

Formaliųjų ontologijų vaidmuo šiuolaikinėje informacinių sistemų inžinerijoje

Donatas ČIUKŠYS, Albertas ČAPLINSKAS* (MII)

el. paštas: donatas.ciuksys@maf.vu.lt

1. Įvadas

Formaliųjų ontologijų svarba informacinių sistemų (IS) inžinerijoje ir informatikoje pastaraisiais metais labai išaugo. Jeigu prieš keletą metų ontologijos buvo naudojamos tik kaip formalizuoti dalykinės srities žodynai, tai šiandien jos naudojamos semantinėms užklausoms realizuoti, elektroniniams dokumentams analizuoti, programų sistemoms projektuoti ir daugeliui kitų tikslų. Ontologijos naudojamos visur, kur grupė asmenų, organizacinių vienetų ar kitokių agentų turi bendrauti tarpusavyje ir vienodai interpretuoti vartojamus terminus bei sąvokas. Ontologijų svarba informacinių sistemų inžinerijoje visų pirma yra sąlygota silpno šios disciplinos formalizavimo laipsnio. Netgi bazinės šios disciplinos sąvokos yra apibrėžtos neformaliai [11], ko pasekoje skirtingos mokslinės mokyklos jas interpretuoja skirtingai. Tai sukelia tam tikrą painiavą, apsunkina bendravimą ir mokslinių publikacijų supratimą. Ontologijos gali bent iš dalies šią problemą išspręsti [1]. Tačiau čia susiduriama su nauja problema, skirtingomis pačių ontologijų traktuotėmis. Nepaisant mokslinių publikacijų ontologijų ir jų naudojimo klausimais gausos, kol kas nėra vieningos, visuotinai pripažintos ontologijos sampratos. Iš dalies tai taikytina ir formaliosioms ontologijoms. Šio straipsnio tikslas – panagrinėti skirtingas formaliosios ontologijos sampratas, pabandyti jas apibendrinti ir pasiūlyti, kokia traktuotė būtų priimtinausia informacinių sistemų inžinerijos kontekste.

2. Ontologijos samprata

Dirbtinio intelekto srityje ontologijos paprastai yra traktuojamos [10] kaip *formaliosios loginės konceptų teorijos*, apibrėžiančios ne tik dalykinės srities terminus ir jų tarpusavio sąryšius, bet ir jų vartojimo kontekstą bei išvestinius terminus ir sąryšius. Šitokios teorijos yra pakankamai formalios, kad būtų galima tikrinti jų išsamumą ir neprieštaringumą. Duomenų bazių ir kai kuriose kitose srityse ontologijos dažniausiai suprantamos panašiai kaip koncepciniai modeliai, taksonomijos ar schemas. Šitai suvokiamomis technologijomis aprašyti terminų vartojimo kontekstą ar kitas sudėtingesnes priklausomybes

*Darbas atliktas Matematikos ir informatikos institute, vykdant planinę temą „Ontologijomis grindžiamų komponentinių programų, informacinių ir verslo sistemų inžinerijos problemos“.

neįmanoma. Dar kitaip traktuojamos ontologijos kompiuterinėje lingvistikoje. Pavyzdžiui, leksinėje anglų kalbos duomenų bazėje WordNet [4] ontologija traktuojama kaip tezaurus, aprašantis sinonimijos, antonimijos ir kitus panašius santykius tarp sąvokų, bet nenusakantis tikslios sąvokos prasmės. Atskirai dalykinei sričiai, kurioje nagrinėjamas siauras problemų ratas, ontologijos vaidmenį gali atlikti netgi reliacinės duomenų bazės schema, nes visos dalykinės programos bendrauja tarpusavyje, naudodamosis ta schema aprašyta dalykinės srities samprata (konceptualizacija).

Gruber [6] pasiūlė ontologijos apibrėžimą, plačiai pripažintą dirbtinio intelekto srityje, teigianti, kad „ontologija yra konceptualizacijos specifikacija išreikštiniu būdu“. Šis apibrėžimas yra lakoniškas, bet tuo pačiu ir nevienareikšmiškai interpretuojamas. Guarino pastebėjo, kad, pirma, specifikacija gali būti įvairaus detalumo lygio, antra, apibrėždamas konceptualizacijos sampratą, Gruber citavo Genesereth ir Nilsson 1987 metų darbą, kuriame konceptualizacija suprantama kaip ekstensionalinė semantinė struktūra. Pirmiausiai aptarkime pastarąją pastabą.

3. Konceptualizacija – intensionalinė ar ekstensionalinė semantinė struktūra?

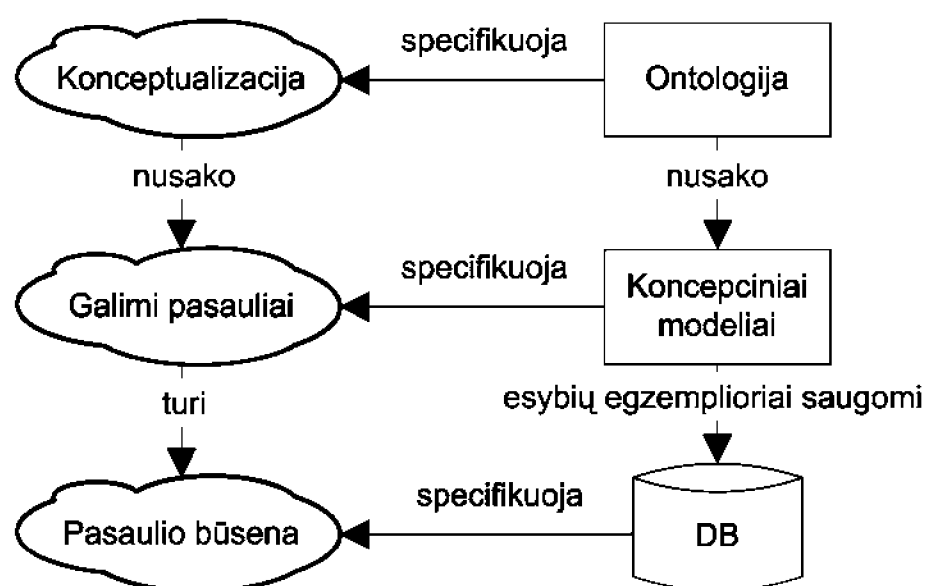
Genesereth ir Nilsson [5] pasiūlė konceptualizaciją formalizuoti kaip struktūrą $\langle D, R \rangle$, čia D – nagrinėjamos srities esybių egzempliorių aibė (angl. universe of discourse), R – nagrinėjamoje srityje svarbių ryšių aibė. Pavyzdžiui, knygų lentynos su keturiomis knygomis b, c, a, d konceptualizacija galėtų atrodyti taip:

$$\langle \{a, b, c, d\}, \{pirmaKnyga, stoviKairiau\} \rangle$$

Čia a, b, c, d – knygos, $pirmaKnyga$ ir $stoviKairiau$ – atitinkamai unarinis ir binarinis ryšiai: $pirmaKnyga = \{b\}$, $stoviKairiau = \{\{b, c\}, \{c, a\}, \{a, d\}\}$. Jei pirmą knygą sukeistume vietomis su antrąja, abu ryšiai jau būtų apibrėžti kitaip, iš ko sektų, kad ir konceptualizacija būtų kitokia. Tačiau intuityviai yra aišku, kad šių ryšių prasmė nepriklausomai nuo knygų išsidėstymo išlieka ta pati. Problema yra ta, kad pasiūlytas konceptualizacijos formalizavimo būdas yra ekstensionalinis – priklausantis nuo konkretios pasaulio būsenos. Siekdamas pašalinti šį trūkumą, Guarino [7] pasiūlė struktūrą $\langle D, R \rangle$ pavadinti pasaulio struktūra, o pačią konceptualizaciją formalizuoti kaip intensionalinę struktūrą $\langle D, W, \mathbf{R} \rangle$, čia W – galimų pasaulių aibė, \mathbf{R} – taip vadinamų koncepcinių ryšių aibė. n -ariškumo koncepcinis ryšys r^n apibrėžiamas kaip funkcija $r^n: W \rightarrow \wp(D^n)$, t.y., funkcija, su kiekvienu galimu pasauliu susiejanti ekstensionalinį ryšį.

Tokiu būdu gauname, kad konceptualizacija nusako tam tikrą skaičių galimų pasaulių, o kiekvienas pasaulis turi savo būseną (ekstensionalinius ryšius). Pasaulių idėja nėra nauja, ji yra naudojama modalinės logikos semantikai apibrėžti [3].

Bishr [2] pastebėjo, kad ontologija nusako koncepciniams modeliams bendrus principus, o koncepciniai modeliai aprašo konkretų pasaulį. Būtų galima pridurti, kad pasaulio būseną informacinėse sistemose paprastai saugoma duomenų arba žinių bazėse (1 pav.).



1 pav. Konceptualizacijos, ontologijos ir koncepcinių modelių sąryšiai.

4. Konceptualizacijos specifikuojimo detalumas

Kuo detaliau yra specifikuota konceptualizacija, tuo labiau vienareikšmiški yra vartojami terminai. Aksiomos yra naudingos apribojant daugiaprasmiškumą. Pavyzdžiui, bet kuriai knygai x teisingas teiginys: $\forall x \neg stoviKairiau(x, x)$. Verta pastebėti, kad šis teiginys teisingas visuose galimuose pasauliuose, todėl galima pasinaudoti modaline logika ir šį teiginį užrašyti šitaip: $\Box \forall x \neg stoviKairiau(x, x)$. Taip pat galima pastebėti, kad jei kuriame nors pasaulyje esybė yra knyga, tai ji bus knyga ir visuose pasauliuose: $\Box(\forall x(knyga(x) \Rightarrow \Box knyga(x)))$. Ši savybė [9] vadinama *ontologiniu griežtumu* (angl. *ontological rigidity*) ir nėra būdinga visiems unariniam ryšiams, todėl yra naudinga, kai siekiama apriboti unarinio ryšio (koncepto) galimas interpretacijas. Greta ontologinio griežtumo, [9] aksiomatizuojami ir kai kurie kiti panašaus pobūdžio ribojimai. Jų visuma uždraudžia neprasmingus sąryšius (ryšio ekstensionalo elementus). Naudojant šiuos ribojimus, konceptualizacijos specifikuojimas yra iš dalies formalizuojamas.

PAVYZDYS. Tarkime, turime dalykinę teoriją T1:

$$\begin{aligned} &\forall x(\text{\textit{šuo}}(x) \Rightarrow \text{\textit{žinduolis}}(x)) \\ &\forall x(\text{\textit{vilkas}}(x) \Rightarrow \text{\textit{žinduolis}}(x)) \\ &\text{\textit{šuo}}(a) \\ &\text{\textit{naminis}}(a) \end{aligned}$$

Šioje teorijoje galime išskirti nuo konkretaus pasaulio nepriklausančias vadinamąsias *ontologines žinias*. Paprastai jas sudaro visuose pasauliuose galiojančios aksiomos. Mūsų atveju tai aksiomos, aprašančios konceptų taksonomiją, vadinamąją *Tbox*¹. Jos sudaro *ontologinę teoriją T2*:

$$\begin{aligned} &\forall x(\text{\textit{šuo}}(x) \Rightarrow \text{\textit{žinduolis}}(x)) \\ &\forall x(\text{\textit{vilkas}}(x) \Rightarrow \text{\textit{žinduolis}}(x)) \end{aligned}$$

Guarino [8] įvedė terminą *ontologinė teorija*, įvardinantį loginę teoriją, skirtą išreikšti ontologines žinias (aibę aksiomų, galiojančių visuose pasauliuose).

¹ Aksiomos, aprašančios teiginius apie konkretaus pasaulio faktus, sudaro vadinamąją *Abox*.

Teorija T2 specifikuoja konceptualizaciją labai paviršutiniškai. Pasitelkus modalinę logiką, galima sudaryti teoriją T3, labiau apribojančią galimą konceptų interpretavimą:

$$\begin{aligned} &\Box(\forall x(\text{\textit{šuo}}(x) \Rightarrow \text{\textit{žinduolis}}(x))) \\ &\Box(\forall x(\text{\textit{vilkas}}(x) \Rightarrow \text{\textit{žinduolis}}(x))) \\ &\Box(\forall x(\text{\textit{žinduolis}}(x) \Rightarrow \Box\text{\textit{žinduolis}}(x))) \\ &\Box(\forall x(\text{\textit{šuo}}(x) \Rightarrow \Box\text{\textit{šuo}}(x))) \\ &\Box(\forall x(\text{\textit{vilkas}}(x) \Rightarrow \Box\text{\textit{vilkas}}(x))) \\ &\neg\Box(\forall x(\text{\textit{naminis}}(x) \Rightarrow \Box\text{\textit{naminis}}(x))) \end{aligned}$$

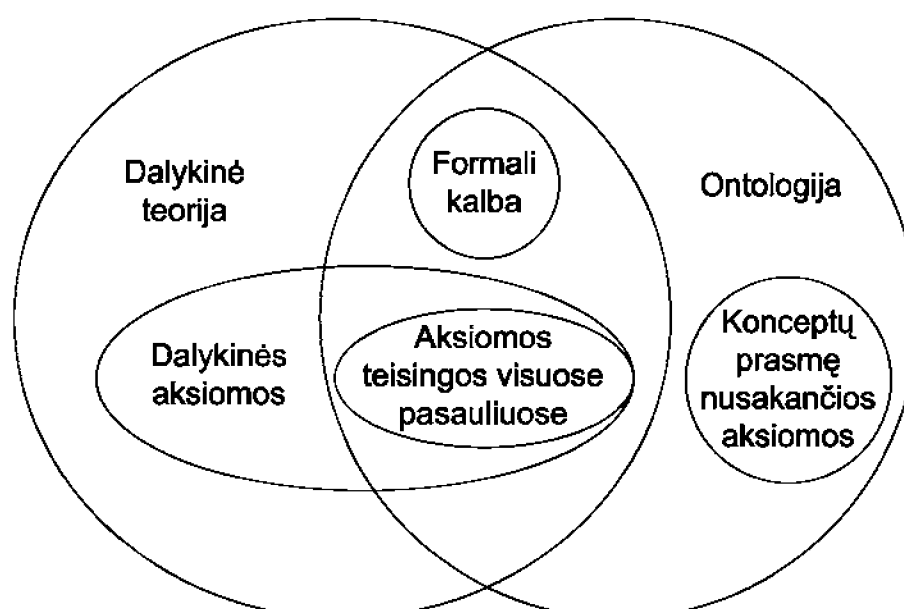
Teorija T3 nusako pačius bendriausius konceptų prasmės interpretavimo ribojimus. Guarino siūlo šią teoriją vadinti teorijos T1 *ontologinės nuostatos specifikacija*. Pažymėtina, kad T3 – tai teorijos T1 ontologinė teorija, besiskirianti nuo T2 tikslesniu konceptualizacijos specifikavimu.

Remdamasis šiais samprotavimais, Guarino [8] pasiūlė dvi ontologijos apibrėžtis, besiskiriančias konceptualizacijos specifikavimo tikslumu:

- ontologija – tai ontologinės teorijos sinonimas (pagal šią apibrėžtį, pateiktajame pavyzdyje ir T2, ir T3 yra ontologijos);
- ontologija – tai ontologinės nuostatos specifikacija (pagal šią apibrėžtį, pateiktajame pavyzdyje ontologija yra tik T3).

Pastebėsime, kad abi šios apibrėžtys yra tiesiog Gruber pasiūlytos ontologijos apibrėžties detalizacijos.

Galima išvesti paralelę tarp dalykinės teorijos ir ontologijos (2 pav.): jei dalykinė teorija yra loginė teorija, papildyta dalykinėmis aksiomomis, tai ontologija yra loginė konceptų teorija, papildyta visuose pasauliuose teisingomis dalykinėmis aksiomomis bei (jei priimame antrąją ontologijos apibrėžtį) konceptų prasmės interpretavimą apribojančiomis aksiomomis. Taigi, vieną ontologiją atitinka kelios dalykinės teorijos, nes tai pačiai visuose pasauliuose teisingų aksiomų aibei galime parinkti daug skirtingų aksiomų aibių, aprašančių teiginius apie konkretaus pasaulio faktus (t.y. vieną *Tbox* atitinka daug skirtingų *Abox*).



2 pav. Ontologijos ir dalykinės teorijos sąryšiai.

5. Išvados

Priklausomai nuo reikiamo konceptualizacijos specifikavimo detalumo, arba ontologinė teorija, arba ontologinės nuostatos specifikacija gali būti traktuojama kaip ontologija. Pastarosios traktuotės privalumas yra detalesnis konceptualizacijos specifikavimas, trūkumas – turtingesnės kalbos poreikis, kas savo ruožtu reiškia, kad reikės sudėtingesnių ir mažiau efektyvių išvadų aparatų (angl. inference engines). Kai labiau svarbus yra ontologijos panaudojimas išvadų darymui, labiau tinkama yra pirmoji traktuotė.

Literatūra

- [1] J-P. Van Belle, A critique of current system engineering methods: the case for ontology-augmented methodologies, in: *Proceedings of the Workshop on Evaluation of Modeling Methods in Systems Analysis and Design*, K. Siau, W. Wand (Eds.), Crete, Greece (1996).
- [2] Y. Bishr, W. Kuhn, Ontology-based modelling of geospatial information, *Proceedings of 3rd AGILE Conference on Geographic Information Science*, Helsinki/Espoo, Finland (2000).
- [3] B.F. Chellas, *Modal Logic: An Introduction*, Cambridge University Press (1980).
- [4] Ch. Fellbaum (Ed.), *WordNet. An Electronic Lexical Database*, Bradford Books (1998).
- [5] M.R. Genesereth, N.J. Nilsson, *Logical Foundation of Artificial Intelligence*, Morgan Kaufmann, Los Altos, California (1987).
- [6] T.R. Gruber, *A Translation Approach to Portable Ontology Specifications*, Knowledge Systems Laboratory, Computer Science Department, Stanford University, Technical Report KSL 9271 Revised April 1993.
- [7] N. Guarino, Formal ontology in information systems, *Proceedings of FOIS'98*, Trento, Italy, IOS Press, Amsterdam (1998), pp. 3–15.
- [8] N. Guarino, P. Giaretta, Ontologies and knowledge bases: towards a terminological clarification, in: *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, N. Mars (Ed.), IOS Press, Amsterdam (1995), pp. 25–32.
- [9] N. Guarino, Ch. Welty, Conceptual modeling and ontological analysis, *AAAI-2000 Tutorial on Conceptual Modeling and Ontological Analysis (MP-2)* (2000).
- [10] V. Vasudevan, *Notes on ontologies, (and their relevance to service trading in an internet service market)*, Technical notes, Object Services and Consulting, Inc. DARPA coABS Agility Project Public Homepage (February 1999), <http://www.objs.com/agility/tech-reports/9902-ontology.html>
- [11] Y. Wand, R. Weber, An ontological model of an information system, *IEEE Transactions on Software Engineering*, **16**(11), 1282–1292 (1990).

Role of Formal Ontologies in Modern Information System Engineering

D. Čiukšys, A. Čaplinskas

Paper discusses the importance of role formal ontologies play in information system engineering, analyses different conceptions of formal ontology, analyses relations of conceptualization, ontology and domain theory, gives formal definition of conceptualization as intensional semantic structure. Definition of formal ontology, acceptable in field of information system engineering, is proposed.