

VERSLO SISTEMOS LYGMENS TAISYKLIŲ TRANSFORMAVIMAS Į INFORMACIJOS APDOROJIMO LYGMENS TAISYKLES

Jurgita Rozancevaitė¹, Olegas Vasilecas²

Vilniaus Gedimino technikos universitetas

El. paštas: ¹ vilnius.jurgita@gmail.com; ² olegas.vasilecas@vgtu.lt

Santrauka. Nagrinėjama verslo taisyklių rinkinio išgavimo iš verslo taisyklių ir šių taisyklių paruošimo realizavimui problema. Apžvelgti taikomi taisyklių rinkinių išgavimo metodai, atsižvelgiant į jų privalumus bei trūkumus. Analizuoti pagrindiniai taisyklių užrašymo principai, pagal kuriuos būtų galima sudaryti taisyklių užrašymo šablonus. Sumodeliuotas įrankis, kuris naudojantis verslo objektų žodyną, padėtų struktūrizuoti anglų kalba užrašytame verslo dokumente aptikti galimas taisykles ir padėtų užrašyti šias taisykles apibrėžtu formatu, kurį galima realizuoti programiniame kode. Atlikti eksperimentai parodė, kad įrankis leidžia atpažinti ir užrašyti apibrėžtu formatu apie 80 proc. taisyklių, esančių verslo dokumente.

Reikšminiai žodžiai: verslo taisyklė, reikalavimų identifikavimo metodai, verslo taisyklės šablonas.

Įvadas.

Nei vienas verslas negali egzistuoti be verslo taisyklių. Šaltiniai, iš kurių gaunamos verslo taisyklės, yra skirtingi: verslo atstovai, teisės aktai, vyriausybės įstatymai, verslo politika. Programų sistemos kūrimas yra seka vykdomų transformacijų, kurios prasideda reikalavimų identifikavimu ir baigiasi realizavimu. Problema yra ta, kad rašytiniuose dokumentuose taisyklės užrašytos natūralia kalba sudėtiniais sakiniais, nesinaudojant jokių formatu. Toks tekstinis užrašymas gerai suprantamas verslo atstovams, tačiau ne kompiuteriams. Tokius reikalavimus sunku analizuoti. Net ir nustačius, kad tai yra verslo taisyklė, gali būti sudėtinga automatiškai transformuoti į informacijos apdorojimo taisykles, ypač jei ji yra užrašyta kaip sudėtinis sakiny. Analizuojant tokį tekstą, stengiamasi nustatyti verslo taisykles išskaidyti į atomines taisykles.

Paskutiniu metu ieškoma metodų, kurie kuo patogiau ir su kuo mažiau resursų padėtų aptikti ir transformuoti tekstinius reikalavimus į formą, kurią būtų galima realizuoti.

Reikalavimų transformacijos metodai.

Reikalavimai gali transformuojami į analizės modelį tiesiogiai arba per tarpinį modelį. Gali būti

taikomos keturių tipų transformacijos: taisyklių, ontologinė, tapatumo ir šabloninė.

Literatūroje dažniausiai naudojama taisyklių transformacija, kuri naudoja iš anksto apibrėžtas transformavimo taisykles.

Ontologija yra padalintas žodynas, aprašantis su žinoma veikimo sritimi susijusias sąvokas; jos semantika yra apibrėžta nedviprasmiška ir kompiuteriniu būdu vykdoma forma. Ontologijos modelis kuriamas tada, kai yra realizuojami natūralios kalbos sakiniai. Šis ontologijos modelis veikia kaip tarpinis modelis, kuris vėliau transformuojamas į analizės modelį. Tokio tipo transformacija vadinama ontologine transformacija. Paprastai ši informacija reikalauja daug palaikymo resursų.

Tapatumo transformacija transformuoja pradinį (šaltinio) modelį į kitą (tikslą) modelį nekeisdama turinio informacijos: du modeliai aprašo tuos pačius konceptus, bet juos vaizduoja skirtingai. Šio tipo transformacija gali būti taikoma kaip papildoma priemonė.

Šabloninė transformacija transformuoja šaltinio šablonus į tikslo šablonus. Šaltinio šablonas aprašo ir jungia šaltinio elementų aibę, o tikslo šablonas aprašo ir jungia tikslo elementų aibę.

Taisyklių identifikavimo metodai.

OMG apibrėžė žodyną, taisyklės verslo žodyną ir taisykles verslo žodyną, verslo faktų ir verslo

taisyklių semantikai dokumentuoti (OMG. 2008). Taisyklių apibrėžimui jie siūlo naudoti struktūrizuotą anglų kalbą. Pateikiamas sąrašas naudotinių kvantorių (pavyzdžiui, kiekvienas, bent vienas), loginių operacijų (pavyzdžiui, ir, arba), modalinių operacijų (pavyzdžiui, privaloma, draudžiama, niekada).

Technikos, skirtos transformuoti reikalavimus užrašytus tekstone kalba, dažniausiai taiko pradinio apdorojimo procesą. Paprastai taikomi penkių pradinio apdorojimo technikų tipų algoritmai, kurie gali būti naudojami tiek atskirai, tiek bendrai: leksikos analizė, sintaksės analizė, semantinė analizė, skaidymas į kategorijas ir pragmatinė analizė (Bajwa et al. 2009, Yue et al. 2009).

Leksikos analizė, dar vadinama sakinio dalių formavimu, yra procesas perverčiantis simbolius į sakinio dalis. Šį procesą sudaro etapai: sakinio į dalis skaidymas, sakinių suskaldymas, kalbos dalies žymėjimas ir morfologinė analizė. Sakinio į dalis skaidymas naudojamas žodžiams ir skyrybai atskirti bei skaičiams identifikuoti. Sakinių suskaldymas apibrėžia sakinio ribas duotame tekste. Kalbos dalies žymėjimas identifikuoja daiktavardžius, veiksmažodžius, būdvardžius. Morfologinė analizė apibrėžia žodžių struktūrą ir transformaciją.

Sintaksės analizė, dar vadinama sintaksiniu nagrinėjimu, yra sakinio dalių skaidymo sekos analizavimo procesas, nusakantis gramatinę struktūrą atsižvelgiant į duotą formalią gramatiką. Paprastai išvestis yra sintaksinio nagrinėjimo medis. Jei tekstas užrašytas struktūrizuota anglų kalba, tai šiame etape pervedama iki galiojančių frazių ir sakinio pagal anglų kalbos gramatines taisykles. Be to, šis etapas padeda identifikuoti pagrindines sakinio dalis: objektą, subjektą, veiksmus, atributus.

Semantinė analizė yra procesas, pridantis semantinės informacijos į nagrinėjamą medį, dažniausiai naudodamas srities specifinę informaciją (terminų žodynas, apibrėžtys ar srities modelis) (Yue et al. 2009). Iš kitos pusės, šis etapas nusako konkretaus objekto vykdomus veiksmus ir atributų rinkinį, priklausantį šiam objektui (Bajwa et al. 2009).

Skaidymas į kategorijas yra procesas, kurio metu reikalavimai yra atpažįstami, atskiriami ir suklasifikuojami pagal tikslus. Paprastai šis procesas vykdomas rankiniu būdu.

Pragmatinė analizė pašalina reikalavimų dviprasmiškumą ir prieštarumą. Pavyzdžiui, pragmatinė analizė gali būti naudojama patikrinti

naujos informacijos darną prieš ją fiziškai patalpinant su kitais egzistuojančiais reikalavimais.

Šie metodai patogūs tekstiniais dokumentams, kurie yra užrašyti natūralia anglų kalba arba struktūrizuota anglų kalba, nes taisyklių identifikavimui naudoja anglų kalbos gramatikos taisykles.

Taisyklių ribojimai.

Norint išvengti taisyklių persidengimo, taikomi ribojimai natūralia kalba parašytam tekstui. Natūralios kalbos ribojimų taisykles galima klasifikuoti į tris ribojimų tipus: sakinio, sakinio struktūros ir žodžių.

Sakinys yra žodžių grupė, kurie kartu sudaro reikšmingą vienetą ir yra tikimasi, kad turi subjektą ir veiksmažodį. Pavyzdžiui, veiksmažodžio laikų pasirinkimų ribojimai, daiktavardžio vienaskaitos ar daugiskaitos formų pasirinkimai yra laikomi sakinio ribojimais.

Sakinio struktūros ribojimai įveda ribojimus sudėtinio sakinio struktūrai. Sudėtinį sakinį sudaro daug sakinių. Šie sakiniai jungiami vienas su kitu jungtimis, skyrybos ženklais arba abiem.

Žodžių ribojimas riboja žodžių pasirinkimą ir jų naudojimo būdą. Pavyzdžiui, „naudoti tik „būti“ arba „tapti“ reikšminius žodžius apibendrinimo ryšiui tarp sakinio subjekto ir objekto išreikšti.

Pritaikius šiuos ribojimus, lengviau suprasti bei realizuoti užrašytas taisykles. Ribojimai suteikia taisyklėms aiškumo.

Formalios logikos taisyklių šablonas.

Dalį taisyklių galima aprašyti formalia logika, samprotaujant teiginiais (Vasilecas et al. 2008). Formalios logikos sakiniai formuojami naudojant: konstantas, predikatus, funkcijas, kintamuosius, logines jungtis ir operacijas bei kvantorius. Norint verslo taisykles aprašyti formalia logika, turi būti taikomi sakinio struktūros ribojimai. Formalia logika užrašytoms taisyklėms lengviau pritaikyti struktūros šabloną.

Nustatyta, kad kiekvienai verslo taisyklei yra reikšmingos keturios dalys: terminas, pavadinimas, veiksmažodis ir reikšminis žodis (OMG. 2008).

Terminas žymi daiktavardinį konceptą, tai kas yra naudojama ir apibrėžta kaip verslo objektas verslo objektų žodyne. Terminai turi būti apibrėžiami

vienaskaitine forma, tuo tarpo daugiskaitinės formos yra suprantamos kaip daugybinis panaudojimas.

Pavadinimas naudojamas išreikšti individualų konceptą – pasireiškimą. Pavadinimai dažniausiai būna tikriniai daiktavardžiai. Lygiai taip pat išreiškiamos ir skaitinės pavadinimo reikšmės formaliuose nuostatuose (pavyzdžiui, 25).

Veiksmažodis išreiškia fakto tipus – paprastai veiksmažodžius, prielinksnius arba jų abiejų kombinaciją. Kiekviena išraiška yra apibrėžiama kaip fakto tipo formos kontekstas. Fakto tipo formoms naudojamos vienaskaitinės, aktyvios veiksmažodžių formos su tam tikromis išimtimis. Bendratys, tariamoji nuosaka, neveikiamoji forma ir daugiskaitinės veiksmažodžių formos taip pat naudojamos nuostatuose ir aprašuose.

Reikšminis žodis yra naudojamas išreikšti lingvistinius simbolius, naudojamus nuostatams konstruoti. Reikšminiai žodžiai skirstomi į kvantorius, loginius operatorius, modalinius operatorius ir kita (jungtukai ir pan.).

Natūralia anglų kalba užrašytas taisyklės formuluojant formalia logika, galima suteikti tokį taisyklės struktūros šabloną:

<Modalinis operatorius> <Kvantorius> <Terminas>
<Veiksmažodis> <Kvantorius> <Terminas> (1).

Kur modalinis operatorius nėra griežtai privalomas. Tokiu atveju, taisyklės struktūros šablonas:

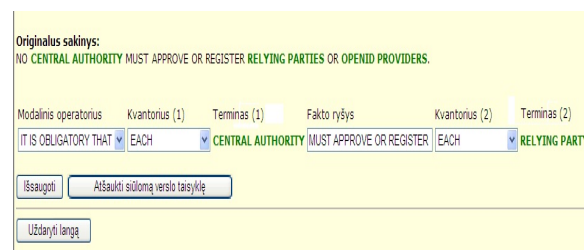
<Kvantorius> <Terminas> <Veiksmažodis>
<Kvantorius> <Terminas> (2).

Automatizuotas verslo taisyklių identifikavimas.

Realizuotas įrankis, kuris padeda struktūrizuoti anglų kalba užrašytoje specifikacijoje identifikuoti galimas verslo taisykles. Pirmu etapu nustatomas verslo žodynas, identifikuojami reikšminiai verslo terminai. Antru etapu atliekama nurodyto teksto analizė, taikant pradinio apdorojimo proceso algoritmus. Tekste išryškinami visi sakiniai, kuriuose pasitaikė reikšminiai verslo žodyno terminai.

Kitu etapu pasirenkamas konkretus sakinytis ir pusiau automatizuotu būdu jis transformuojamas taip, kad atitiktų apibrėžtą taisyklės struktūros šabloną (1 pav.). Šio etapo metu, taikomi sakinio struktūros ribojimai. Iš vienos verslo taisyklės galima sudaryti vieną ar daugiau realizuotinių taisyklių. Jei randama daugiskaitinė termino forma, ji automatiškai perverčiama į vienaskaitinę, kartu pasiūlant atitinkamą

kvantorių. Veiksmažodis ar veiksmažodinė frazė tarp dviejų terminų suprantamas kaip fakto ryšys.



1 pav. Verslo taisyklės identifikavimas

Paskutiniu etapu iš visų suformuluotų ir patvirtintų taisyklių sudaromas rinkinys, kurį galima eksportuoti ir pritaikyti vėlesniuose informacinės sistemos kūrimo etapuose.

Naudojantis šiuo įrankiu, verslo taisyklių atpažinimas tuo efektyvesnis, kuo gausesnis ir tikslesnis verslo terminų žodynas.

Išvados.

1. Išnagrinėjus siūlomus verslo taisyklių užrašymo būdus, nustatyta, kad konstruktyviu būdu apibrėžtas taisyklės lengviau realizuoti, nes jos pasižymi geros kokybės taisyklės savybėmis (svarbi, atominė, deklaratyvi, darni ir kt.) ir turi apibrėžtą formalizuotos kalbos šablonų formą, kuri leidžia tokias taisykles realizuoti kode.

2. Remiantis atliktais eksperimentais, nustatyta, kad naudojant pradinio apdorojimo procesą struktūrizuota anglų kalba užrašytose specifikacijose, automatizuotu būdu galima identifikuoti 80 proc. verslo taisyklių, esančių dokumente, bei jas paruošti realizavimui.

Literatūra.

Bajwa, I. S.; Samad, A; Mumtaz, S. 2009. Object Oriented Software Modeling Using NLP Based Knowledge Extraction. *ISSN 1450-216X Vol.35 No.1(2009), pp 22-33.* [interaktyvus] [žiūrėta 2010 m. balandžio 15 d.]. Prieiga per internetą:

<http://www.eurojournals.com/ejsr_35_1_03.pdf>

Yue, T; C. Briand, L; Labiche, Y. 2009. A Systematic Review of Transformation Approaches between User

Requirements and Analysis Models. *Carleton universitetas, TRSC09-03*.

OMG. 2008. Semantics of Business Vocabulary and Business Rules (SBVR), v1.0. *OMG dtc/07-09-04, formal/2008-01-03*. [interaktyvus]. [žiūrėta 2010-05-17]. Prieiga per internetą:

<<http://www.omg.org/spec/SBVR/1.0/PDF>>

Vasilecas, O.; Saulis, A. 2008. *Informacinių sistemų kūrimo metodai ir technikos*. Klaipėdos universiteto leidykla.

THE RULES OF BUSINESS SYSTEM LAYER TRANSFORMATION TO THE RULES OF INFORMATION PROCESSING LAYER

J. Rozancevaitė, O. Vasilecas

Abstract

The problem of business rules extracting from business rules and those extracted preparing for realization has been analyzed. Applicable sets of rules extraction methods reviewed, according to their strengths and weaknesses. Basic written principles of rules, that according to templates could be built, were analyzed as well. The tool, using business objects dictionary, that could help to detect possible rules in provided business document, written in structured English language, has been simulated. That tool would also help to build rules in defined rules format that could be realized in programming code. Performed experiments showed that the tool allows to identify and record in defined format 80 percent of rules, found in the business document.

Keywords: business rule, taxonomy of requirements, the business rules pattern.