

NÁVRH DATOVÉHO ROZHRAŇÍ MEZI GLIF MODELEM A INFORMAČNÍM SYSTÉMEM

DESIGN OF DATA INTERFACE BETWEEN GLIF MODEL AND INFORMATION SYSTEM

David Buchtela

Anotace:

Znalosti z různých oborů lidské činnosti je možné reprezentovat obecným GLIF (*GuideLine Interchange Format*) modelem. Pro podporu rozhodování je nezbytné tento model navázat na konkrétní data uložená v informačním systému. Tento příspěvek se zabývá implementační fází GLIF modelu a návrhu datového rozhraní mezi GLIF modelem a informačním systémem.

Klíčová slova: GLIF model, implementační fáze, datové rozhraní

Annotation:

Knowledge acquired in different branches is possible to represent in a general GLIF (*GuideLine Interchange Format*) model. For decision support it's necessary to link this model to concrete data which are stored in an information system. This article put mind to an implementation stage of the GLIF model and to design of a data interface between the GLIF model and the information system.

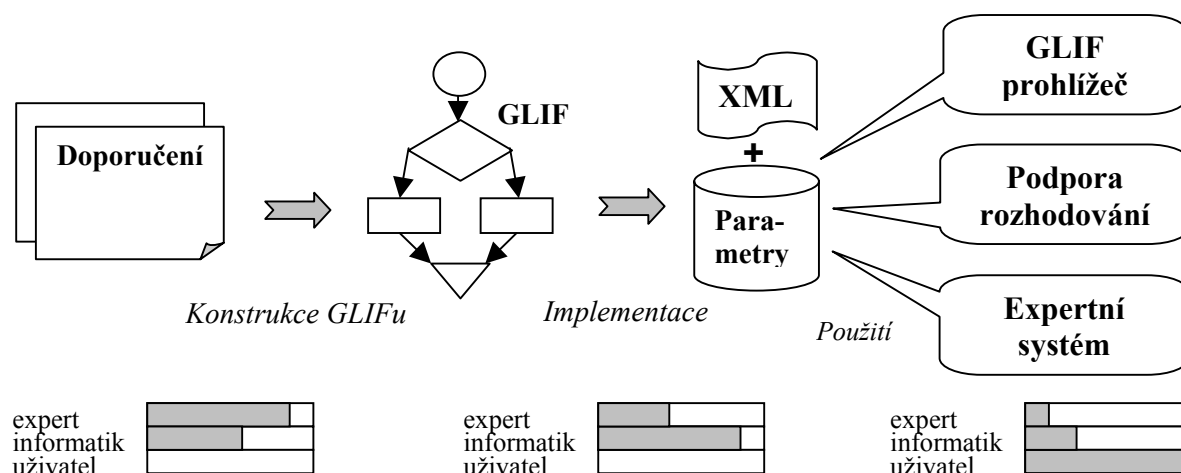
Keywords: GLIF model, implementation stage, data interface

ÚVOD

Oborová doporučení jsou obvykle distribuována v textové podobě, pro počítačovou implementaci a zpracování je však nutné mít tato doporučení ve strukturované formě. Proces formalizace textových doporučení do strukturované podoby (GLIF modelu) není zcela triviální a lze jej rozdělit do několika fází (viz obrázek 1). Ve fázi konstrukce GLIF modelu z textových doporučení je důležité najít procesní strukturu doporučení a všechny podstatné parametry modelu a jejich vzájemné vztahy. Ve fázi implementace GLIF modelu je grafický model doporučení zakódován v XML. Mimoto je vytvořen i seznam základních a odvozených parametrů modelu (viz [1]).

CÍL PŘÍSPĚVKU

Pro použití GLIF modelu pro podporu rozhodování je nezbytné navázat výsledný model rozhodování na data o konkrétním případě. Data jsou obvykle uložena v příslušné databázi či informačním systému. Cílem tohoto příspěvku je návrh obecného datového rozhraní mezi GLIF modelem a konkrétními daty uloženými v databázi.



Obrázek 1: Proces konstrukce a implementace GLIF modelu

METODIKA

GLIF model (*GuideLine Interchange Format*), nejčastěji používaný pro přehlednou reprezentaci oborových doporučení, vznikl spoluprací Univerzity Columbia, Harvardské, McGillovy a Stanfordské univerzity. GLIF poskytuje objektivě a procesně orientovaný pohled na oborová doporučení (viz [2]). Výsledným modelem je orientovaný graf skládající se z pěti hlavních částí (kroků):

akce - představuje specifickou činnost nebo událost. Akcí může být i podgraf, který dále zjemňuje danou činnost.

rozhodování – představuje větvení (výběr) na základě automatického splnění logického kritéria, kdy další postup grafem je dán výsledkem aritmetického nebo logického výrazu nad konkrétními daty, a nebo rozhodnutí uživatele, kdy akce reprezentovaná následným krokem může ale nemusí být provedena, nebo může probíhat paralelně s jinými akcemi. Uživatel má v tomto místě možnost se rozhodnout, kterou částí grafu bude dále pokračovat.

větvení a synchronizace – větvení se používá při modelování nezávislých kroků, které mohou probíhat paralelně a synchronizace slouží pro tyto kroky jako slučovací bod.

stav – značí stav, ve kterém se zkoumaný objekt nachází při vstupu do modelu nebo po provedení některého předchozího kroku.

Průchod jednotlivými větvemi GLIF modelu je závislý na splnění či nesplnění podmínek v rozhodovacích krocích nebo na interaktivním výběru příslušné větve grafu uživatelem. Pro modelování všech variant se používají tyto typy podmínek:

- **strict-in** – jestliže je splněna podmínka (např. logický výraz) nezávisle na uživateli, určitě se bude pokračovat příslušnou větví
- **strict-out** – jestliže je splněna podmínka, příslušnou větví se určitě nebude pokračovat – tato větev je zakázána
- **rule-in** – při splnění této podmínky je pouze doporučeno pokračovat následnou větví – je vyžadován zásah uživatele
- **rule-out** – při splnění této podmínky není doporučeno následnou větví – opět vyžadován zásah uživatele

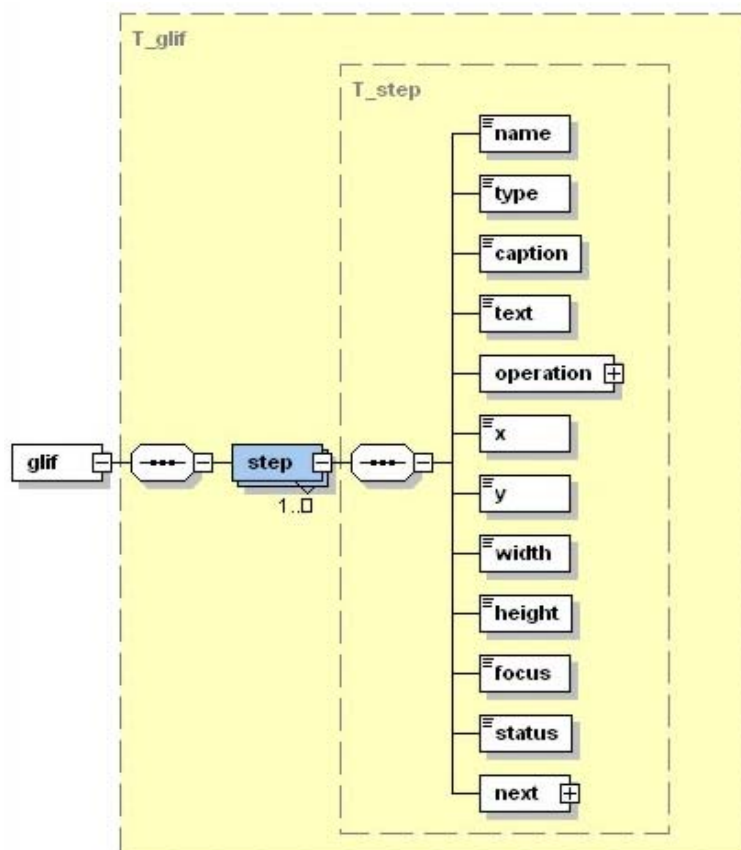
Všechny typy rozhodovacích podmínek mají obvykle tvar logické formule, kde proměnnými jsou základní a odvozené parametry GLIF modelu. Základní parametry představují přímo měřitelné (nebo jinak získatelné) hodnoty, odvozené parametry se získají aritmetickou, logickou či logicko-aritmetickou operací nad základními parametry.

VÝSLEDKY A DISKUZE

GLIF model je grafický, pro další počítačové zpracování je proto nutné jej zakódovat. Získaný GLIF model je možné, pro účely dalšího počítačového zpracování, uložit ve formátu XML (*eXtensible Markup Language*), který je v současné době přehledným a rozšířeným způsobem uložení veškerých dat v počítačových (nejen znalostních) systémech.

XML reprezentace GLIF modelu

Celý model se skládá z jednotlivých kroků (*step*) reprezentujících příslušné prvky GLIF modelu. Kromě atributů vyplývajících přímo z modelu obsahuje navržená syntaxe XML i atributy vhodné pro další zpracování, například pro grafické zobrazení nebo pro podporu rozhodování. Navržený XML model je patrný z obrázku 2.



Obrázek 2: Elementy jednotlivých kroků GLIF modelu

Význam jednotlivých XML elementů GLIF modelu je následující:

- **glif** - kořenový element GLIF modelu, skládá se z kroků (elementů **step**)
- **step** - parametry jednoho kroku
- **name** - jednoznačná identifikace (název) kroku
- **type** - typ kroku, je možný pouze výběr z typů *action*, *case*, *branch*, *synchronization*, *state*, *subgraph*
- **caption** - text v záhlaví grafického symbolu kroku
- **text** - vlastní text popisující krok

- **operation** - seznam operací probíhajících „na pozadí“ kroku. Jedná se především o operace načtení a nastavení parametrů modelu
- **x** - x-ová (horizontální) souřadnice grafického symbolu
- **y** - y-ová (vertikální) souřadnice grafického symbolu
- **width** - šířka (v pixelech) grafického symbolu
- **height** - výška (px) grafického symbolu
- **focus** - zvýraznění kroku, možné hodnoty jsou *not* = nezvýrazněn, *auto* = zvýrazněn (automaticky), *user* = zvýrazněn (uživatelská simulace)
- **status** - postavení kroku v rámci modelu, možné hodnoty jsou *start* = počáteční krok, *end* = koncový krok, *in* = vnitřní krok (všechny ostatní)
- **next** - seznam následných větví kroku. Pro každou následnou větev (volbu) jsou kromě dalších atributů definovány všechny typy rozhodovacích podmínek (*strict-in*, *strict-out*, *rule-in*, *rule-out*)

Rozhodovací a synchronizační podmínky

Zápis rozhodovacích podmínek (*strict-in*, *strict-out*, *rule-in*, *rule-out*) dodržuje syntaxi jazyka **XPath**⁸. Rozhodovací podmínky pracují s parametry uloženými v odpovídajícím paramodelu (modelu parametrů).

Příklad definice podmínky testující pravdivost obou parametrů PRIZNAKY a NALEZY:
 /params/param[pid="PRIZNAKY"]/pvalue **and** params/param[pid="NALEZY"]/pvalue

Zápis synchronizačních podmínek je obdobný (opět v jazyku **XPath**), předpokládá ale definici „on-fly“ parametrů definujících průchod příslušným krokem.

Příklad definice synchronizační podmínky, kdy je vyžadován průchod větví S3 nebo větví S4:
 /params/param[pid="DG_S3"]/pvalue **or** /params/param[pid="DG_S4"]/pvalue

XML reprezentace paramodelu

Pro reprezentaci paramodelu (modelu parametrů) je, stejně jako u GLIF modelu, využito jazyka XML. Celý model se skládá z XML elementů reprezentujících příslušný parametr (viz obrázek 3).

Parametry GLIF modelu mohou být trojího typu a to:

basic – základní parametr, tj. přímo měřitelná nebo získatelná hodnota

derived – parametr odvozený ze základních provedením logické, aritmetické či logicko-aritmetické operace

onfly - pomocný parametr pro definici synchronizačních podmínek v GLIF modelu

Význam jednotlivých XML elementů paramodelu je následující:

params - kořenový element parametrového modelu, skládá se z parametrů (elementů **param**)

param - popis jednoho parametru

pid - jedinečná identifikace (název) parametru

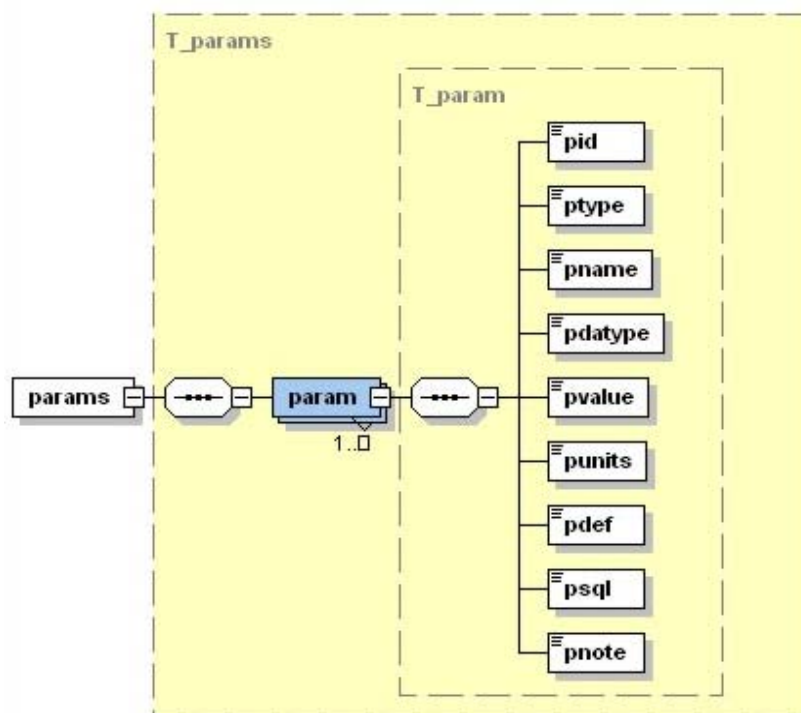
pctype - typ parametru (*basic*, *derived*, *onfly*)

pname - text úplného názvu parametru

pdatatype - datový typ parametru odpovídající některému z XML datových typů

⁸ **XPath** slouží k reprezentaci XML dokumentu v podobě stromu uzlů. Uzlem může být element, atribut nebo text. K popisu se používá podobná syntaxe jako k popisu cesty v souborovém systému. Navíc lze použít standardní predikáty a výrazy. Bližší specifikaci XPath lze nalézt na <http://www.w3.org/TR/xpath>.

- pvalue** - vlastní hodnota parametru
- punits** - jednotky parametru
- pdef** - definice operace na základních parametru v syntaxi XPath pro odvozené parametry
- psql** - definice dotazu do databáze na hodnotu parametru v jazyce SQL (*Structured Query Language*)
- pnote** - text poznámky k parametru



Obrázek 3: Elementy parametrového modelu

ZÁVĚR

Parametry podstatné pro modelování oborových doporučení pomocí GLIF modelu nemusí a často ani nebývají uloženy způsobem vhodným pro přímé použití GLIF modelem. Výhodnějším řešením je navržené rozhraní mezi GLIF modelem a konkrétními uloženými daty ve formě paramodelu reprezentovaném v XML. Skutečné hodnoty se pak získají vyhodnocením SQL dotazu, který by měl být dostatečně obecným nástrojem pro různé databáze a informační systémy. Vyhodnocování rozhodovacích podmínek se potom provádí s odkazem na parametry uložené v paramodelu.

LITERATURA

- [1] Buchtela D., Peleska J., Vesely A., Zvarova J.: *Method of GLIF model Construction and Implementation*, The XIX International Congress of the European Federation for Medical Informatics, 2005, 6 pp.
- [2] 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*, Journal of the American Medical Informatics Association 1998, 5(4), pp. 357-372, ISSN 1067-5027.

KONTAKT

Ing. David Buchtela
 KII PEF ČZU Praha
 e-mail: buchtela@pef.czu.cz