KNOWLEDGE REPRESENTATION SYSTEM FOR DECISION SUPPORT*

D. Buchtela, J. Brožek

Czech University of Life Sciences, Faculty of Economics and Management, Department of Information Engineering, Prague, Czech Republic

The aim of this article is to present a design of a Knowledge Representation System (KRES) for decision support. The system automatically offers relevant formalized knowledge by extended GLIF (Guidelines Interchange Format) models to participants on the basis of acquired data. This selection algorithm is based on key attributes and cooperation with knowledge ontologies or other registry of recognized attributes.

Knowledge Representation; GLIF model; ontology

INTRODUCTION

Nowadays the conceptual knowledge (concepts, targets, actions, …) plays more prominent role. Conceptual knowledge is related to given task but expressed on higher level of abstraction and relatively independent of the form which is used by a knowledge system or expert. These factors are related to the creation and development of structured methodologies which support model based knowledge mining and using

The aim of this article is to present a design of a Knowledge Representation System which is based on GLIF models of branch knowledge. The GLIF models are extended by key attributes. The KRES offers the participants (patient, physician, operator,…) relevant model(s) based on acquired data.

MATERIAL AND METHODS

GLIF model

The GLIF (*Guideline Interchange Format*) model is a result of collaboration among Columbia University, Harvard University, McGill University and Stanford University. The main goal of GLIF was to enable sharing of guidelines among institutions and across computer applications (Ohno-Machado et al., 1998).

GLIF specifies a process-oriented model for guidelines representation and syntax for guidelines utilization in software systems as well as for their transport. GLIF guidelines are mostly given as a flowchart representing a temporarily ordered sequence of steps. The nodes of the graph are guideline steps and edges represent continuation from one step to the other one (see Fig. 1). Guideline steps are an *action step*, *decision step*, *branch and synchronization steps* and a *patient state step.*

- **Action steps** specify clinical actions that are to be performed. It can be an application of some therapy, carrying out some examination or measurement etc. Action step also may name sub-guidelines (subgraph), which provide a detail for the action.
- **Decision steps** are used for conditional branching. There are two kinds of decision steps: **Case step** is used, when branching is determined by evaluation of defined logical criteria based on data items. **Choice step** is used when the decision cannot be precisely specified in guidelines themselves and decision should be made by the user.
- **Branch** and **synchronization steps** enable concurrence in the model. Guideline steps that follow branch step can be performed concurrently. Branches with root in branch step eventually converge in a synchronization step. In this step all branches are synchronized after evaluation of synchronizing condition.
- **State step** characterizes a surveyed object state after execution of the previous step or in the beginning of the model.

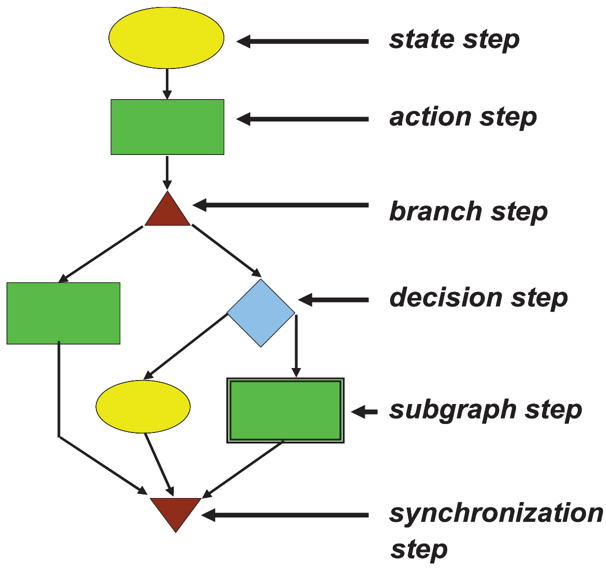

Fig. 1. Graphical symbols of GLIF model

* The paper was supported by the grant projects No. 2C06004 and No. 1M06014.

Criteria of conditions

The decision step specifies several criteria of condition for each decision option:

- The strict-in criterion is used to specify a decision condition that could be computed automatically (for example if systolic blood pressure is 130 or higher). If a strict-in is true then the control flows to the guideline step that is specified by that decision option's destination.
- The strict-out criterion is analogous to an absolute contraindication (for example if a patient is gouty he could not be cured by thiazides diuretics). If a strict-out is true then the decision option's destination is forbidden.
- The rule-in criteria rank a choice as the best among several options. For example, when there are competing diagnoses for a disease, a pathognomonic condition would be a rule-in for the disease. This criterion is analogous to conditions favouring the use (indications).
- A rule-out takes precedence over rule-in when ranking options. If an option contains both a rule-in criterion and a rule-out criterion, and both are evaluated as true, then that option should be the last choice. This criterion is analogous to contraindications.

The strict-out criterion is evaluated at first. If strict-out criterion is evaluated as true the rest of the criteria is not evaluated. This option is forbidden. In the opposite case the strict-in criterion is evaluated. If the strict-in criterion is false too, the rule-in and rule-out criteria are evaluated. The ranking of rule-ins and rule-outs is left to the users who may use their clinical judgement or may develop their own ranking schemes.

GLIF implementation

GLIF model is graphical so it is necessary to code it in a formal language. These formats are available for implementation of GLIF model and its parameters:

- *GELLO* (Guideline Expression Language) – the object oriented query and expression language for decision support (Ohno-Machado et al., 1998).
- *RDF* (Resource description framework) – the built-in form of saved knowledge base in a Protége system (Gennari et al., 2003). The Protége is a development environment used for knowledge system design. The environment enables knowledge ontology construction, data inserting and user formatting of input and output forms.
- *XML* (eXtensible Markup Language) – The encoded model consists of a sequence of guideline steps. Some attributes of a guideline step contain next guideline steps. It enables sequential representation of a graph structure in the guideline language.

Knowledge ontology

Knowledge ontologies follow the development in an artificial intelligence in an area of knowledge representation. Ontologies are thoroughly taken in logic theories and their link to real objects (instances) is relatively open. The ontology is primarily used to describe concepts (classes) and not to describe facts about concrete objects. Classes (concepts) and relations are systematically defined by a formal language (Uschold, Gruninger, 1996).

The *class* is the base of the knowledge ontology and describes a set of objects. In some formalisms the *class* corresponds to a *concept* or a *category* and closely corresponds to a *frame* which is the base construction of any artificial intelligence systems. The ontology classes do not contain procedural methods in opposite to classes in object-oriented models.

On a set of classes there is defined a hierarchy (taxonomy). The philosophical view of ontology sometimes requires a strict tree structure but all of the main ontology languages support multiple heredity.

The *individuum* corresponds to a concrete object of the real world and is the opposite of the class in a way. The individuum can be inserted into the ontology without a link to any class.

RESULTS

Knowledge model

Knowledge Representation Model (KREM) of the whole system is based on a branch knowledge formalisation trough the use of the GLIF model. The formalisation process, i.e. construction of the graphic GLIF model of knowledge contained in free text and the model coding into the formal language (XML) (Buchtela et al., 2005), is illustrated in Fig. 2.

In the stage of a GLIF model construction from a free text it is important to find a logical and process structure of the knowledge, all fundamental parameters and their interrelationships. The result of this stage is a graphic GLIF model corresponding to the knowledge in the text. The construction stage is the most important and difficult of all stages.

In the stage of GLIF model coding the graphic model of knowledge is coded into XML. Some steps of the encoded model contain description of next steps (next options). It enables sequential representation of a graph structure (see Fig. 3).

The resulting GLIF model is extended by key attributes. The list of the key attributes is in the Table 1. These key attributes are used in selection algorithm and they are coded in XML along with the GLIF model.

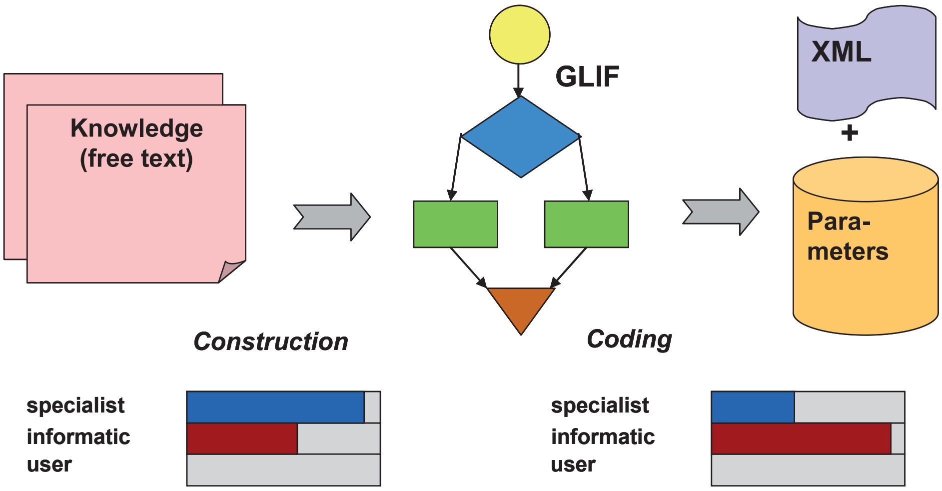

Fig. 2. Process of GLIF model construction and implementation

Fig. 3. XML elements of GLIF model and next option

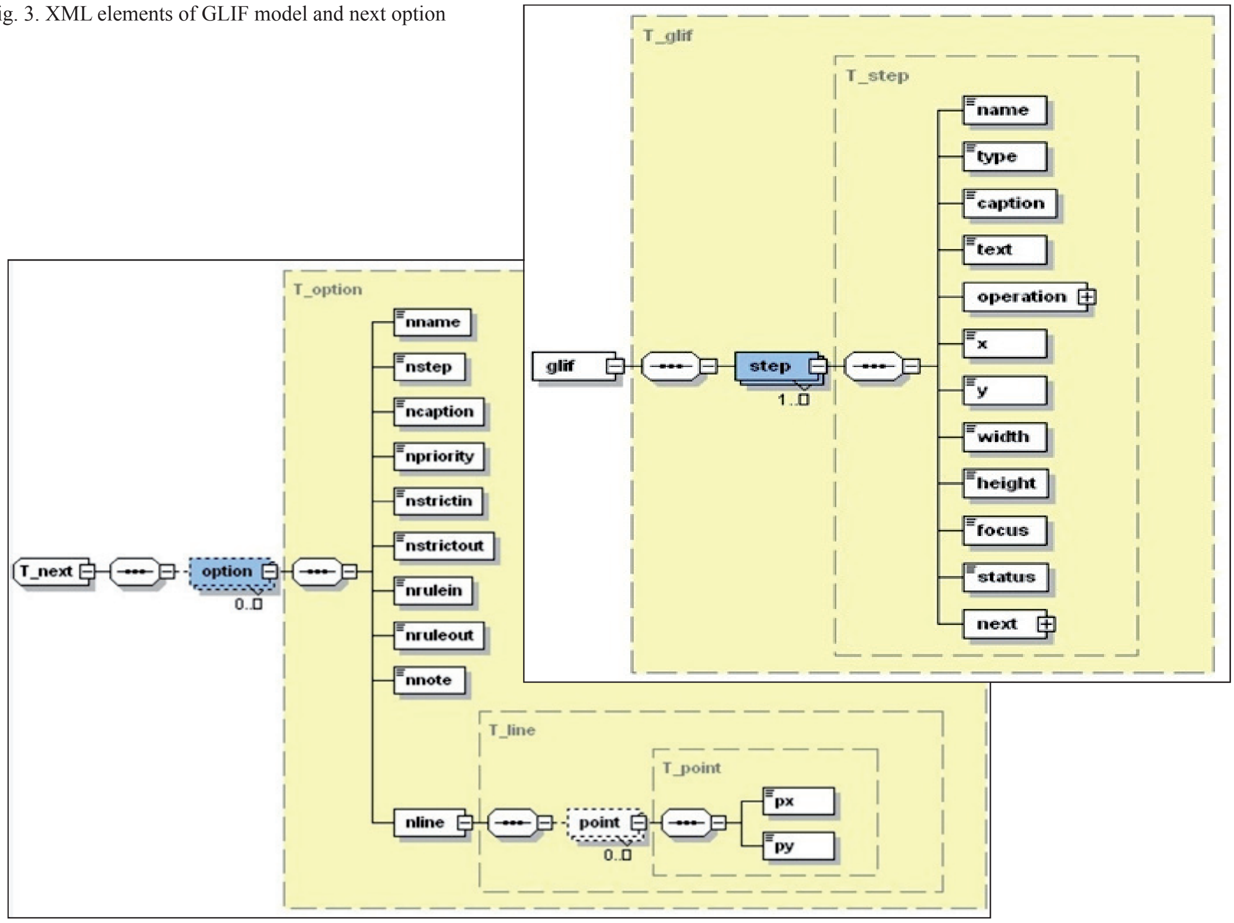

Table 1. Key attributes of knowledge model

Key attribute	Attribute description
branch	Branch described by the GLIF model – e.g. cardiology
BID	Branch identification – e.g. International Classification of Diseases (ICD)
user	User of system to whom the GLIF model is primary determined (patient, physician, operator, …)
status	The GLIF model validity
key	List of keys described by the GLIF model – e.g. blood pressure, diabetes …
key_name	Name of the key
key_weight	Weight of the key – the GLIF model description rate of the key

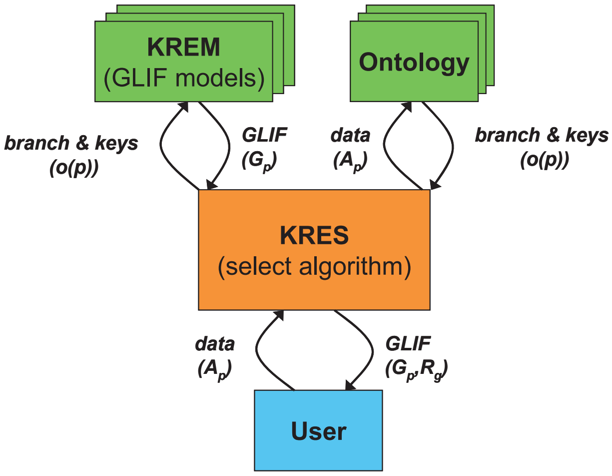

Fig. 4. Knowledge Representation System

Selection algorithm

The principle of the knowledge representation system (KRES) and the algorithm of the relevant GLIF model selection is illustrated in Fig. 4.

The selection algorithm can be described subsequently:

- For every specific participant (user, patient, physician, operator, …) p and his attributes A_p a set of all branches (areas) exp(B) which corresponds to participant state (attributes A_p) is determined. This function o is determined using the knowledge ontology or other registry of recognized attributes:

  $o: A_p \rightarrow \exp(B)$,

  where B is a finite set of all branches.
- For each branch $B \in o(p)$ a subset $K(p, B)$ of recognized attributes (keys) of the branch B and affected by attributes of the participant p is determined.
- Models G_p are chosen from a finite set of all GLIF models M thus:

  $$G_p = \{g \in M : g \in K(p, B) \cap \\ \cap\, key\_weight(branch(g), P(g)) \neq 0\},$$

  where each GLIF model contains attribute *branch* and a finite set P of keys (attribute *key_name*) and their weights (attribute *key_weight*).
- For each model G_p a general aggregate operator r is defined :

  $r: \{G_p\} \rightarrow \Re$,

  where $\Re$ is the set of real numbers.

  The operator r can be for example defined as $R_g = \sum_{k \in P(g)} key\_weight(k)$. The participant is then offered the GLIF model with the highest relevance value $R = \max(R_g)$ or a list of models ordered by the value of R_g.

CONCLUSION

Designed knowledge representation system is based on the GLIF model which is the universal method of modelling of mainly procedural-oriented knowledge. The system offers a relevant GLIF model or an ordered list of models from a list of available formal models on the basis of participant attributes. Additional information describing the concrete situation gives the participant possibility of better decision.

REFERENCES

BUCHTELA, D. – PELESKA, J. – VESELY, A. – ZVAROVA, J.: Method of GLIF model Construction and Implementation. In: Proc. XIX Int. Congr. of the European Federation for Medical Informatics, 2005.

GENNARI, J. H. – MUSEN, M. A. – FERGESON, R. W. – GROSSO, W. E. – CRUBEZY, M. – ERIKSSON, H. – NOY, N. F. – TU, S. W.: The evolution of Protégé: an environment for knowledge-based systems development. Int. J. Hum.-Comput. St., *58*, 2003 (1).

OHNO-MACHADO, L. – GENNARI, J. H. – MURPHY, S. N. – JAIN, N. L. – TU, S. W. – OLIVER, D. et al.: The GuideLine Interchange Format: A model for representing guidelines. J. Am. Med. Inform. Assn., 1998.

USCHOLD M., GRUNINGER M.: Ontologies: principles, methods and applications. Knowl. Engin. Rev., *11*, 1996: 2.

Received for publication on January 14, 2008
Accepted for publication on March 3, 2008

BUCHTELA, D. – BROŽEK, J. (Česká zemědělská univerzita, Fakulta provozně ekonomická, katedra informačního inženýrství, Praha, Česká republika):

Systém reprezentace znalostí pro podporu rozhodování.

Scientia Agric. Bohem., *39*, 2008: 97–101.

Cílem tohoto příspěvku je návrh systému reprezentace znalostí (KRES) pro podporu rozhodování. Systém nabízí účastníkům systému relevantní GLIF (Guidelines Interchange Format) model formalizovaných znalostí na základě konkrétních získaných dat. Algoritmus výběru je založen na klíčových atributech a spolupráci se znalostními ontologiemi nebo registrem uznávaných atributů (termínů).

Znalostní model (KREM – Knowledge Representation Model) celého systému je založen na oborových znalostech formalizovaných pomocí GLIF modelu. GLIF model je nejčastěji používaný pro přehlednou reprezentaci lékařských doporučení. Vznikl spoluprací Kolumbijské univerzity, Harvardské, McGillovy a Stanfordské univerzity. GLIF poskytuje objektově a procesně orientovaný pohled na oborová doporučení. Výsledným modelem je orientovaný graf skládající se z pěti hlavních částí (kroků): *Akce*, *Rozhodování*, *Větvení* a *synchronizace, Stav*. GLIF model je grafický, pro další počítačové zpracování je proto nutné jej zakódovat do formálního jazyka (XML – eXtensible Markup Language).

Výsledný GLIF model je doplněn o klíčové atributy. Přehled klíčových atributů je uveden v tab. 1. Klíčové atributy se uplatní v algoritmu výběru relevantních formalizovaných znalostí účastníkům systému. Klíčové atributy (a případně další pomocné atributy) jsou zakódovány opět v jazyce XML společně s GLIF modelem.

Pro všechny dostupné GLIF modely týkající se vybrané oblasti (atributy *BID*, *branch*, *status* a *user*) je stanovena hodnota relevance R podle následujícího algoritmu:

- Pro každého pevně zvoleného uživatele (pacienta) p a jeho atributy A_p určíme, pomocí znalostní ontologie (registru uznávaných atributů), množinu všech oborů (oblastí), které se týkají jeho stavu (A_p).
- Pro každý obor B určíme konečnou podmnožinu K (p, B) uznávaných atributů (klíčů) daného oboru B ovlivněných hodnotami atributů pacienta p.
- Vybereme z konečné množiny všech GLIF modelů M modely G_p týkající se oboru B a atributů uživatele.
- Pro každý model G_p je definován agregační operátor r. Operátor r může být například definován:

$$R_g = \sum_{k \in P(g)} key\_weight(k)$$

Uživateli je pak nabídnut GLIF model s největší hodnotou relevance $R = \max(R_g)$, respektive seznam dostupných modelů uspořádaných podle hodnoty R_g.

reprezentace znalostí; GLIF model; ontologie

Contact Address:

Ing. David B u c h t e l a , Česká zemědělská univerzita v Praze, Fakulta provozně ekonomická, katedra informačního inženýrství, Kamýcká 1076, 165 21 Praha 6-Suchdol, Česká republika, tel.: +420 224 383 806, e-mail: buchtela@pef.czu.cz