

# Prezentacija znanja stablom odlučivanja

## Knowledge Presentation in Decision Tree

Mentor: Prof. dr. sc. Branko Jeren

## Zadatak

U sklopu diplomskog zadatka potrebno je razmotriti korištenje prezentacije znanja kroz stablo odlučivanja (eng. Decision tree). Istražiti stanje znanosti i tehnike u tom području te evaluirati raspoložive programske alate. Naglasak staviti na jednostavnost upotrebe za korisnike s minimalnim informatičkim vještinama te na dobavljalivost i troškove korištenja alata. Prilagoditi jedan takav alat i istražiti njegovu upotrebljivost za savjetovanje pri odabiru literature za učenje ili odgovarajućeg obrazovnog programa.

U vezi dodatnih informacija obratiti se mentoru.

## Kronologija

### 1. tjedan (5.-11.3.2012.)

#### Izveštaj

Upoznala se s osnovnim pojmovima vezanima za temu diplomskog rada (Knowledge Management, Knowledge Representation, Decision Tree). Počela voditi Zotero bibliografiju i započela razradu mape područja. Za izradu mape područja koristim besplatni alat FreeMind dostupan na:

[http://freemind.sourceforge.net/wiki/index.php/Main\\_Page](http://freemind.sourceforge.net/wiki/index.php/Main_Page)

#### Plan za naredni tjedan:

- na sastanku s mentorom točno odrediti koje teme u preglednom radu obraditi
- detaljno proučiti i isprobati alate poput WEKA-a, Orange-a, KNIME-a, ELKI-ja i ostalih
- započeti rad na preglednom radu (po mogućnosti dovršiti do 21.3)

### 2. tjedan (12.-18.3.2012.)

#### Izveštaj sa sastanka, ponedjeljak, 12.3.2012, 9:30

- rasprava o načinima definiranja znanja u stablima odlučivanja
- Je li predikatna logika jedino što se koristi za definiranje znanja? Je li ona dovoljno jednostavna za neinformatičare?
- definiran tjedni zadatak
- motivacija: formaliziranja i reprezentacija znanja kao način za provjeru znanja

## Tjedni zadatak

- istražiti sve načine prezentiranja znanja
- staviti prezentiranje znanja (Knowledge Representation) u kontekst s umjetnom inteligencijom (Artificial Intelligence) i strojnim učenjem (Machine Learning) i ostalim povezanim pojmovima

## **Izveštaj**

Proučila pojmove znanje (engl. knowledge), prezentacija znanja (engl. knowledge representation), tehnologije prezentacije znanja (engl. knowledge representation technologies). Prezentacija znanja je jedno od područja kojim se bavi umjetna inteligencija (engl. artificial intelligence).

Tehnologije prezentacije znanja:

- Logic based representation – first order predicate logic, Prolog
- Procedural representation – rules, production system
- Network representation – semantic networks, conceptual graphs
- Structural representation – scripts, frames, objects

Plan za naredni tjedan:

- dogovoriti daljnje smjernice na sastanku s mentorom
- prijedlog: proučiti detaljno stabla odlučivanja, gdje se ona uklapaju u tehnologije prezentacije znanja?
- je li znanje prezentirano u semantičkim mrežama i konceptualnim grafovima za čovjeka najintuitivniji način prikaza?
- uklapa li se takav način prikaza u prikaz stablom odlučivanja?

## **3. tjedan (19.-25.3.2012.)**

**Izveštaj sa sastanka, ponedjeljak, 19.3.2012, 11:00**

- prezentirala načine prezentacije znanja; rasprava; potrebno dodati primjere
- prezentacija: [knowledge\\_representation\\_research.pptx](#)
- rasprava o daljnjem radu
- koja se područja još bave prezentacijom znanja

Plan za daljnji rad:

- napraviti wiki stranicu o načinima prezentacije znanja po uzoru na Teorije učenja; nazvati ju Knowledge Representation
- pronaći jedan primjer koji je potrebno prikazati na sva četiri načina (logički, proceduralni, mrežni i strukturalni) npr. fotografija, kolači
- istražiti područje prezentacije znanja u nekim drugim područjima različitim od područje umjetne inteligencije

Smjernice:

- potražiti materijale na teme "How to develop a course?" "How to write a textbook?" "structuring/presenting knowledge"
- kako se formaliziraju znanja za pisanje enciklopedije?
- kako razlučiti što je bitno?

#### 4. tjedan (26.3-1.4.2012.)

Wiki stranica sa sadržajima o prezentaciji znanja: [Knowledge Representation](#)

#### 5. tjedan (2.4-8.4.2012.)

Prezentacija znanja u udžbenicima. Istržila karakteristike dobrih udžbenika i metodologiju pisanja kvalitetnih udžbenika. Pretraživanje u okviru područja "educational psychology", a po ključnim riječima: methodology in textbooks, textbook design, writing textbooks, developing textbooks, design of instructional texts, courseware design, textbook evaluation.

Članak: [Hypertext Principles for Text and Courseware Design](#)

- vizionarski članak u kojem autor objašnjava pojmove hiperteksta i dinamičkog teksta koji su preteča današnjih e-udžbenika
- predstavlja se koncept individualiziranog učenja, koje omogućava studentu učenje na način koji mu najbolje odgovara (različiti tijek usvajanja znanja, pretraživanje teksta, povezivanje pojmova, dodavanje bilješki, uređivanje teksta)
- "We need to be able to represent subject matter knowledge and model its structure, including knowledge representation, in a way that will provide principles and strategies of instruction."
- representation of subject knowledge - network representation - schema theory
- modeling the structure of knowledge - nove informacije postaju novi čvorovi u mreži: web learning principles → web teaching principles: "The effective teacher presents the material in a way that allows learners to develop some framework for relating materials to each other (web) and then elaborating the material (covering it in depth)
- mapping the network structure of the subject matter with the semantic network of the learner - prepoznavanje što student ne zna, način za ispitivanje znanja?
- procedure for designing hypertexts:

1. Identify all key concepts
2. Map the structure of the content
3. Verify the structure
4. Determine the type of hypertext structure
5. Prepare the concept blocks
6. Provide links and cues to other concepts
7. Debug the system

- bitna je struktura i organizacija sadržaja: top-level struktura odlomka pridonosi boljem razumijevanju i učenju
- upute za strukturiranje teksta: namjena i struktura trebaju biti jasno izražene kroz naslove, podnaslove; struktura teksta mora iskazivati autorov stav; tekst mora biti organiziran u okvire s konzistentom strukturom

Članak: [Strategy for the Development of the Textbook](#)

- “education—science—technology—production” system
- utjecaj razvoja tehnologije na obrazovni sustav i stvaranje udžbenika
- problem kvalitete udžbenika: nepostojanje standarda i loš prikaz znanja (didaktičke karakteristike vizualizacije objekata i fenomena)
- smjernice za rješavanje problema: osnovni i primijenjeni pristup

#### Članak: Schema Theory and the Design of Content-Area Textbooks

- čitatelji razumiju tekst tako što konstruiraju odgovarajuće sadržajne sheme (organizirano znanje o objektima, događajima, situacijama) i tekstualne sheme (organizirano znanje o tekstovima od dijelova teksta do bajki i znanstvenih članaka)
- tekstualne sheme: istraživanja su pokazala da čitatelji bolje pamte informacije iz tekstova organiziranih prema određenim konvencijama
- signali - naziv za sredstva i načine kojima se autori služe za strukturiranje tekstova: veznici (uzročni, sastavni, rastavni, vremenski), riječi naglašavanja i izražavanja stavova poput nažalost; tipografska naglašavanja poput bold, italics, podcrtavanje; “preview” statements; sažeci
- sadržajne sheme: analogije i usporedbe

#### Textbook Writer: A tool for writing effective online college textbooks

- jednostavan koncept “flipped classroom” - stjecanje teorijskih znanja izvan učionice, a praktičnih vještina u učionici uz pomoć eksperta

### **6. tjedan (9.4-15.4.2012.)**

Prikazi znanja na primjeru.

#### Plan za naredni tjedan

- do sada istražen pojam prezentacije znanja na područjima umjetne inteligencije i psihologije odgoja i obrazovanja
- istražiti KR u poslovnom svijetu - što se koristi u kompanijama?

### **7. tjedan (16.4-22.4.2012.)**

#### **Izveštaj sa sastanka, ponedjeljak, 16.4.2012, 10:00**

- rasprava o dosadašnjim rezultatima: psihologija obrazovanja i prezentacija znanja
- smjernice za daljnji rad:
- prezentacija znanja s obzirom na vrstu znanja (deklarativno i proceduralno)
- pronaći vizualne alate za formalizaciju znanja s naglaskom na dva pitanja:

1. jednostavnost uporabe: Koliko je intuitivno korištenje alata?
2. mogućnost obrade: Može li prezentirano znanje alat iskoristiti za rješavanje nekog konkretnog problema? Također predstavlja jedan od načina za ispitivanje znanja.

- istražiti koncepte mind (Mind42 i FreeMind) i concept mappinga (cmap.ihmc.us) -

sličnosti/razlike

Koncept Personal Knowledge Managementa. Personal Knowledge Management vs. Knowledge Management. Mogu li se PKM alati koristiti i za KM?

### Deklarativno znanje

- Architectures with declarative representations have knowledge in a format that may be manipulated, decomposed and analyzed by its reasoners
- Advantages of declarative knowledge are: The ability to use knowledge in ways that the system designer did not foresee
- logic

### Proceduralno znanje

- Architectures with procedural representations encode how to achieve a particular result
- Advantages of procedural knowledge are: Possibly faster usage
- scripts

### Mind Maps

- za strukturiranje tema; hijerarhijske; stablasta građa
- pogodne za brain-storming i pregled određenog područja
- moguća su povezivanja pojmova, ali onda mapa postaje nepregledna

### Concept Maps

- struktura grafa; pogodne za povezivanja
- nema tako jasne hijerarhije kao kod mind mapa
- Članak: The Theory Underlying **Concept Maps** and How to Construct and Use Them
- graphical tool for organizing and representing knowledge
- **concepts**, usually enclosed in circles or boxes of some type, and **relationships** between concepts indicated by a connecting line linking two concepts
- concepts are represented in a **hierarchical fashion** with the most inclusive, most general concepts at the top of the map and the more specific, less general concepts arranged hierarchically below
- it is best to construct concept maps with reference to some particular question we seek to answer, which we have called a **focus question**
- **cross-links** are relationships or links between concepts in different segments or domains of the concept map
- There are two features of concept maps that are important in the facilitation of creative thinking: **the hierarchical structure** that is represented in a good map and **the ability to search** for and characterize new cross-links
- CmapTools program: <http://cmap.ihmc.us/> - instaliran

### Članak: Conceptual Data Structures for Personal Knowledge Management

- predstavlja razvijeni alat iMapping (zajedno s QuiKeyem) <http://www.imapping.info/>
- postojeće vizualne tehnike mapiranja: mind maps, concept maps i slobodan tekst (natuknice - koncept pinboarda)
- QuiKey - mind-reading semantic command line; authoring (adding entities and statements), querying, browsing

## 8. tjedan (23.4-29.4.2012.)

### Izveštaj sa sastanka, ponedjeljak, 23.4.2012, 09:30

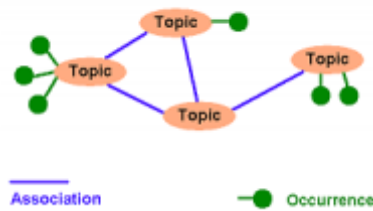
- rasprava o konceptu personal knowledge management i alatu iMapping te njegovim mogućnostima
- smjernice za daljnji rad
- detaljnije proučiti alat iMapping i usredotociti se na postavljanja upita putem QuiKeya
- proučiti Ihmc i druge concept mapping alate - mogu li eksportirati podatke u formatu prikladnom za strojno pretraživanje?

### IHMC CmapTools (Florida Institute for Human and Machine Cognition)

- IHMC podučja djelatnosti: prezentacija znanja u formi pogodnoj za računalnu obradu; istraživanje i razvijanje standarada za razvoj semantičkog Interneta - Interneta koji pruža bogatiji i bolje povezani sadržaj; ključne tehnologije za razvoj semantičkog Interneta su Resource Description Framework (RDF), Web Ontology Language (OWL) i SPARQL; Common Logic (CL) - standard predložen od ISO-a za razmjenu i prijenos znanja u računalima, baziran na predikatnoj logici
- omogućuje gradnju, pretragu i dijeljenje znanja modeliranog kroz koncept mape
- svaki entitet je moguće dodatno pojasniti linkom na novu koncept mapu, sliku, video ili dokument
- korisnici mogu spremati svoje mape na vlastitim računalima ili na Cmap serverima te tako dijeliti svoje mape s ostalim korisnicima
- pri svakom spremanju mape automatski se generira njena web stranica na nekom od servera
- stranica koja omogućava pretraživanje svih javno dostupnih koncept mapa i sadržaja o njima na webu: <http://www.cmappers.net/>
- mogućnost eksportiranja podataka u HTML, CXL, XTM/XCM, IVML; koncept mapa o eksportiranju [http://cmapskm.ihmc.us/servlet/SBReadResourceServlet?rid=1064009710027\\_522212290\\_27111&partName=htmltext](http://cmapskm.ihmc.us/servlet/SBReadResourceServlet?rid=1064009710027_522212290_27111&partName=htmltext)
- CXL - Open Extensible Concept Map File Format - jezik baziran na XML-u za opisivanje koncept mapa (Import Cmap from CXL); primjer <http://cmap.ihmc.us/xml/PlantsSimple.cxl> koji prikazuje <http://cmap.ihmc.us/xml/img/PlantsSimple.jpg>
- XTM - Topic Maps

### Topic Maps

- **standard** for the representation and interchange of knowledge, with an emphasis on the **findability** of information
- topic Maps are similar to concept maps and mind maps in many respects, though only Topic Maps are standardized
- A topic map represents information using
  - **topics**, representing any concept, from people, countries, and organizations to software modules, individual files, and events,
  - **associations**, representing hypergraph relationships between topics
  - **occurrences** representing information resources relevant to a particular topic



- topic maps make information findable by giving every concept in the information its own identity and providing multiple redundant navigation paths through the information space
- these paths are semantic, and all points on the way are clearly identified with names and types that tell you what they are
- Charles Goldfarb called topic maps “the GPS of the information universe”
- topic maps also help by making it possible to relate together information that comes from different sources through merging and published subjects
- XML vs. RDF vs. XTM: use XML for interchange and document contents, RDF for fine-grained metadata, and topic maps for making information findable and anything that is mostly about relationships
- alat za kreiranje, prikaz i pretraživanje topic mapa The Omnigator, dostupan na <http://www.ontopia.net/download.jsp>

## 9. tjedan (30.4-6.5.2012.)

### iMapping alat

- omogućava prikaz i povezivanje nestrukturiranog sadržaja poput osobnih bilješki i formalnog sadržaja kao što su to ontologije
- pregledan prikaz velikih mapa - hijerarhijsko gniježđenje i mogućnost uvećavanja
- moguć eksport u html format: jasan pregled - prikaz entiteta i veza
- QuiKey: pretraživanje - moguć upis samo prvih slova traženog pojma; standard i chain browsing; mogućnost davanja naziva upitima; dokumentacija [http://semanticweb.org/wiki/QuiKey/Documentation#Simple\\_Queries](http://semanticweb.org/wiki/QuiKey/Documentation#Simple_Queries):
- napravljena probna mapa i testni upiti: problem - kako napisati negaciju? poslati upit autoru
- negacija u trenutnoj verziji nije implemetirana; složeniji upiti mogući kao “chain” queries ili spremanjem upita (stavljanje znaka “?” ispred imena)

### Solvr - Ideas. Issues. Interesting Discussions.

- online alat koji omogućava jasno i jednostavno strukturiranje znanja i problema kroz strukturu sličnu stablu <http://a.freshbrain.com/solvr/>
- neslužbeni video tutorial: [http://www.youtube.com/watch?v=VU5M4jW\\_kv](http://www.youtube.com/watch?v=VU5M4jW_kv)
- moto: Problem-solving as a collaborative game
- svaki problem dobije svoj jedinstveni URL; unosi u Solvr mogu biti problemi, ideje i komentari
- problem se smatra riješenim ako ima svoju ideju
- primjer: How to represent knowledge? <http://www.a.freshbrain.com/solvr/d/kpcijla36b>
- primjer: How to make money with Solvr? <http://a.freshbrain.com/solvr//d/MakeMoneyWithSolvr>
- sličan alat, više orijentiran na brainstorming i dijeljenje ideja <http://wridea.com/>

### Korištenje konceptnih mapa za testiranje učenikova znanja

- testiranje kako učenici primjenjuju koncepte i koliko je duboko njihovo razumijevanje

- powerful learning and evaluation tool
  - specifications for concept-mapping assessments usually include:
1. defining the task demands in terms of a given set of concepts or terms within a content area and electronic information links that are provided to the student in a finite database
  2. scoring the concept maps using the Herl metric for scoring semantic content, organizational structure, number of links used in the map
  3. comparison with expert teacher maps that serve as criterion maps

#### Članak: AN APPROACH TO COMPARISON OF CONCEPT MAPS REPRESENTED BY GRAPHS

- predlaže automatiziranu usporedbu konceptnih mapa za različite primjene, ali prvenstveno u obrazovanju
  - ideja je zasnovana na sličnosti pojmovna konceptna mapa i graf: entiteti predstavljaju vrhove u grafu, a veze predstavljaju bridove
  - u matematičkim terminima je problem poznat pod nazivom The Graph Matching Problem tj, Graph Isomorphism Problem (pronalaženje sličnosti između grafova iste veličine - isti broj vrhova i bridova)
  - primjena na konceptne mape:
1. Ekstrakcija entiteta i veza i semantička usporedba korištenjem algoritama za obradu prirodnog jezika poput korjenovanja riječi i pronalaženja sinonima iz rječnika
  2. Gradnja matrica sličnosti za entite i veze (numeričke vrijednosti)
  3. Rješavanje funkcije heurističkim algoritmom GRASP (Greedy Randomized Adaptive Search Procedure) koji daje dovoljno dobre rezultate u dovoljno kratkom vremenu
- eksperimentalni rezultati su pokazali uspješnost algoritma (u gotovo 90% slučajeva je otkirvena visoka sličnost testnih primjera)

#### Inspiration alat

- dostupan na: <http://www.inspiration.com/Inspiration>
- plaćeni alat koji omogućuje strukturiranje znanja kroz concept i mind mape, pisanje zabilješki, planiranje, pisanje prezentacija
- moguć eksport u pdf, doc, html, ppt
- jednostavno generiranje prezentacija
- pristup rječniku i velikom broju grafičkih simbola
- moguć rad po templateima: od planiranja učenja i istraživanja do pisanja CV-a i prezentacija

#### **Concept Map of Knowledge Representation**



more weight to propositions that were considered to be critical by experts

- Concept Maps can be used in *formative* or *summative* assessment procedures
  - formacijska ocjena: ocjena tijekom učenja - učitelj može vidjeti koliko učenik razumije, ali i dobiti ocjenu svojih nastavnih materijala i metoda te ih prilagoditi na odgovarajući način
  - ukupna (zbrojna) ocjena: ocjena po završetku lekcije - učitelj može provjeriti učenikovo znanje i razumijevanje te mu dodijeliti odgovarajuću ocjenu
  - izvori: IHMC Literature Review on Concept Mapping i Propositional Analysis Model to the Comparison of Expert Teachers' Concept Maps
1. <http://www.ihmc.us/users/acanas/Publications/ConceptMapLitReview/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf>
  2. <http://cmc.ihmc.us/cmc2008papers/cmc2008-p182.pdf>

### Concept Map of Knowledge Representation

- ispravljena i nadopunjena verzija, dostupna na:

<http://cmapspublic.ihmc.us/rid=1KLNWVFB7-148JDV6-34RV/Knowledge%20Representation.cmap>

### 11. tjedan (14.5-21.5.2012.)

- decision trees in decision analysis and decision trees in artificial intelligence (machine learning, data mining)
- decision trees for graphical representations of IF-THEN rules
- expert systems for solving problems
- decision trees vs. expert systems: decision tree present graphical knowledge base while expert systems give answers based on interaction with the user, knowledge base and inference engine
- good tutorial on decision trees: [http://dms.irb.hr/tutorial/tut\\_dtrees.php](http://dms.irb.hr/tutorial/tut_dtrees.php)

Decision tree representation:

- Each internal node tests an attribute
- Each branch corresponds to attribute value
- Each leaf node assigns a classification

When to use decision trees:

- Instances describable by attribute-value pairs
- Target function is discrete valued
- Disjunctive hypothesis may be required
- Possibly noisy training data

Examples of application

- Equipment or medical diagnosis
- Credit risk analysis
- Modeling calendar scheduling preferences

## WEKA - Data Mining Tool for Machine Learning

- contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality
- Weka supports several standard data mining tasks: data preprocessing, clustering, classification, regression, visualization, and feature selection
- the data is available as a single flat file or relation, where each data point is described by a fixed number of attributes (normally, numeric or nominal attributes, but some other attribute types are also supported)

## SmartDraw

- komercijalni alat za izradu različitih vrsta vizualizacija, među ostalima i za izradu stabla odlučivanja

## **Izveštaj sa sastanka, četvrtak, 17.5.2012, 10:00**

- rasprava o stablima odlučivanja i data miningu
  - plan za dalje:
1. istražiti detaljnije stabla odlučivanja u području analize odlučivanja (engl. decision analysis); istražiti za koja područja, prednosti i nedostatke; dostupne alate
  2. unaprijediti konceptnu mapu istraživanja

## **12. tjedan (21.5-28.5.2012.)**

### Decision trees in decision analysis

#### **Definition:**

A decision making tree is essentially a diagram that represents, in a specially organized way, the decisions, the main external or other events that introduce uncertainty, as well as possible outcomes of all those decisions and events.

#### **Components of a decision tree**

1. Decision node: a square node; represents decisions you can make
2. Chance (uncertainty) node: a circle node; shows the occurrence of events over which the decision maker has no direct control
3. Consequences (outcomes: utilities or costs)



### Process of building a decision tree

1. Interview decision makers and construct a preliminary tree
2. Present the tree and show how various concerns are captured
3. Solicit a list of new concerns
4. Revise the tree

### Applications:

Decision trees are used in the field of operational research, a discipline that deals with the application of advanced analytical methods to help make better decisions. Some examples of the fields where it is used:

- business and economics
- health economics
- public health
- law suits

From: <http://studentski-izvjestaji.zesoi.fer.hr/> - Studentski izvještaji

Permanent link: [http://studentski-izvjestaji.zesoi.fer.hr/doku.php?id=studenti:mihita\\_cvitanovic:mc\\_start&rev=1337723997](http://studentski-izvjestaji.zesoi.fer.hr/doku.php?id=studenti:mihita_cvitanovic:mc_start&rev=1337723997)

Last update: 2023/06/19 16:20

