

Lyginamosios analizės metodai informatikoje

Jelena GASPEROVIČ, Albertas ČAPLINSKAS* (MII)

el. paštas: j.gasperovic@algoritmusistemas.lt, alcapl@ktl.mii.lt

1. Įvadas

Metodika – esminis bet kurio tyrimo elementas. Skirtingai nuo tokių mokslų kaip, tarkime, matematika, informatikoje tyrimai naudojami ne tik moksliniame darbe, bet ir praktinėje veikloje, pavyzdžiui, atliekant dalykinių sričių analizę, parenkant projekto instrumentariją ir pan. Todėl tyrimų metodų, taip pat ir lyginamosios analizės metodų, reikšmė čia yra ypač didelė. Netinkamai atlikti tyrimai veda prie nepagrįstų išvadų ir, kalbant apie tyrimus komercinių projektų kontekste, gali sąlygoti netgi reikšmingus finansinius nuostolius. Vienok, šitą suvokia ne visi tyrėjai, taip pat ir ne visi Lietuvos informatikai. Netgi informatikos ir informatikos inžinerijos universitetinėse studijose, įskaitant ir doktorantūros studijas, metodikos klausimams skiriama nepakankamai dėmesio. Studentai su tyrimų metodikomis supažindinami atsitiktinai ir fragmentiškai, į tyrimo metodiką beveik neatsižvelgiama vertinant bakalaurų ir magistrų baigiamuosius darbus. Šio straipsnio tikslas – aptarti informatikoje naudojamų kokybinių lyginamosios analizės metodų ypatumus ir pasiūlyti tų metodų klasifikaciją.

2. Bendrieji lyginamosios analizės principai

Paprastai, lyginamoji analizė naudojama dviejų ar daugiau fenomenų (objektų, procesų ar procedūrų), gvildenamų iš analizės tikslų išplaukiančiais aspektais, panašumams ir skirtumams nustatyti. Siekiama rasti tam tikru požiūriu geriausius ar tinkamiausius egzempliorius analizuojamų fenomenų klasėje. Dažniausiai, taip pat bandoma išsiaiškinti panašumų bei skirtumų priežastis ir poveikį nagrinėjamiems fenomenams. Galutinis tikslas paprastai esti susisteminti ar suklasifikuoti nagrinėjamus fenomenus. Gali būti naudojami tiek kokybiniai (vertinamieji), tiek ir kiekybiniai lyginamosios analizės metodai. Kiekybiniai metodai tikslesni, tačiau jais pasinaudoti pavyksta gana retai, nes esti per daug sudėtinga surinkti visus reikiamus duomenis. Todėl šiame straipsnyje nagrinėsime tik kokybinius metodus.

Svarbus vaidmuo lyginamojoje analizėje tenka vadinamiesiems *lemiamiesiems* (*esmianiams*) analizuojamų fenomenų elementams. Tai elementai, apsprendžiantys fenomenų pobūdį. Greta lemiamųjų elementų fenomenai turi ir antrinius, darančius poveikį tik tam tikriems, individualiems atskiro fenomeno aspektams. Lemiamieji elementai siejami su fenomenų panašumais (bendrybėmis), antriniais – su jų skirtumais. Lemiamieji elementai paprastai esti susieti tam tikromis priklausomybėmis (ryšiais) arba, kitaip tariant, gali

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

būti aprašomi tam tikru šablonu. Todėl vienu iš svarbiausių lyginamosios analizės metodų yra fenomeno sugretinimo su šablonu metodas (angl. *pattern matching*). Šis uždavinys gali būti sprendžiamas daugeliu būdų, bene populiariausiu iš kurių yra atstumo įvertinimas atitinkamu būdu sukonstruotoje metrinėje erdvėje. Bendruoju atveju atstumas ir panašumas yra skirtingos priklausomybės, nes atstumai yra skaičiuojami daugiamatėje erdvėje, o panašumai gali būti vertinami tik pagal vieną iš erdvės metmenų. Panašumas, skirtingai nuo atstumo, gali būti ir netranzytīvus. Pavyzdžiui, dienis drėkinamasis kremas panašus į naktinį drėkinamąjį kremą (abu drėkinamieji), naktinis drėkinamasis kremas panašus į naktinį sausinamąjį (abu naktiniai), bet dienis drėkinamasis kremas neturi nieko bendro su naktiniu sausinamuoju kremu.

3. Panašumų ir skirtumų sugretinimas

Panašumų ir skirtumų sugretinimo metodą pasiūlė Milas [5]. Yra visa tokio pobūdžio metodų klasė. Atliekant panašumų išskyrimą, nustatomos lyginamųjų fenomenų bendrybės. Atliekant skirtumų išskyrimą, nustatomos jų skirtybės. Po to sudaroma lentelė, kurioje sugretinami skirtingų fenomenų panašumai bei skirtumai. Panašumų išskyrimas grindžiamas prielaida, kad jeigu du ar daugiau analizuojamos fenomenų klasės egzempliorių turi tik vieną bendrą savybę, tai pagal šią savybę tos klasės egzemplioriai yra panašūs. Skirtumų išskyrimas grindžiamas prielaida, kad jeigu du ar daugiau analizuojamos fenomenų klasės egzemplioriai turi tam tikrų bendrų savybių, o vieną iš savybių turi tik vienas iš egzempliorių, tai ši savybė ir yra analizuojamos klasės egzempliorių skirtybė.

Informatikoje panašumų ir skirtumų sugretinimo metodai yra naudojami, pavyzdžiui, dviem panašiausioms arba dviem labiausiai skirtingoms sistemoms rasti. Ieškant panašiausių sistemų, daroma prielaida, kad programų arba informacinės sistemos yra identiškos tam tikrų lemiamųjų (esminių) elementų požiūriu. Ieškant skirtingiausių sistemų, sulyginami besiskiriantys atvejai. Tai daroma, parodant ryšių stiprumą tarp sistemų priklausomų ir nepriklausomų kintamųjų. Nustatant skirtumus išryškėja antriniai sistemų elementai.

Panašumų ir skirtumų sugretinimo metodai turi keletą rimtų trūkumų. Visų pirma, vien tik išskiriant analizuojamų fenomenų panašumus ir skirtumus, nustatyti tinkamiausias savybes, praktiškai yra neįmanoma. Antra, nėra skiriamos fenomeno prigimtinės (būtiniosios) ir kontingentinės savybės, kitaip tariant, neleistina apibendrinami atskiri atvejai. Jeigu yra iškelta hipotezė ir rastas atvejis, kuris ją patvirtina, visumoje tai dar nieko neįrodo. Taigi panašumų ir skirtumų sugretinimo metodai nėra pakankamai objektyvūs.

4. Normatyvinė analizė

Normatyvinės analizės metodas, lyginant jį su panašumų ir skirtumų sugretinimo metodu, yra objektyvesnis. Jo esmė – parinkti etaloninį fenomeno egzempliorių (normą) ir visus kitus fenomenus vertinti nuokrypio nuo normos požiūriu.

Tegu $A = \{a_i\}$ – normos savybių aibė, $B = \{b_i\}$ – analizuojamo fenomeno savybių aibė, $R \subset A \times B$ – sugretinimo ryšys. Lyginant fenomeną su norma galima nustatyti šiuos fenomeno bruožus:

- **neišsamumą** (angl. *incompleteness*): kokių normoje numatytų savybių analizuojamasis fenomenas neturi, t.y.

$$\exists x \forall y ((x \in A) \& (y \in B) \& \neg R(x, y));$$

- **praturtinamumą**: kokiomis normoje numatytais savybėmis galima praturtinti jų neturintį analizuojamąjį fenomeną;

- **susiliejiama** (angl. *overload*): kokios normoje numatytos savybės analizuojamame fenomene pasireiškia kaip viena savybė, t.y.

$$\exists x_1, x_2, y ((x_1 \in A) \& (x_2 \in A) \& (y \in B) \& R(x_1, y) \& R(x_2, y));$$

- **pasikartojamumą** (angl. *redundancy*): kokios skirtingos analizuojamo fenomeno savybės normoje pasireiškia kaip viena savybė, t.y.

$$\exists x, y_1, y_2 ((x \in A) \& (y_1 \in B) \& (y_2 \in B) \& R(x, y_1) \& R(x, y_2));$$

- **pertekliškumą** (angl. *excess*): kokias perteklines (nenumatytas normoje) savybes turi analizuojamasis fenomenas, t.y.

$$\exists x \forall y ((x \in A) \& (y \in B) \& \neg R(x, y)).$$

Informatikoje normatyvinės analizės metodas naudojamas IS specifikuojimo kalbų ontologinei analizei. Pasiūlytos dvi normatyvinės ontologijos – Bunge [1] ir Chisholmo [2]. Bunge ontologija naudojama Bunge-Wando-Weberio (BWW) modeliuose [8]. BWW modeliai pateikia, autorių nuomone, išsamų konstrukcijų rinkinį, tinkamą bet kuriai IS specifikuoti. Lyginant specifikuojimo kalbas su šia norma, reikalaujama, kad kiekvieną BWW vaizdavimo modelio konstrukciją atitiktų bent viena vertinamos kalbos konstrukcija. Tačiau praktiniai eksperimentai parodė, kad ši metodika nėra pakankamai pagrįsta. Pavyzdžiui, kilo abejonių, ar norma nėra perteklinė, nes kai kurių jos savybių atitikmenų nebuvo rasta nei vienoje iš analizuotų specifikuojimo kalbų. Kita vertus, kai kurių aiškiai reikalingų savybių norma neturi. Pavyzdžiui, pasigendama kai kurių verslui specifikuoti būtinų konstrukcijų.

Chisholmo ontologija skiriasi konceptualizavimo priemonių rinkiniu. Ji yra artimesnė vadinamajai „sveikos nuovokos“ pasaulio sampratai. Tačiau svarbesnė negu šie skirtumai yra metodinė prielaida, jog skirtingi analitikai tam pačiam tikrovės fragmentui gali rašyti skirtingas specifikuojimo kalbas, nes skiriasi jų tikslai. Jiems gali prireikti ir skirtingų specifikuojimo priemonių. Todėl specifikuojimo kalbas siūloma vertinti ne jų turimų konstrukcijų, bet jomis specifikuojamų situacijų požiūriu. Lyginant su BWW metodika, šis būdas geriau atsižvelgia į specifinius konkrečios IS specifikuojimo poreikius, bet vis vien yra siekiama sukonstruoti „patį geriausią“ (etaloninį) tikrovės konceptualizavimo būdą.

5. Konceptinės analizės metodas

Pagrindinė metodo idėja – išskaidyti fenomenus į atskirus komponentus ir lyginti tik tarpusavyje sulyginamus, t.y. tos pačios funkcinės paskirties komponentus. Tuo tikslu sudaromi lyginamų fenomenų konceptiniai modeliai, aprašantys fenomenų struktūrą. Lyginami ne patys fenomenai, bet jų modeliai.

Informatikoje šis metodas naudojamas programų sistemų projektavimo metodikoms lyginti [7]. Jis gana objektyvus, nes lyginimas atliekamas formalios bazės pagrindu. Tai leidžia atkartoti vertinimus ir išitikinti jų objektyvumu. Vienok metodas taip pat turi rimtų trūkumų. Pirma, sudėtinga suskaidyti fenomenus į vienodus komponentus, nes jie dažniausiai esti skirtingai konceptualizuoti (pvz., dvi skirtingos metodikos yra aprašytos vartojant skirtingas sąvokų sistemas). Antra, funkcinės paskirties kriterijus yra nepakankamas. Vadovaujantis šiuo kriterijumi galima išsiaiškinti, kokias priemones turi, tarkime,

viena ar kita projektavimo metodika, bet visiškai ignoruojamas metodikos naudojimo kontekstas, o jos tinkamumas priklauso nuo projekto pobūdžio. Trečia, koncepcinis modelis aprašomas konkrečia modeliavimo kalba, o ši primeta tam tikrą ontologinį požiūrį, kas gali iškreipti lyginimo rezultatus [7].

6. Aspektinės analizės metodas

Pagrindinė metodo idėja – išskirti analizės tikslams relevantiškus fenomeno aspektus, įvertinti juos pagal pasirinktąją kokybinę skalę ir, panaudojant nuo išskirtųjų aspektų pobūdžio priklausančią formulę, paskaičiuoti suminį fenomeno įvertį.

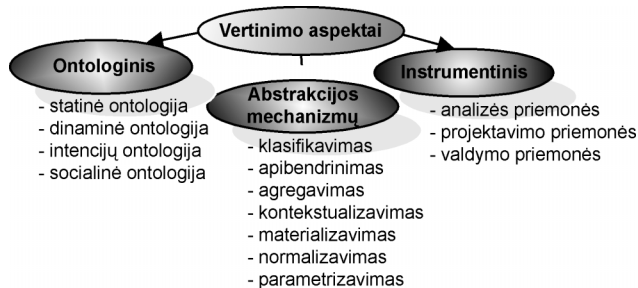
Informatikoje šis metodas naudojamas, pavyzdžiui, specifikuojant sistemoms vertinti [6]. Išskiriami ontologinis, abstrakcijos mechanizmų ir instrumentinis sistemos aspektai (1 pav.). Ontologinis aspektas toliau skaidomas į keturis aspektus, kiekvienas iš kurių aprašomas sava ontologija. Jis nusako kokius socialinės tikrovės reiškinius galima specifikuoti vertinamosios sistemos specifikuojamoje kalboje. Abstrakcijos mechanizmų aspektas nusako leistinus informacijos organizavimo būdus, instrumentinis – priemonės, skirtas padidinti specifikuojamą sudarančio asmens darbo našumą. Visi trys aspektai vertinami pasirinktoje kokybinėje skalėje. Pagal bendrąjį įvertį sprendžiama apie sistemos tinkamumą konkrečiam projektui.

Svarbiausieji metodo trūkumai yra du. Pirma, neaišku, kaip įsitikinti, kad išskirti visi relevantiški fenomeno aspektai ir kad jie yra pakankamai detalizuoti (pvz., ar 1 pav. siūlomas ontologijų rinkinys iš tiesų yra pakankamas?) Antra, įvertiniai yra subjektyvūs, gaunami neatliekant jokių matavimų [3].

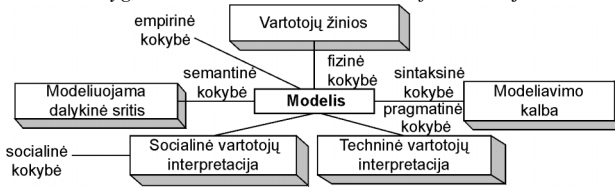
7. Lyginimas kokybės modelio pagrindu

Metodo esmė – sukonstruoti analizuojamų fenomenų klasės kokybės charakteristikų hierarchiją, vadinamąją kokybės modelį, ir vertinti tuo modeliu numatytas charakteristikas. Charakteristikoms įvertinti konstruojama kiekybinių arba kokybinių matų sistema.

Informatikoje tipiškas šio metodo pavyzdys yra Lindlando-Sindre-Solvbergo (LSS) karkasas [4], naudojamas specifikuojant kalboms vertinti. Jame kokybės modelis turi du lygmenis (2 pav.). Viršutiniame lygmenyje išskiriamos dvi kokybės atributų grupės, aprašančios kalbos konstrukcijas koncepciniu ir vaizdavimo aspektais. Antrojo lygmens atributai charakterizuoja kalbą dalykinės srities, vartotojų ir technologijos aspektais. Mūsų nuomone, LSS karkasas labiausiai priartėja prie tikslo sukurti objektyvią specifikuojamą



1 pav. Ontologinės analizės aspektai.



2 pav. LSS karkaso schema.

kalbų vertinimo metodiką. Tačiau čia siūlomas kokybės modelis nėra iki galo išbaigtas, neaišku kaip matuoti ir vertinti kokybės charakteristikas, ir kaip, vertinant kalbos kokybę, atsižvelgti į konkretaus projekto pobūdį [3].

8. Išvados

Išanalizavus ir apibendrinus informatikoje naudojamus lyginamosios analizės metodus, nustatyta, kad galima išskirstyti šešias svarbiausias tokių metodų klases: šablonų sugretinimo metodai, panašumų ir skirtumų sugretinimo metodai, normatyvinės analizės metodai, koncepcinės analizės metodai, aspektinės analizės metodai ir kokybės modelių grindžiami metodai. Kiekviena klasė turi savų privalumų ir savų trūkumų. Atliekant lyginamosios analizės tyrimus informatikos srityje (rašant magistro tezes, daktaro disertaciją, parenkant technologines konkretaus projekto priemones ir pan.), būtina kruopščiai išanalizuoti, kokios klasės metodai geriausiai tinka tiems tyrimams ir pritaikyti tuos metodus, atsižvelgiant į jiems būdingus ypatumus. Kitaip gali būti iškreipti viso tyrimo rezultatai.

Literatūra

- [1] M. Bunge, *Treatise on Basic Philosophy, Ontology I*, 3, Boston, Reidel, *The Furniture of the World* (1977).
- [2] R.M. Chisholm, *The basic ontological categories*, in: K. Muligan (Ed.), *Language, Truth, and Ontology*, Kluwer Academic Publishers (1992).
- [3] J. Gasperovič, A. Čaplinskas, *Informacinių sistemų specififikavimo kalbų lyginamosios analizės metodai*, *Informacijos mokslai*, 26, VU leidykla (2003).
- [4] O.I. Lindland, G. Sindre, A. Solvberg, *Understanding quality in conceptual modelling*, *IEEE Software*, 2(11), 42–49 (1994).
- [5] J.S. Mill, *System of Logic*, New York, Harper & Row, Publishers, 278 (1970).
- [6] J. Mylopoulos, *Characterizing information modeling techniques*, in: P. Bernus, K. Mertins, G. Schmidt (Eds.), *Handbook on Architectures of Information Systems*, Berlin/Springer (1998), pp. 17–57.
- [7] X. Song, L.J. Osterweil, *A framework for classifying parts of software design methodologies*, in: R.W. Selby (Ed.), *Proceedings 2nd Irvine Software Symposium IRUS*, March 1992 (1992), pp. 49–68.
- [8] Y. Wand, R. Weber, *An ontological evaluation of systems analysis and design methods*, in: E.D. Falkenberg, P. Lindgreen (Eds.), *Information Systems Concepts: An In-depth Analysis*, North-Holland, Amsterdam (1989), pp. 79–100.

Comparative analysis methods in informatics

J. Gasperovič, A. Čaplinskas

The paper discusses comparative analysis methods in informatics. It examines nature of each class of methods and presents examples of its application. The purpose of the paper is to propose a classification of comparative analysis methods.