

APLIKACE KLASICKÝCH A MODERNÍCH ONTOLOGIÍ PŘI TVORBĚ ZNALOSTNÍCH BÁZÍ LÉKAŘSKÝCH PROGRAMŮ

Petr Lesný, Kryštof Slabý, Tomáš Holeček, Jan Vejvalka

Anotace

Algoritmy a doporučené postupy jsou uznávanou součástí klasické medicíny. Se vzrůstem jejich počtu a počtu jejich praktických aplikací vzrůstá potřeba vytvářet explicitní specifikace konceptualizace jak algoritmů samotných, tak dat, která zpracovávají. Pro popis lékařských algoritmů jsme vypracovali doménovou ontologii, jejímž základem je fenomenologický popis zpracovávaných dat jako indikátorů a transformací těchto indikátorů v indikátory jiné. Aby mohl být námi navržený systém součástí medicíny založené na důkazech, musí být pro všechny popisované koncepty dostupná plná sémantická informace, která také umožňuje propojování znalostí napříč jednotlivými znalostními doménami.

Klíčová slova

Doménové ontologie, Biomedicínské algoritmy, Sémantika, Medicína založená na důkazech

1. Úvod

Lékařské algoritmy, ať už v podobě výpočtů nebo doporučených postupů (guidelines) jsou součástí klasické medicíny. Současně je k dispozici celá řada nástrojů, včetně lékařského software, které usnadňují implementaci těchto algoritmů v praxi. Při vzrůstání počtu publikovaných algoritmů (například v rámci projektů MEDAL [1], MedCalc [2] apod.) se však do popředí dostávají otázky dokumentace (a dokumentovatelnosti) těchto algoritmů a otázky důvěry v tyto algoritmy.

Jedním z nástrojů, které jsou dnes využívány při dokumentování lékařských algoritmů jsou ontologie. Ontologie, resp. explicitní specifikace konceptualizace lékařských algoritmů, umožňují doplnit klasické archívy lékařských algoritmů i software pro jejich zpracování o nové funkce, které umožní využití lékařských algoritmů v rámci striktních pravidel medicíny založené na důkazech [3].

2. Aristoteléská ontologie

Studium anatomie a hierarchického uspořádání lidského těla ve starověku přineslo (mimo jiné) i klasické ontologie, které mohou být aplikovány na biomedicínu. Tyto ontologie, které nazýváme „Aristoteléské“ vycházejí z potřeby nalézt nástroje pro jednotný popis celého univerza. Klasicky

vycházejí ze základního prvku, který nazývají například „koncept“ nebo „entita“, a opisují svět pomocí relativně malého množství vztahů mezi popisovanými jevy. Některé klasické ontologie v biomedicíně (SNOMED [4], ON9.2 [5] apod.) vycházejí jen ze vztahů „is-a“ a „part-of“, jiné (např. UMLS [6]) používají až 80 vztahů mezi jevy.

Klasické ontologie nám pomáhají hledat nové pohledy na biomedicínu, nicméně při jejich praktickém použití narážíme na zcela pragmatickou nemožnost popsat celou biomedicínu (jejíž součástí je například hierarchie ... → člověk → orgány → tkáně → buňky → organely → bílkoviny → ...). Komplexnost klasických ontologií se projevuje také v snadném zavlečení chyb, které je možné odhalit například porovnáním dvou klasických ontologií mezi sebou [7]. Obtížnost použití klasických ontologií pro popis dat zpracovávaných lékařskými algoritmy jsme popisovali již dříve [8].

3. Moderní ontologie

Základní rozdíly mezi klasickými a moderními ontologiemi jsou shrnuty v tabulce 1. Příkladem moderních ontologií v biomedicíně jsou například MeSH [9], IUPAC [10] nebo McKusickova syndromologie dědičných chorob [11]. Rozdíly mezi klasickými a moderními ontologiemi jsou přehledně shrnuty v tabulce 1.

Klasické (Aristotelské) ontologie	Moderní ontologie
Konstruovány shora (entita → ...)	Konstruovány zdola
Obecné	Specifické pro oblast poznání
	Specifické pro očekávanou aplikaci
Malý počet vztahů (is_a, part_of, ...)	Komplexní vztahy mezi prvky

Tab. 1 – Shrnutí rozdílů mezi klasickými a moderními ontologiemi.

Moderní ontologie, někdy nazývaná také „doménově specifická ontologie“ pro lékařské algoritmy by měla mít následující vlastnosti

- Filosofický model, který umožňuje korektní počítačové zpracování dat, aniž by od zpracovávajícího software vyžadoval „porozumění“.
- Explicitně vyjádřenou sémantickou informací, která je vázána na současné paradigma medicíny založené na důkazech (*Evidence Based Medicine*).
- Podpora kontroly a jistění kvality lékařských algoritmů a zacházení s důvěrou v tyto algoritmy (*Trust management*).

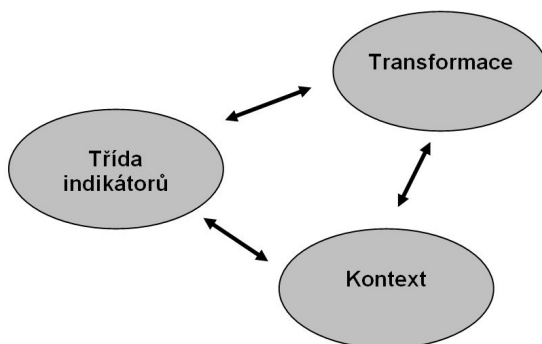
- Umožňuje organizovat předávání zpracovávaných dat mezi jednotlivými popisovanými algoritmy.

4. Moderní ontologie pro popis lékařských algoritmů

Moderní ontologii pro popis lékařských algoritmů je možné založit například na fenomenologii [12], obr. 1. Zpracovávaná data zde popisujeme jako indikátory (indikující čtenáři, kterým je vždy člověk, nějakou skutečnost). Tyto indikátory mohou být transformovány v jiné indikátory (bez ohledu na to, zda jsou transformovány člověkem, nebo automatem). Indikátory se podle své role v transformaci sdružují do tříd indikátorů (například třída indikátorů popisujících tělesnou výšku).

Indikátory se vyskytují v kontextu (například indikátory, které patří k jednomu pacientovi jsou ve stejném kontextu) a v kontextu jsou také transformovány. Výskyt indikátorů ve společném kontextu označuje, že tyto indikátory patří definovaným způsobem k sobě (podobně, jako by byly uvedeny pod sebou na listu papíru). Z pragmatického hlediska může být naopak kontext určen v něm se vyskytujícími indikátory.

Popis transformací, tříd indikátorů i kontextů (dále je budeme uvádět jako *popisované entity*) musí obsahovat strukturovanou **sémantickou informaci** nutnou k tomu, aby uživatel porozuměl jejich použití a případným omezením; sémantická informace je provázána na publikované práce (citace) a na osoby nebo instituce, které jsou autory jednotlivých sémantických popisů. Struktura sémantické informace může vycházet například ze zkušeností, kterých bylo dosaženo v projektu *S.M.A.R.T.I.E.* [13]. Vzhledem ke změnám, ke kterým v biomedicíně pravidelně dochází a které mohou ovlivňovat využití popisovaných transformací apod., jsou však pro každou popisovanou entitu zajištěna pravidla revizí, která umožňují udržovat kvalitu popisu a možnost vyjádření důvěry (review).



Obr. 1 – Grafické vyjádření indikátorové ontologie

Jednoznačný sémantický popis tříd indikátorů, který je navíc doplněn o vztahy ke transformacím a dalším popisovaným entitám [8] pak slouží jako základ pro spojování jednotlivých transformací do komplexních proudů pro zpracování dat (*workflow*).

5. Závěr

Námi popisovaná moderní ontologie pro popis lékařských algoritmů byla úspěšně vyzkoušena v rámci projektu MediGrid [14] pro řadu odborných domén (pneumologie, růstová medicína, kardiologie). Podporováno z grantu 1ET202090537.

6. Literatura

- [1] Sen D.: The Medical Algorithms Project. *Occupational Medicine* (2005) 645.
- [2] Stroud S.D., Erkel E.A., Smith C.A.: The use of personal digital assistants by nurse practitioner students and faculty. *J Am Acad Nurse Pract* (2005) 67–75.
- [3] Birnbaum M.L. Guidelines, algorithms, critical pathways, templates, and evidence-based medicine. *Prehosp Disaster Med.* 1999 Jul–Sep;14(3):114–5.
- [4] Cote R.A., Robboy S.: Progress in medical information management. Systematized nomenclature of medicine (SNOMED). *Jama* (1980) 756–62.
- [5] Pisanelli DM, Gangemi A, Steve G. The role of ontologies for an effective and unambiguous dissemination of clinical guidelines. In R. Dieng and O. Corby (eds.), *Knowledge Engineering and Knowledge Management*. Springer Verlag, 2000.
- [6] Lindberg C.: The Unified Medical Language System (UMLS) of the National Library of Medicine. *J Am Med Rec Assoc* (1990) 40–2.
- [7] Ceusters W, Smith B, Kumar A, Dhaen C. Ontology-based error detection in SNOMED-CT(R). *Medinfo* 2004;482–6.
- [8] Vejvalka J., Lesny P., Holecek T., Slaby K., Jarolimkova A., Bouzkova H.: *MediGrid – Facilitating Semantic-Based processing of Biomedical Data and Knowledge*; Karopka T., Correia R.J., (eds): *Open Source in European Health Care: The Time is Ripe*. Porto, Portugal, INSTICC PRESS, 2009.
- [9] Lipscomb C.E.: Medical Subject Headings (MeSH). *Bull Med Libr Assoc* (2000) 265–6.
- [10] <http://www.iupac.org>
- [11] Hamosh A., Scott A.F., Amberger J., Valle D., McKusick V.A. Online Mendelian Inheritance in Man (OMIM). *Hum Mutat.* 2000;15(1):57–61.

- [12] Husserl E. Logical Investigations, Investigation I. (translated by J. N. Findlay). Routledge, London (2001) 184.
- [13] Smart Medical Applications Repository of Tools For Informed Expert (S.M.A.R.T.I.E.) and MedNotes™. IST (2000). <http://www.smartie-ist.org>
<http://medigrid.sourceforge.net>

Kontakt:

MUDr. Petr Lesný
Ústav neurověd UK 2. LF
V Úvalu 84
150 06 Praha 5
e-mail: lesny@cesnet.cz