kodiranja postavk bi bil najbolj ustrezen glede na različne tipe problemov, glede na različne načine reprezentacije podatkov in glede na različne tipe uporabnikov, npr. z družboslovnim ali naravoslovnim predznanjem. Predpostavljamo namreč lahko, da imajo ljudje različno razvite oz. bolj ali manj avtomatizirane strategije pomnjenja za numerične kot za verbalne postavke, kakor je nakazala tudi raziskava Turnerja in Engla (1989). Na osnovi pričujoče raziskave pa je možno zaključiti, da so s stališča hitrega in



nenapornega razumevanja predstavljenih podatkov najbolj ustrezne take opredelitve domen kriterijev, za katere je značilna razumljivost, nedvoumnost, lahka predstavljenost označenih količin oz. veličin in jasna razločljivost posameznih vrednosti tako semantično kot tudi fonološko. Nadalje so analize zbranih podatkov jasno pokazale, da informacijsko procesiranje bistveno olajšajo tiste predstavitve podatkov, ki vsebujejo enotno opredeljene domene pri vseh kriterijih.

Rezultati raziskave so nadalje pokazali, da so za *trodimensionalne grafe* značilni pomembno nižji  $I_R$  kot za odločitvena drevesa in tabele izpeljanih pravil pri vseh konceptih, razen pri BSPOR in ASTAN, kjer se približajo razumljivosti tabel izpeljanih pravil. Tudi na vseh treh 7-stopenjskih lestvicah so jih respondenti ocenili signifikantno nižje kot prvi dve reprezentaciji, in sicer so na lestvici razumljivosti dobili povprečno oceno 4.22, na lestvici težavnosti 3.85, na lestvici preglednosti pa so dobili najnižjo oceno izmed vseh štirih reprezentacij, namreč 3.63. Sicer pa so iz grafične predstave  $I_R$  tudi za trodimenzionalne grafe razvidna podobna gibanja in stopnje variacij  $I_R$  v odvisnosti od kompleksnosti, semantične vsebine problemskega področja in tipa informacij, ki jih vsebuje koncept, kot pri tabelah izpeljanih pravil. Torej, razumljivost trodimenzionalnih grafov izrazito upada s stopnjo kompleksnosti reprezentacije, odvisna pa je tudi od semantične vsebine koncepta ter tipa vsebovanih informacij. Trodimenzionalni grafi dosežejo podobno kot izpeljana pravila najnižjo stopnjo razumljivosti pri konceptu BSTAN, za katerega je značilna visoka stopnja kompleksnosti, s stališča respondentov nepoznana vsebina in kognitivno obremenjujoča opredelitev domen kriterijev. Po drugi strani pa so za oba problema, ki so ju respondenti označili kot lahka, značilni višji  $I_R$  tako v pogojih nizke (AAVTO, ASPOR) kot tudi v pogojih višje (BAVTO, BSPOR) stopnje kompleksnosti. Ta značilnost, namreč majhna razlika v  $I_R$  med manj in bolj kompleksno predstavitvijo, je razvidna tudi pri problemu KADR, ki se uvršča med semantično lažje probleme, čeprav dosega nizke  $I_R$ . Podobno kot teste izpeljanih pravil, so tudi teste trodimenzionalnih grafov respondenti reševali sorazmerno počasi, zaradi česar število pravih odgovorov pri obeh zadnjih konceptih (AKADR, BKADR) opazno upade. Kot smo pokazali za izpeljana pravila, je učinek interakcije med obema dejavnikoma, ki vplivata na obremenjenost delovnega spomina, tj. kompleksnost in pomenljivost, iz prezentiranih rezultatov jasno razviden tudi v primeru trodimenzionalnih grafov.

Kot je nazorno razvidno iz grafičnega prikaza  $I_R$ , je za trodimenzionalne grafe značilen izrazit upad razumljivosti v primeru BPORT. Kot pokaže primerjalna analiza, gre za koncept, za katerega je značilna visoka stopnja kompleksnosti, po semantični vsebini pa se uvršča med težje razumljive. Poleg tega je reprezentacija problema PORT (vrednotenje proizvodnih programov s portfolio analizo) edina, ki lomi običajno konvencijo, da vrednosti spremenljivke na vertikalni osi naraščajo od spodaj navzgor. V primeru PORT pa naraščajo vrednosti na vertikalni osi od zgoraj navzdol, kar respondentov, kot je razvidno iz prikazov  $I_R$ , ni motilo v primeru APORT - v testu koncept z najnižjo stopnjo kompleksnosti (6), medtem ko je očitno povzročilo pomembno kognitivno obremenitev pri konceptu BPORT -v pogojih višje stopnje kompleksnosti (15). Poleg tega so pri konceptu APORT domene enotno opredeljene pri vseh kriterijih, pri konceptu BPORT pa različno.

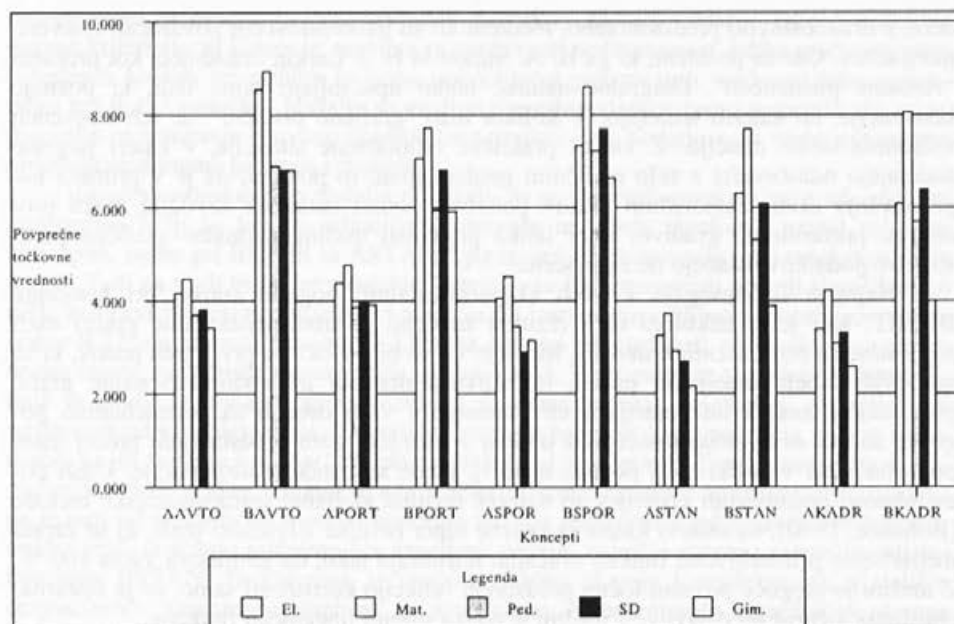
Kot je pokazala analiza razlogov, kaj je respondentom pri reševanju testa povzročalo težave, so pogosto navajali, da "trodimenzionalnih grafov ne znajo brati", težave pa so imeli tudi zaradi "slabe razločljivosti ravnin". Za razliko od odločitvenih dreves in tabel izpeljanih pravil, katerih razumevanje temelji na bolj splošno poznanih strategijah branja, je za trodimenzionalne grafe značilno, da ustreznih operacij (mentalnih korakov), potrebnih za njihovo razumevanje, ne obvladujejo predvsem respon-

denti z družboslovno predizobrazbo, medtem ko so naravoslovcem povzročali bistveno manj težav. Gre za problem, ki ga H. A. Simon in H. J. Larkin označujeta kot problem "vizualne pismenosti". Diagrame namreč lahko uporabljajo samo tisti, ki poznajo konvencije, na katerih temeljijo. V kolikor niso "grafično pismeni", se teh konvencij večinoma lahko naučijo. Z vidika praktične odločitvene situacije, v kateri pogosto sodelujejo odločevalci z zelo različnim predznanjem, to pomeni, da je v primeru nepoznavanja trodimenzionalnih grafov potrebno podati natančna navodila, kako procesirati prezentirano gradivo; sicer lahko prednosti trodimenzionalne grafične predstavitve podatkov ostanejo neizkoriščene.

Najnižje  $I_R$  dosegajo v vseh eksperimentalnih pogojih (razen pri konceptu BPORT, kjer je le nekoliko nižji rezultat značilen za trodimenzionalne grafe) *uteži pomembnosti posameznih kriterijev*. Medtem ko so bile v okviru prvih treh testov, ki so vsebovali tabele izpeljanih pravil, odločitvena drevesa in trodimenzionalne grafe, predstavitve znanja informacijsko ekvivalentne v vseh desetih eksperimentalnih pogojih, so bili sicer enaki koncepti z utežmi v zadnjem testu predstavljeni precej manj podrobno, kar v veliki meri pojasni upad  $I_R$  glede na ostale reprezentacije. Uteži pomembnosti posameznih kriterijev so namreč rezultat globalne aproksimacijske metode (Bohanec, 1990), na osnovi katere iz enačbe hiper ravnine izluščimo uteži, ki se zaradi medsebojne primerljivosti funkcij običajno normirajo tako, da je njihova vsota 100 %. Z utežmi je mogoče povsem točno predstaviti funkcijo koristnosti samo, če je linearna; absolutna točnost predstavitve z utežmi je enaka stopnji linearnosti funkcije.

Uteži pomembnosti posameznih kriterijev so respondenti ocenili na 7-stopenjskih lestvicah kot najmanj razumljive (4.18) in najtežje od vseh reprezentacij ekspertnega sistema DEX, medtem ko so na lestvici preglednosti dobile sorazmerno visoko oceno (4.02), ki jih v tem pogledu uvršča takoj za odločitvena drevesa. Sicer pa je tudi za uteži značilen upad  $I_R$  z rastjo stopnje kompleksnosti reprezentacije in variiranje  $I_R$  glede na semantično vsebino koncepta ter tip vsebovanih informacij. Tudi uteži dosežejo najnižjo točko razumljivosti pri konceptu BSTAN, ki se, kot smo pokazali, uvršča med semantično težje predstavitve logičnih pravil. Učinek interakcije med obema dejavnikoma, ki vplivata na obremenjenost delovnega spomina, tj. kompleksnost in pomenljivost, je iz prezentiranih rezultatov, podobno kot pri izpeljanih pravilih in trodimenzionalnih grafih, jasno razviden tudi v primeru reprezentacije logičnih pravil z utežmi pomembnosti posameznih kriterijev. Za problema AVTO in SPOR, ki sta semantično manj zahtevna, je namreč tako kot pri izpeljanih pravilih in grafih tudi pri utežeh razvidna majhna razlika v  $I_R$  med manj in bolj kompleksno predstavitvijo, medtem ko je za oba semantično zahtevnejša problema STAN in PORT značilen velik upad  $I_R$  v pogojih višje stopnje kompleksnosti.

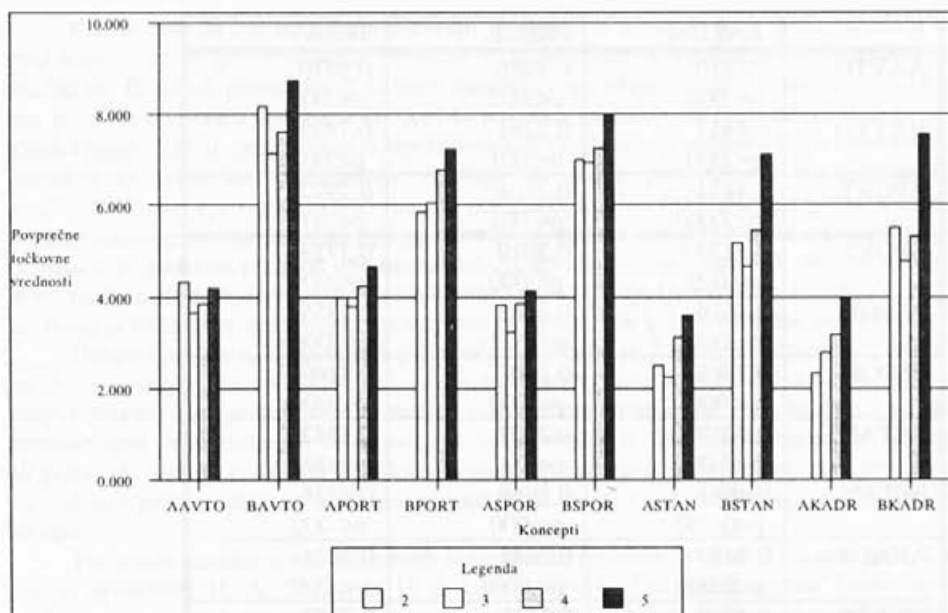
Razumevanje, pomnjenje in priklic prezentiranih informacij je odvisno od kompleksnih interakcij med značilnostmi predstavitve gradiva po eni strani, kar smo predhodno obravnavali, ter kognitivnimi in motivacijskimi značilnostmi recipienta, kot so pozornost, znanje, izkušnje, učne strategije in stališča, po drugi strani. Tudi iz rezultatov pričujoče raziskave so razvidne signifikantne *individualne razlike* med respondenti v stopnjah razumevanja in hitrosti razbiranja logičnih pravil iz obravnavanih reprezentacij. Kot je pokazala multivariantna analiza variance, obstajajo med respondenti glede na njihovo študijsko usmeritev statistično visoko signifikantne razlike v doseženih točkovnih vrednostih v vseh eksperimentalnih pogojih (Wilks' Lambda=.696,  $p < .000$ ). Najvišje testne rezultate dosegajo študenti matematike, ki jim sledijo študenti elektrotehnike in računalništva, medtem ko dosegajo precej slabše rezultate študenti pedagogike in Visoke šole za socialne delavce, najnižje točkovne vrednosti pa so značilne za dijake Gimnazije Poljane (Slika 4).



Slika 4. Grafični prikaz odnosa med povprečnimi točkovnimi vrednostmi pri desetih konceptih in študijsko usmeritvijo.

Medtem ko respondenti z naravoslovno predizobrazbo pomembno boljje obvladujejo strategije branja obravnavanih reprezentacij, povzročajo te operacije družboslovno usmerjenim respondentom očitno precej več težav. Rezultat ni presenetljiv, saj v okviru družboslovnih študijskih programov študenti relativno redkeje uporabljajo npr. znanje, izraženo v utežeh oz. se seznanjajo z grafičnimi načini prikazovanja in interpretacije podatkov, logičnih pravil in zakonitosti, kot v okviru naravoslovnih študijskih usmeritev. Učenje in spomin imata pomembne učinke na relativni kognitivni napor pri uporabi različnih strategij razumevanja prezentiranih podatkov ter na hitrost branja. V tej zvezi se pojavijo pomembna vprašanja o interakciji med kapaciteto delovnega spomina in stopnjo, do katere je uporabljena strategija (zaporedje mentalnih korakov) avtomatizirana in tako zahteva manj kognitivnih naporov ter dopušča človeku več svobode pri uporabi delovnega spomina za druge operacije.

Multivariantna analiza variance je pokazala tudi signifikantne razlike med dosežki respondentov glede na spremenljivko "učni uspeh" (Wilks' Lambda=.784,  $p<.000$ ): najvišje točkovne vrednosti dosegajo respondenti z odličnim učnim uspehom, sledijo jim respondenti s prav dobrim uspehom, slabši pa so bili respondenti z dobrim in zadostnim uspehom (Slika 5). Medtem ko dosežke respondentov glede na njihovo študijsko usmeritev lahko pojasnimo z vidika stopnje avtomatiziranosti ustreznih strategij branja, lahko predpostavljamo, da so razlike med posamezniki, ki jih pokaže variabla "učni uspeh", v večji meri odraz individualnih razlik v kognitivnih sposobnostih in v motivaciji respondentov.



Slika 5. Grafični prikaz odnosa med povprečnimi točkovnimi vrednostmi pri desetih konceptih in učnim uspehom.

Izračun Pearsonovih korelacijskih koeficientov med povprečnimi točkovnimi vrednostmi in povprečnimi ocenami na 7-stopenjskih lestvicah razumljivosti, preglednosti in težavnosti reprezentacij znanja je nadalje pokazal pozitivno in statistično visoko signifikantno povezanost med obema spremenljivkama v vseh eksperimentalnih pogojih (Tabela 2). Respondenti, ki so dosegali višje točkovne vrednosti, so tudi reprezentacije znanja ocenjevali višje, tisti pa, ki so dosegali slabše testne rezultate, so reprezentacije znanja ocenili kot manj razumljive, manj pregledne in težavnije. Z vidika praktične odločitvene situacije to pomeni, da uporabnikova opredelitev na omenjenih treh lestvicah lahko služi kot sorazmerno dober prediktor (ne)razumevanja uporabljenih reprezentacij logičnih pravil. V kolikor je izražena ocena nižja od 4 na sedem-stopenjski lestvici, je mogoče pričakovati precej slabo razumevanje in počasno razbiranje logičnih pravil iz prezentiranih podatkov.

	RAZUM	PREGL	TEŽAV
AAVTO	0.4530 p<.000	0.1986 p<.002	0.3979 p<.000
BAVTO	0.4457 p<.000	0.2241 p<.001	0.3455 p<.000
APORT	0.4151 p<.000	0.2104 p<.001	0.2992 p<.000
BPORT	0.5139 p<.000	0.3019 p<.000	0.3728 p<.000
ASPOR	0.4616 p<.000	0.1483 p<.022	0.3355 p<.000
BSPOR	0.3986 p<.000	0.2307 p<.000	0.4169 p<.000
ASTAN	0.4017 p<0.000	0.2002 p<.002	0.3542 p<.000
BSTAN	0.4981 p<0.000	0.2966 p<.000	0.4436 p<.000
AKADR	0.4489 p<0.000	0.2645 p<.000	0.3821 p<.000
BKADR	0.4290 p<.000	0.2530 p<.000	0.3927 p<.000
VSOTA	0.5732 p<.000	0.3074 p<.000	0.4841 p<.000

*Tabela 2. Korelacije (Pearsonovi korelacijski koeficienti) med povprečnimi točkovnimi vrednostmi in povprečnimi ocenami na 7-stopenjskih lestvicah razumljivosti, preglednosti in težavnosti reprezentacij znanja.*

Na tej osnovi lahko zaključimo, da je v postopku izgradnje baze znanja pomembno, da ekspert oz. skupina ekspertov lahko razmišlja in rešuje relevantne probleme na podlagi različno predstavljenih podatkov in zakonitosti, kajti, kot smo predhodno poudarili, različni, čeprav informacijsko ekvivalentni načini reprezentacije podatkov, vzpodbujajo povsem različne kognitivne procese. Medtem ko vzpodbujajo grafične reprezentacije bolj paralelno, sintetično obdelovanje informacij, značilno predvsem za desno cerebralno hemisfero, pospešujejo verbalne bolj sekvenčne, analitične kognitivne procese, ki so značilni za levo cerebralno hemisfero. Ker povezuje obe hemisferi interhemisferno živčno nitje (corpus callosum), ki omogoča notranjo koordinacijo in sodelovanje, potekata oba tipa predelovanja informacij hkrati in povezano med reševanjem različnih nalog.

V tej zvezi je pomemben tudi izsledek, da je človek zmožen sprejemati in obdelovati večjo količino informacij, v kolikor so predstavljene v dveh modalnostih, tj. grafično in verbalno, kot če so vse informacije predstavljene zgolj v eni modalnosti; kar so pokazale tako nevropsihološke raziskave (Farah, 1989) kot tudi psihološki eksperimenti, v katerih so uporabili tehnike sekundarnih nalog (Baddeley, 1992, 1994). Zato je v praktični odločitveni situaciji pomembno premostiti komunikacijski prepad med inženirjem znanja in ekspertom oz. skupino ekspertov ter podati natančna navodila, kako procesirati prezentirano gradivo, sicer lahko prednosti, zlasti trodimenzionalnih grafičnih predstavitev podatkov in uteži, ostanejo neizkoriščene.



Kot je bilo že nakazano, pa odpirajo ugotovljeni rezultati tudi vrsto vprašanj in problemov, ki bi jih bilo treba proučiti z nadaljnjimi teoretičnimi in empiričnimi študijami. Eden od pomembnih vidikov razumljivosti obravnavanih reprezentacij znanja je način kodiranja domen kriterijev, ki je lahko verbalen ali numerični. Kot je pokazal Huber (1980), lahko različni tipi informacij (verbalni vs. numerični), s katerimi so opredeljene postavke, pomembno vplivajo na način procesiranja prezentiranega problema. Nadaljnje raziskovalno delo bi pokazalo, kakšen tip kodiranja postavk bi bil najustreznejši glede na različne tipe problemov, glede na različne načine reprezentacije podatkov in glede na različne tipe uporabnikov, npr. z družboslovnim ali naravoslovnim predznanjem. Predpostavljamo namreč lahko, da imajo ljudje različno razvite oz. bolj ali manj avtomatizirane strategije pomnjenja za numerične kot za verbalne postavke.

Drugo vprašanje, ki ga odpira pričujoča raziskava, je, katera od kombinacij obravnavanih reprezentacij podatkov, npr. tabela izpeljanih pravil + trodimenzionalni graf oz. uteži pomembnosti posameznih kriterijev + trodimenzionalni graf ipd., bi dala z vidika razumevanja predstavljenih informacij najboljše rezultate. Nadaljnje empirične študije bi pokazale, katere kombinacije obravnavanih reprezentacij so najustreznejše ter kateri od načinov prikazovanja (sekvenčni vs. simultani) bi dal z vidika razumevanja najboljše učinke.

Pričujoča raziskava na več mestih opozarja na problem "vizualne pismenosti", kot sta ga opredelila H. A. Simon in H. J. Larkin (1987). Diagrame namreč lahko uporabljajo samo tisti, ki poznajo konvencije, na katerih temeljijo. V kolikor niso "grafično pismeni", se teh konvencij večinoma lahko naučijo. V okviru nadaljnjega dela bi bilo treba posvetiti posebno pozornost tudi navodilom oz. učenju strategij branja posameznih reprezentacij ter pri tem upoštevati individualne razlike med recipienti.

## ZAHVALA

Vsem sodelavcem Laboratorija za umetno inteligenco na Inštitutu "Jožef Stefan" v Ljubljani se iskreno zahvaljujemo za vsestransko pomoč in nasvete pri zasnovi in izvedbi pričujoče študije, še posebej prof. dr. Ivanu Bratku, vodji laboratorija, doc. dr. Nadi Lavrač, doc. dr. Marku Bohancu, dr. Bojanu Cestniku in Božu Urhu, dipl. ing. Ministrstvu za znanost in tehnologijo Republike Slovenije se zahvaljujemo za sredstva, ki so v okviru akcije "2000 novih raziskovalcev" omogočila to raziskovalno delo.

## LITERATURA

- BADDELEY, A. D. (1992). Working Memory. *Science*, 255, 556-559.
- BADDELEY, A. D. (1994). *Your Memory*. London: Penguin Books, 1994.
- BARON, J. (1988). *Thinking and deciding*. Cambridge: Cambridge University Press.
- BERGADANO, F., MATWIN, S., MICHALSKI, R.S. & ZHANG, J. (1988). Measuring quality of concept descriptions. *Proceedings of the Third European Working Session on Learning (EWSL)*, 1-14. Glasgow.
- BOHANEČ, M. (1990). *Metode za vrednotenje variant in razlago znanja v večparametrskem odločanju*. Doktorska disertacija. Univerza v Ljubljani, Fakulteta za elektrotehniko.
- BOHANEČ, M., GYERGYEK, L., RAJKOVIČ, V. (1988). Večparametrsko odločanje, podprto z lupino eksperimentalnega sistema. *Elektrotehniški vestnik*, 55, 3-4, 189-198.
- BOHANEČ, M. & RAJKOVIČ, V. (1990). DEX: An expert system shell for decision support. *Sistémica*, 1, 145-158.
- BOHANEČ, M., URH, B. & RAJKOVIČ, V. (1992). Evaluating options by combined qualitative and quantitative methods. *Acta Psychologica*, 80, 67-89.

- BOHANEC, M. & BRATKO, I. (1994). Trading accuracy for simplicity in decision trees. *Machine Learning*, 15, 223-250.
- BOHANEC, M., RAJKOVIĆ, V., SEMOLIĆ, B. & POGAČNIK, A. (1995). Knowledge-based portfolio analysis for project evaluation. *Information & Management*, 28, 293-302.
- BRATKO, I. & KODRATOFF, Y. (1989). An analytical report of EWSL-88. *AI Communications*, 2, 24-29.
- BRATKO, I. (1990). *Prolog programming for artificial intelligence*. Wokingham: Addison-Wesley.
- DANFMAN, M., CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 19, 450-466.
- DE VEGA, M., DIAZ, J. M. (1991). Building referents of indeterminate sentences in the context of short narratives. V: R. H. Logie, M. Denis (Ur.), *Mental Images in Human Cognition*. Amsterdam: Elsevier (North-Holland).
- EVANS, J. ST B. T. (1988). The knowledge elicitation problem: a psychological perspective. *Behaviour and Information Technology*, 7, 111-130.
- FARAH, M. J. (1989). Knowledge from text and pictures: A neuropsychological perspective. V: H. Mandl, J. R. Levin (Ur.), *Knowledge Acquisition from Text and Pictures*. Amsterdam: Elsevier (North-Holland).
- FLETCHER, C. R. & BLOOM, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language*, 27, 235-244.
- GAINES, B.R. (1989). Social and cognitive processes in knowledge acquisition. *Knowledge Acquisition*, 1, 39-58.
- GIARRATANO, J. & RILEY, G. (1994). *Expert systems: Principles and programming*. Boston: PWS.
- HAWKINS, D. (1983). An analysis of expert thinking. *International Journal of Man-Machine Studies*, 18, 1-47.
- HIRST, W. (1988). Improving memory. V: M. S. Gazzaniga (Ur.), *Perspectives in Memory Research*. Cambridge MA: MIT Press.
- HUBER, O. (1980). The influence of some task variables on cognitive operations in an information-processing decision model. *Acta Psychologica*, 45, 187-196.
- JOHNSON-LAIRD, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press.
- JOHNSON-LAIRD, P. N. (1988). *The Computer and the Mind: An Introduction to Cognitive Science*. Cambridge, MA: Harvard University Press.
- KINTSCH, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, 95, 163-182.
- KINTSCH, W. & VAN DIJK, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- KODRATOFF, Y. (1987). Is AI a sub-field of computer science - or is AI the science of explanations? V: I. Bratko, N. Lavrač (Ur.), *Progress in Machine Learning*. Wilmslow: Sigma Press.
- LALOMIA, M. J., COXWERT, M. D. & SALAS, E. (1992). Problem-solving performance as a function of problem type, number progression, and memory load. *Behaviour and Information Technology*, 11, 268-280.
- LARKIN, H. J. & SIMON, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.
- MICHIE, D. & BRATKO, I. (1986). *Expert systems: Automatic knowledge acquisition*. Wokingham: Addison-Wesley.
- MONTAZEMI, A. R. (1991). The impact of experience on the design of user interface. *International Journal of Man-Machine Studies*, 34, 731-749.
- MUSEK, J. (1990). *Simboli, kultura, ljudje*. Ljubljana: Znanstveni inštitut Filozofske fakultete.
- MUSEK, J. (1993). *Znanstvena podoba osebnosti*. Ljubljana: Educy.
- NEALE, I. M. (1988). First generation expert systems: a review of knowledge acquisition methodologies. *The Knowledge Engineering Review*, 3, 105-145.
- NEWELL, A. & SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- NOSEK, J. T., ROTH, I. (1990). A comparison of formal knowledge representation schemes as communication tools: predicate logic vs semantic network. *International Journal of Man-Machine Studies*, 33, 227-239.
- ORMEROD, T. C., MANKTLOW, K. I., ROBSON, E. H., STEWARD, A. P. (1986). Content and representation effects with reasoning tasks in Prolog form. *Behaviour and Information Technology*, 5, 157-168.
- PAIVIO, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston.
- PUPPE, F. (1993). *Systematic introduction to expert systems: Knowledge representation and problem solving methods*. Heidelberg: Springer-Verlag.
- RAJKOVIĆ, V. (1987). *Večparametrski odločitveni postopek, ki temelji na metodah kibernetike in umetne inteligence*. Doktorska disertacija. Univerza v Ljubljani, Fakultetaza elektrotehniko.

- RAJKOVIČ, V. & BOHANEČ, M. (1991). Decision support by knowledge explanation. V: H.G. SOL & J. VECSENYI (Ur.), *Environments for supporting decision processes*, 47-57. Amsterdam: North-Holland.
- RAJKOVIČ, V., KRISPER, M., SAGAJIN, T., BUKVIČ, V. (1991). *Razvoj ekspertnega sistema za podporo strateškemu vodenju na podjetniški ravni*. Celje: Razvojni center, Inštitut za raziskave in razvojni inženiring.
- SUBRAMANIAN, G. H., NOSEK, J., RAGHUNATHAN, S. P. & KANITKAR, S. S. (1992). A comparison of the decision table and tree. *Communications of the ACM*, 35, 89-94.
- TRSTENJAK, A. (1980). Vertikalnost in lateralnost mišljenja v dilemah ustvarjalnosti. *Anthropos*, 1-2, 157-194.
- TRSTENJAK, A. (1981). *Psihologija ustvarjalnosti*. Ljubljana: Slovenska matica.
- TURNER, M. L., ENGLE, R. W. (1989). Is working memory task dependent? *Journal of Memory and Language*, 28, 127-154.
- VAN DEN BROEK, P. (1990). Causal inferences and the comprehension of narrative texts. V: A. C. GRAESSER & G. BOWER (Ur.) *Inferences and text comprehension. The psychology of learning and motivation*, zv. 25, 175-196. New York: Academic Press.
- VAN DIJK, T. A. & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- WAERN, Y. (1989). *Cognitive aspects of computer supported tasks*. New York: Wiley.
- VESSEY, I., WEBER, R. (1986). Structured tools and conditional logic: An empirical investigation. *Communications of the ACM*, 29, 48-57.
- WINN, W. (1989). The design and use of instructional graphics. V: H. Mandl, J. R. Levin (Ur.), *Knowledge Acquisition from Text and Pictures*. Amsterdam: Elsevier (North-Holland).
- VIZJAK PAVŠIČ, M. (1994). *Razumljivost baz znanja kot dejavnik učinkovitosti ekspertnih sistemov za podporo odločanju*. Doktorska disertacija. Univerza v Ljubljani, Filozofska fakulteta.