

ROZŠÍŘENÍ SPECIFIKACE GLIF MODELU

EXTENSION OF GLIF MODEL SPECIFICATION

David Buchtela

Anotace:

Pro reprezentaci oborových doporučení se nejčastěji používá model GLIF (*GuideLine Interchange Format*). Při praktickém použití je však modelování některých situací jen velmi obtížné nebo nemožné. Proto je nutné provést rozšíření specifikace GLIF modelu. V tomto článku je popsáno rozšíření tak, aby řešilo problematické modelové situace při zachování univerzality GLIF modelu pro libovolná oborová doporučení.

Klíčová slova:

oborová doporučení, rozšířený GLIF model, implementace GLIF modelu

Abstract:

GLIF model is mostly used for branch guidelines representation. But modelling of any situations is too difficult or impossible in a practice use. Therefore an extension of GLIF model specification is necessary. This article describes the extension, which tries to find a solution of problematical model situations. The GLIF model universality is kept for any branch guidelines.

Keywords:

branch guidelines, extended GLIF model, GLIF model implementation

ÚVOD A CÍL

Znalosti získané v různých oborech je možné formalizovat do oborových doporučení usnadňujících rozhodování v daném konkrétním případě opírající se o znalostní bázi oboru. Pro počítačovou implementaci a zpracování je nutné mít tato doporučení ve strukturované formě. Pro přehlednou reprezentaci doporučení byl vybrán model GLIF.

Při praktickém modelování některých situací (popsaných dále v textu) se ukázalo, že se jedná o příliš složitý nebo dokonce nemožný proces. Jinou možností je rozšíření specifikace GLIF modelu, jejíž návrh a popis je cílem tohoto článku.

METODIKA

GLIF model vznikl spoluprací Univerzity Columbia, Harvardské, McGillovy a Stanfordské univerzity. GLIF poskytuje objektivě a procesně orientovaný pohled na oborová doporučení (viz [1], [3] a [4]). 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.

Získaný model je možné implementovat v XML (*eXtensible Markup Language*). Více o implementaci GLIF modelu viz [2].

Typy rozhodovacích podmínek

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 touto větví
- **strict-out** – jestliže je splněna podmínka, touto větví se určitě nebude pokračovat – následující 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

VÝSLEDKY – PROBLÉMOVÉ SITUACE A JEJICH ŘEŠENÍ

Použití třístavové logiky

Při vyhodnocování logických podmínek (*strict-in*, *strict-out*, *rule-in*, *rule-out*) se velice často setkáme se situací, kdy nejsou známy jednoznačné hodnoty všech vyhodnocovaných proměnných. Neznámá hodnota proměnné nemusí být přitom chybou a je třeba takový stav brát v úvahu. Logické formule, složené z proměnných (parametrů modelu) a logických či relačních operandů, mohou nabývat následující stavy:

- **pravda** – formule je splněna – označíme hodnotou 1
- **nepravda** – formule není splněna – označíme hodnotou 0
- **neznámo** – stav formule není jednoznačně určený – označíme hodnotou $\frac{1}{2}$.

Z matematické logiky plyne, že jakoukoliv logickou formuli lze převést do disjunktivního (resp. konjunktivního) normálního tvaru, složeného pouze z logických operací negace, konjunkce (logický součin) a disjunkce (logický součet). Výsledný tvar formule je ekvivalentní s původní formulí.

Definici operací negace, konjunkce a disjunkce v třístavové logice uvádějí následující tabulky:

proměnná p	negace $\neg p$
0	1
1	0
$\frac{1}{2}$	$\frac{1}{2}$

p	q	logický součin $p \wedge q$	logický součet $p \vee q$
0	0	0	0
0	1	0	1
0	$\frac{1}{2}$	0	$\frac{1}{2}$
1	0	0	1
1	1	1	1
1	$\frac{1}{2}$	$\frac{1}{2}$	1
$\frac{1}{2}$	0	0	$\frac{1}{2}$
$\frac{1}{2}$	1	$\frac{1}{2}$	1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Logický součin je možné, pro libovolný počet proměnných ve formuli, definovat vztahem:

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) = \min(p_1, p_2, \dots, p_n)$$

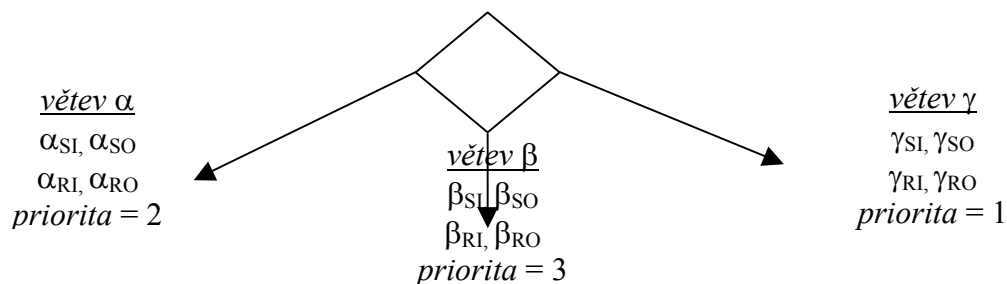
Obdobně je možné definovat i vztah pro logický součet:

$$(p_1 \vee p_2 \vee \dots \vee p_n) = \max(p_1, p_2, \dots, p_n)$$

Pro jednoznačný výběr následné větve grafu v rozhodovacím kroku GLIF modelu je nutné, aby výsledné ohodnocení alespoň typů podmínek *strict-in* a *strict-out* příslušné větve nabývalo jednoznačné hodnoty, tj. 0 nebo 1. V případě, že některý z typů podmínky (*strict-in* nebo *strict-out*) nabývá hodnoty $\frac{1}{2}$ (neznámo), musí uživatel upřesnit hodnoty parametrů vstupujících do logické formule tak, aby bylo dosaženo jednoznačného výsledku.

Priorita rozhodovacích větví

Při důsledném uplatňování předchozího pravidla může nastat situace, ve které bude po uživateli vyžadováno upřesnění hodnot vstupních parametrů i v případě, kdy to pro další průchod GLIF modelem není nezbytně nutné. Mějme např. následnou modelovou situaci:



Pro každou následnou větev jsou definovány typy podmínek α_{SI} – *strict-in*, α_{SO} – *strict-out*, α_{RI} – *rule-in* a α_{RO} – *rule-out*. Při výběru následné větve se postupuje tak, že se nejprve vyhodnotí podmínka typu *strict-out* a při jejím splnění se další typy již nevyhodnocují a z pohledu možného průchodu grafem je příslušná větev zakázána. V opačném případě se vyhodnotí podmínka typu *strict-in*. Je-li splněna pokračuje se touto větví. Celkový model je

nutné navrhnout tak, aby byla splněna podmínka *strict-in* maximálně u jedné z možných větví. Jestliže není splněna ani jedna podmínka typu *strict-in*, je vyžadován zásah uživatele s tím, že na základě vyhodnocení podmínek *rule-in* a *rule-out* je doporučen nebo nedoporučen výběr příslušné větve.

Uvažujeme-li výsledky typů podmínek v třístavové logice, je třeba vyžadovat po uživateli upřesnění vstupních parametrů (s hodnotou neznámo) tak, aby bylo možné jednoznačně tyto výsledky určit. Upřesnění vstupních parametrů může však znamenat provedení nějakého, třeba i ekonomicky náročného, měření nebo speciálního úkonu. V případě, že u některé z větví nabývá typ podmínky *strict-in* hodnoty 1, bude se pokračovat určitě touto větví a vyhodnocování podmínek (a s tím spojené upřesňování parametrů) je vlastně zbytečné.

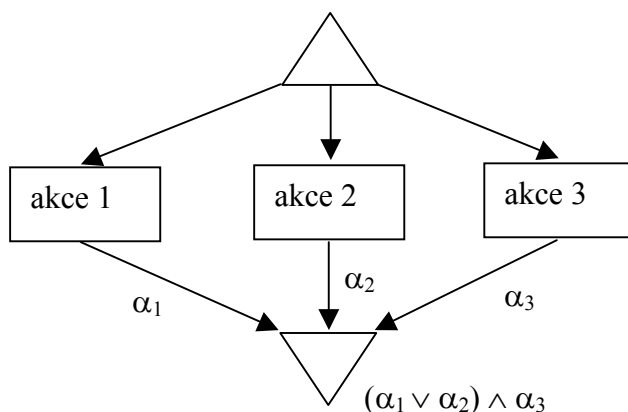
Počet potřebných upřesnění je závislé na pořadí vyhodnocování jednotlivých větví. Z uvedených důvodů je výhodné stanovit pořadí jejich vyhodnocování, tj. určit prioritu. Priorita je přiřazena jednotlivým větvím rozhodování při návrhu GLIF modelu po expertní analýze modelovaného problému. V uvedeném příkladu se tedy nejprve bude vyhodnocovat větev γ (priorita = 1), poté větev α (priorita = 2) a nakonec větev β (priorita = 3). Jestliže bude u větve γ splněna podmínka γ_{SI} , nebudou se větve α a β vůbec vyhodnocovat.

Synchronizační podmínky

Při modelování paralelně probíhajících větví GLIF modelu pomocí prvků větvení a synchronizace je nutné upřesnit, které z větví musí uživatel určitě projít nebo zda stačí, aby vybral některou či některé z nich.

Uvedené případy lze realizovat přidáním tzv. synchronizační podmínky v synchronizačním kroku (viz následující příklad). Podmínku je možné opět vyjádřit v disjunktivním (konjunktivním) normálním tvaru, kde vstupními proměnnými jsou ohodnocení jednotlivých větví:

- **hodnota 1** – touto větví uživatel prošel
- **hodnota 0** – touto větví uživatel neprošel



V ukázkové situaci se bude pokračovat v průchodu grafem, jestliže bude splněna synchronizační podmínka $(\alpha_1 \vee \alpha_2) \wedge \alpha_3$, tj. jestliže uživatel určitě projde větví α_3 a zároveň alespoň jednou z větví α_1 nebo α_2 .

Logicko-numerické operace

GLIF model umožňuje mimo vyhodnocování rozhodovacích podmínek na podkladě vstupních parametrů i ukládání a následné použití parametrů zjištěných (odvozených) až během průchodu grafem. Ve většině případů se jedná o stavy zjištěné jako výsledky

vyhodnocení logických formulí. Hodnota takto uloženého parametru je pak v souladu s třístavovou logikou uvedenou výše.

V některých případech je však potřeba, aby hodnota odvozeného parametru nabývala určité číselné hodnoty v závislosti na vyhodnocení dílčích logických formulí potřebných vstupních parametrů. Například modelování následující tabulky (jedná se pouze o fiktivní hodnoty):

kouření (počet cigaret denně)	procento rizika
< 10	10 %
10 ÷ 40	25 %
> 40	60 %

Uvedenou situaci lze řešit pomocí prvků rozhodování a stav GLIF modelu, zvláště pro každý řádek tabulky. V případě, že by bylo více řádků tabulky, vedl tento způsob k příliš složitému a nepřehlednému modelu. Efektivnějším řešením je zavedení logicko-numerických operací ve tvaru:

odvoz.parametr = (podmínka1:hodnota1; podmínka2:hodnota2; ...; podmínkaN:hodnotaN)

V ukázkovém případě by definice odvozeného parametru *riziko* na základě hodnoty *kour* vypadala následovně:

riziko = ((kour<10):10; ((kour≥10)^(kour≤40)):25; (kour>40):60)

ZÁVĚR

Popsaná rozšíření GLIF modelu (tj. tříhodnotová logika, priorita rozhodovacích větví, synchronizační podmínky a logicko-numerické operace) a jejich důsledné použití umožňuje modelování problémových situací, které se vyskytly během praktické aplikace GLIF modelu. Navržená řešení zachovávají univerzalitu GLIF modelu pro libovolná oborová doporučení a jejich přehledné zobrazení bez příliš složitých konstrukcí.

Literatura:

- [1] Buchtela D.: *E-modelování znalostí pro podporu rozhodování*, Sborník prací - Agrární perspektivy XII., PEF ČZU Praha, 2003, str. 37-44, ISBN 80-213-1056-1
- [2] Buchtela D.: *Implementace GLIF modelu v XML*, Sborník příspěvků, PEF ČZU Praha, 2004, str.48 + CD, ISBN 80-213-1150-9
- [3] 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.
- [4] Ash N., Bernstam E., Greenes R.A., Lacson R., Mork P., Shortliffe E.H.: *Guideline Interchange Format 3.5 Technical Specification* , online - 12.12.2002, dostupné z: http://smi-web.stanford.edu/projects/intermed-web/guidelines/GLIF_TECH_SPEC_Dec_12_2002.pdf

Kontakt:

Ing. David Buchtela, KII PEF ČZU Praha,
tel.: +420 224 343 806, e-mail: buchtela@pef.czu.cz