

Knowledge engineering – modeling activity. Concept modeling in developing knowledge based systems

Drd. Sabina MIHALACHE

Facultatea de Economie și Administrarea Afacerilor, Universitatea Al. I. Cuza, Iași

In this article we present the modeling activity for knowledge-based systems development. We also examine some of its basic characteristics, such as the separation between domain knowledge and inference knowledge, which specifies the reasoning steps followed to solve the task. The mainly aspect is represented by knowledge model.

Keywords: conceptual modeling, knowledge level, task modeling, PSM.

Introducere

În mod obișnuit sistemele bazate pe cunoștințe sunt utilizate la rezolvarea problemelor complexe sau a celor care nu sunt pe deplin înțelese, astfel încât funcționalitatea sistemului nu poate fi decât parțial specificată. Conceptul *funcții incomplet specificate* este utilizat la caracterizarea unei probleme posibile de rezolvat cu ajutorul unui sistem bazat pe cunoștințe.

Procesul de dezvoltare al sistemelor bazate pe cunoștințe era văzut la început ca un proces de transfer al cunoașterii de la expert la calculator. Astăzi procesul de dezvoltare al sistemelor bazate pe cunoștințe este văzut ca o *activitate de modelare*. A construi¹ un sistem bazat pe cunoștințe înseamnă a construi un model informatic care să imite cât mai bine expertul din domeniu. Procesul de achiziție a cunoașterii nu mai este văzut ca un proces de transfer ci ca un *model al cunoașterii* realizat de un observator exterior.²

¹ O problemă centrală care trebuie rezolvată în cadrul procesului de dezvoltare al sistemelor bazate pe cunoștințe este *achiziția și reprezentarea cunoașterii*. Această activitate este în mod obișnuit denumită **ingineria cunoașterii**, deși unii autori folosesc acest termen pentru desemnarea întregului proces de dezvoltare al sistemelor bazate pe cunoștințe. În momentul de față este general acceptat faptul că achiziția cunoașterii constă într-o sarcină constructivă de modelare la nivele diferite, cu accentul pe analiza și înțelegerea naturii sarcinii, posibilitatea descompunerii în unul sau mai multe PSM ca subsarcini, până când se ajunge la nivelul primitivelor inferențiale unde este nevoie să specificăm rolurile de intrare și de ieșire.

² Prima sarcină a observatorului extern este aceea de a modela cunoașterea pentru a o include în sistem și apoi rescrierea acestor modele ținând cont de limitările

Metodele de construire a sistemelor bazate pe cunoștințe oferă în mod obișnuit instrumente pentru analiza cunoașterii sub forma așa-numitelor *modele conceptuale ale cunoașterii* sau mai simplu modele de cunoaștere.

Un model de cunoaștere oferă o specificare de implementare a cunoașterii într-un domeniu aplicativ. În mod obișnuit, un model de cunoaștere oferă formate atât pentru scrierea cunoașterii statice a domeniului (reguli, clase, relații) cât și a strategiilor de raționament în care acest domeniu al cunoașterii este folosit pentru a rezolva o problemă particulară. Un important aspect al ingineriei cunoașterii este acela că oferă modele predefinite și reutilizabile pentru anumite sarcini de cunoaștere intensivă numite de asemenea metode de rezolvare a problemelor (PSM: problem-solving method).

Tendințe actuale în dezvoltarea sistemelor bazate pe cunoștințe

În momentul de față sunt trei dimensiuni în care dezvoltarea sistemelor bazate pe cunoștințe avansează:

1. accentul pe metodologie. Perspectiva metodologică care a contribuit cel mai mult la apropierea dezvoltării sistemelor bazate pe cunoștințe de alte discipline ingineresti se bazează pe:

impuse de nivelul de formalizare și de cunoașterea pierdută în timpul operaționalizării inferențelor și a procedurilor de programare. Această cunoaștere pierdută trebuie injectată în interpretarea funcțiilor sistemelor bazate pe cunoștințe. Nivelul cunoașterii va conține întotdeauna mai multă cunoaștere decât nivelul formalizării. Numai modelele formalizate trec în calculator (figura nr.1).

- introducerea modelării conceptuale pentru nivelul cunoașterii³ potrivit Newell și acceptarea limbajului natural ca o primară sursă de informații de unde sunt extrase entitățile oricărui model al cunoașterii.
 - căutarea abstracțiilor recurente care pot fi detașate de sarcini și de domeniul aplicativ al problemei și care pot fi reutilizate fie direct, fie în urma unor mici modificări.
 - taxonomia cunoașterii pe două straturi.
2. reutilizarea componentelor din ontologii și librăriile de PSMs.
 3. arhitecturile hibride care integrează algoritmi software și hardware, PSMs, inferențe și operatori formali.

Context și temă

Importanța modelării cunoașterii pentru managementul cunoașterii a fost subliniată de Wielinga în 1997. Modelarea conceptuală este o activitate centrală în ingineria cunoașterii. Modelele sunt utilizate pentru a capta aspectele esențiale ale sistemelor reale descompunându-le în părți controlabile care sunt mai ușor de înțeles și manipulat. Modelele sunt foarte mult asociate cu domeniul pe care îl reprezintă.

“Un model este o simplificare a realității”. Modelele sunt folosite în activitățile de dezvoltare ale sistemelor la trasarea granițelor sistemului și la comunicarea dintre membrii echipei de dezvoltare. Aceștia au viziuni diferite asupra sistemului iar modelele îi pot ajuta să înțeleagă aceste viziuni într-o manieră unificată.

În urma procesului de modelare rezultă modele conceptuale ale activităților de cunoaștere intensivă. Pe parcursul achiziției cunoașterii majoritatea cunoașterii este nestructurată și tacită. Cognoscianul va încerca să înțeleagă atât cunoașterea explicită cât și pe cea tacită, iar apoi va utiliza simple diagrame vizuale pentru a stimula discuțiile dintre utilizatori și experți. Cognoscianul va trebui să construiască modelul conceptual pe baza dis-

cuțiilor din faza de achiziție a cunoașterii. Acesta va comunica cunoașterea specialistului în informatică care va transforma modelul în instrucțiuni de cod sau programe.

Cea mai importantă metodologie de dezvoltare a sistemelor bazate pe cunoștințe este metodologia CommonKADS. Aceasta permite construirea unei colecții de modele, unde fiecare model captează aspecte specifice ale sistemului pentru a fi dezvoltate. Se disting astfel: *modelul organizației*, *modelul sarcinii*, *modelul agentului*, *modelul comunicației*, *modelul cunoașterii* și *modelul de proiectare*; primele patru servesc la modelarea mediului organizațional în care sistemul bazat pe cunoștințe va opera în concordanță cu sarcinile organizației; modelul cunoașterii și modelul proiectului descriu aspectele funcționale ale sistemului bazat pe cunoștințe.

Modelarea organizației se preocupă cu modelarea proceselor afacerii în care suntem interesați din punctul de vedere al cunoașterii. Motivul construirii modelelor organizației este acela de a oferi o descriere a întregului context în care modelul cunoașterii trebuie să funcționeze.

Modelul contextului sistemului descrie mediul organizațional direct al unui sistem software. Acesta este folosit în dezvoltarea softului pentru a arăta modul în care acesta interacționează cu mediul său.

Acesta este proiectat cu ajutorul diagramei UML a cazurilor de utilizare (use-case diagram). De asemenea poate include:

- o diagramă UML de stare sau o diagramă UML a activităților pentru controlul interacțiunilor dintre actorii externi și sistem, cu alte cuvinte o diagramă care să prezinte timing-ul interacțiunilor dintre sistem și mediul său extern;
- o diagramă a activităților care să arate o fluxurile de informații din organizație;
- o diagramă a claselor, pentru a descrie structura statică a informațiilor schimbate între actorii externi și sistem.

Trebuie realizate două instanțe ale modelului organizației: unul care să analizeze organizația în momentul prezent și unul al perspectivei dorite. Dezvoltarea sistemelor bazate pe cunoștințe pe baza modelării are urmă-

³ Alan Newell a introdus principiul nivelului cunoașterii. El afirmă următoarele: cunoașterea trebuie modelată la nivel conceptual, urmată fiind de specificațiile de proiectare și implementarea propriu-zisă. Clancey a fost cel care a subliniat importanța acestui principiu

toarele consecințe:

- ca orice model, el nu poate reflecta realitatea în general, el poate aproxima un comportament așteptat. În principiu procesul de modelare este infinit;
- procesul de modelare este un proces ciclic. Noi observații pot conduce la rafinarea,

modificarea, completarea modelului deja construit;

- procesul de modelare este dependent de interpretările observatorului. Este necesară obținerea de feedback-uri repetate până la obținerea unui model adecvat.

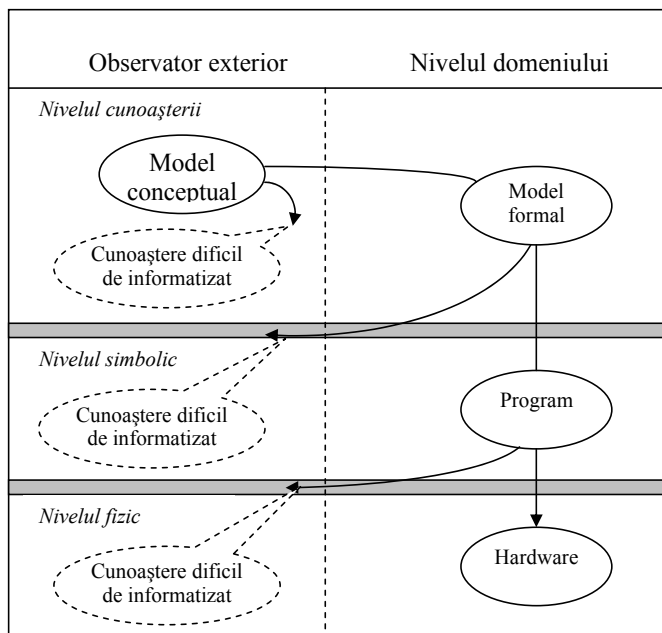


Fig. 1. Locul observatorului exterior în modelarea cunoașterii

Sursa: Knowledge-Based Systems, Jose-Mira, Ana E. Delgado-Garcia, în Artificial Intelligence: Technology with a Future, vol. III, nr.5, octombrie 2002 la <http://www.upgrade-cepis.org/issues/2002/5/up3-5Mira.pdf>

Modelarea cunoașterii

Ce este cunoașterea? Nu este nici materie, nici energie. Este o formă pură. În momentul de față știm de existența sa doar sub forma metodelor de rezolvare a problemelor științifico-tehnice de către experții umani și că acești experți umani posedă cunoașterea. Suntem observatori ai acestui fenomen "a cunoaște" și al achiziției cunoașterii prin construirea așa numitului model conceptual al cunoașterii format din entități și relațiile dintre entități, pe care se presupune că experții umani le folosesc în raționamentele lor. Metodologia CommonKADS distribuie aceste componente ale cunoașterii pe două straturi (figura 2).

Stratul sarcinilor, metodelor (PSMs) și inferențelor

Acest strat joacă un rol echivalent cu cel al circuitelor integrate la nivelul componentelor hardware. Sarcina de cunoaștere oferă o descriere asupra scopurilor sistemului. PSMs

specifică modul de descompunere a acestei sarcini până la nivelul primitivelor inferențiale. Inferența descrie pașii de bază ai raționamentelor prin care PSM descompune o sarcină (cover, select, establish, refine). Descrierea acestor inferențe este completată cu specificarea rolurilor de intrare și ieșire jucate de entitățile domeniului cunoașterii. Astăzi, există chiar librării de PSMs pentru diferite sarcini cum ar fi diagnosticul sau planificarea.

Sarcinile⁴ sunt descompuse până când se ajunge la inferențe de bază care nu sunt des-

⁴ Potrivit metodologiei CommonKADS sarcinile de cunoaștere intensivă se împart în sarcini analitice (clasificarea, diagnosticul, predicția, monitorizarea și evaluarea) și sarcini sintetice (proiectarea, planificarea, desemnarea, modelarea). Pentru fiecare tip de sarcini pot fi găsite în literatură modele de cunoaștere. Ideea constă în descompunerea de către analistul de sistem a sarcinii până la nivelul la care poate identifica asemenea sarcini.

compuse mai departe. Rezultă că o sarcină este compusă dintr-un număr de inferențe combinate într-o diagramă de inferențe (figura 3).

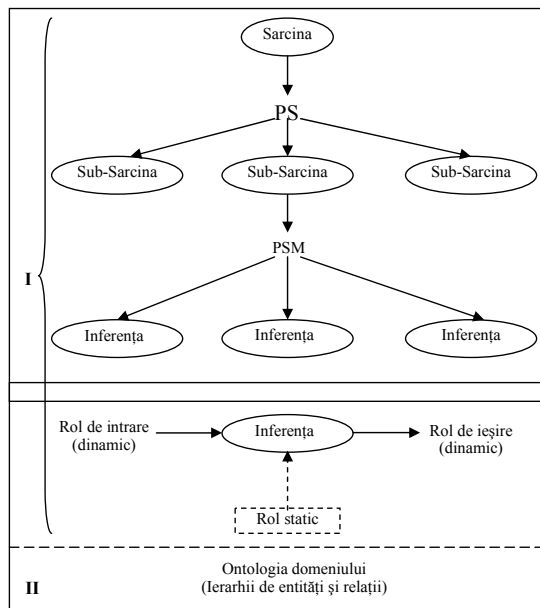


Fig. 2. Modelarea cunoașterii la nivelul cunoașterii: (I) Sarcini, PSMs și inferențe (II) Stratul domeniului

Sursa: Knowledge-Based Systems, Jose-Mira, Ana E. Delgado-Garcia, în Artificial Intelligence: Technology with a Future, vol.III, nr.5, octombrie 2002 la <http://www.upgrade-cepis.org/issues/2002/5/up3-5Mira.pdf>

Rolurile dinamice de cunoaștere sunt intrări și ieșiri de cunoaștere pentru inferențe. Rolurile statice de cunoaștere specifică colecții ale domeniului cunoașterii care sunt folosite pentru a face o inferență. Librării de sarcini și inferențe reutilizabile sunt disponibile sub forma PSM.

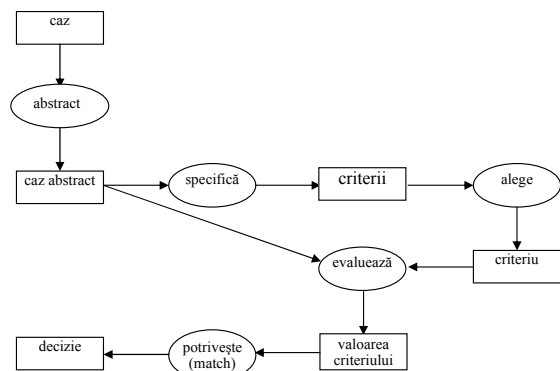


Fig. 3. Diagrama de inferențe generice pentru sarcina de evaluare

Sursa: Conceptual Modelling for Knowledge-Based Systems, Speel, P-H, Schreiber, A.Th., în Encyclopedia of Computer Science and Technology la <http://www.cs.vu.nl/~guus/papers/Speel01a.pdf>

Un important aspect pentru interpretarea corectă a diagramei anterioare este cel al controlului. De observat că diagrama nu specifică ordinea în care au loc acțiunile. Diagrama de activități poate fi folosită pentru a defini controlul întreg al structurii (figura 4).

Stratul domeniului cunoașterii

Cunoașterea din domeniu este statică și descrie ontologia domeniului problemei. Ea poate fi împărțită în module, fiecare reprezentând un punct de vedere particular, acestea fiind denumite în această metodologie „modelele domeniului”: modelul structural, modelul cauzal și modelul comportamental. O descriere a domeniului cunoașterii poate fi comparată cu modelul datelor sau modelul claselor din ingineria programării.

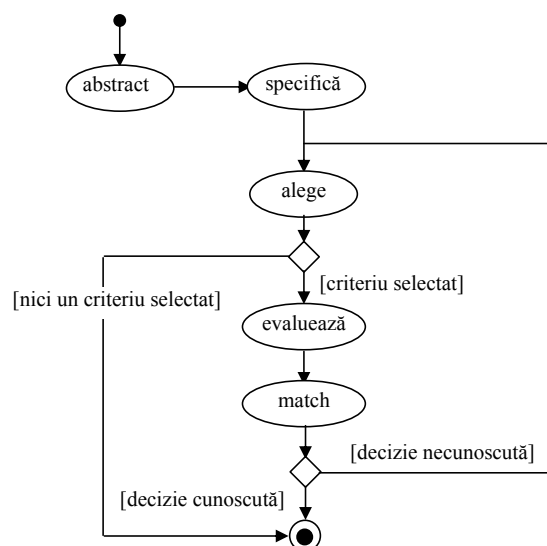


Fig. 4. Diagrama de activități care definește controlul derulării inferențelor

Sursa: Conceptual Modelling for Knowledge-Based Systems, Speel, P-H, Schreiber, A.Th., în Encyclopedia of Computer Science and Technology la <http://www.cs.vu.nl/~guus/papers/Speel01a.pdf>

În ingineria cunoașterii termenul ontologie este utilizat pentru schema care captează vocabularul specific al unui domeniu. În afara ontologiei trebuie definite un număr de baze de cunoștințe diferite. În primii ani de dezvoltare ai sistemelor expert era construită o singură bază de cunoștințe, ceea ce s-a dovedit greu de menținut și dezvoltat. Astăzi sunt dezvoltate baze de cunoștințe diferite pentru fiecare tip de reguli.

Acest strat modelează seturi de entități și re-

lații specifice unui domeniu aplicativ (medicină, economie) și sunt independente de folosirea lor în inferențe. Aceste structuri sunt în mod obișnuit instanțiate din ontologii și includ cunoaștere statică și cunoaștere despre fapte, atât de natură cantitativă cât și de natură calitativă. De exemplu ierarhii de concepte (frames) și relații de cauzalitate între aceste concepte. Aceste structuri de date vor juca un rol echivalent cu cel al semnalelor electrice dintr-un circuit electronic.

Concluzii

Modelele sunt utilizate în dezvoltarea sistemelor inclusiv a sistemelor bazate pe cunoștințe. Există variate tehnici de modelare a cunoașterii utilizate pentru a modela cunoașterea, dar astăzi se observă încercările dezvoltatorilor de sisteme de a unifica aceste tehnici.

Modelarea conceptuală transformă cerințele utilizatorului într-un format accesibil nivelului de proiectare logică, definind problema economică și funcțiile necesare pentru rezolvarea ei. Etapa presupune definirea unui număr mare de scenarii economice sau scenarii de utilizare, destinate identificării cât mai complete a viziunii clientului precum și a funcționalității cerute.

Această fază reprezintă punctul central al modelării economice și nu tratează tehnologiile necesare implementării sistemului.

Bibliografie

1. Abdullah, M.S., Benest, I., Evans, A., Kimble, C., *Knowledge Modelling Techniques For Developing Knowledge Management Systems*, 3rd European Conference on Knowledge Management, Dublin, Ireland, September 2002, pp.15-25 aflat la <http://www.cs.york.ac.uk/mis/docs/ECKM2002.pdf>;
 2. Chandrasekaran, B., Johnson, T.R., Smith, J.W., *Task-Structure Analysis for Knowledge Modeling*, aflat la <http://www.cse.ohio-state.edu/~chandra/task-structure-cacm.PDF>;
 3. Mira, J., Delgado-Garcia, A., *Knowledge-Based Systems*, în *Artificial Intelligence: Technology with a Future*, vol.III, nr.5, octombrie 2002 la <http://www.upgrade-cepis.org/issues/2002/5/up3-5Mira.pdf>;
 4. Rudi, S., et.al., *Situation and Perspective of Knowledge Engineering*, aflat la www.db.stanford.edu/~stefan/paper/2000/ios_2000.pdf;
 5. Schreiber, G., *Pragmatics of the Knowledge Level- PhD thesis*, aflat la <http://www.cs.vu.nl/~guus/papers/Schreiber92c.pdf>;
- Speel, P-H, Schreiber, A.Th., *Conceptual Modelling for Knowledge-Based systems*, în *Encyclopedia of Computer Science and Technology* la <http://www.cs.vu.nl/~guus/papers/Speel01a.pdf>.