VILNIAUS UNIVERSITETAS MATEMATIKOS IR INFORMATIKOS FAKULTETAS INFORMATIKOS KATEDRA

Funkcinio-reaktyvaus programavimo taikymas įvykių kaupimo sistemose

Functional Reactive Programming in Event Sourcing Systems

Mokslo tiriamasis darbas

Atliko: 1 kurso / grupes	studentas	
Žilvinas Kučinskas	(parašas)	
Oarbo vadovas: Viačeslav Pozdniakov	(parašas)	
Recenzentas: orof. Rimantas Vaicekaus	kas(parašas)	

Vilnius, 2014 m.

Turinys

1.	1. Magistro darbo objekto apžvalga bei tyrimo problemos aprašymas		
	1.1.	Tyrimo objektas	3
	1.2.	Darbo tikslai ir uždaviniai	Ç
	1.3.	Tyrimo aktualumas	
	1.4.	Pritaikymo pavyzdys	٦
	1.5.	Tyrimo metodika	6
		Laukiami rezultatai	
2.	Literat	ūros analizė	8
	2.1.	Funkcinis-reaktyvus programavimas	8
		2.1.1. Įvadas	8
		2.1.2. Pagrindinis tiklas	8
		2.1.3. Sąvokos	8
		2.1.4. Semantika	Ć
	2.2.	Įvykių kaupimas	Ć
		2.2.1. Įvadas	Ć
		2.2.2. Momentinė kopija	11
		2.2.3. Įvykių kaupimo privalumai ir trūkumai	11
		2.2.4. Įvykių srautas	12
		2.2.5. Įvykių kaupimas funkciniame programavime	12
	2.3.	Monados	13
	2.4.	Išvados	13
3.	Sąvokų	apibrėžimai	14
4.	Santrui	mpos ir paaiškinimai	15
Lit	eratūro	s caraças	16

Magistro darbo objekto apžvalga bei tyrimo problemos aprašymas

1.1. Tyrimo objektas

Tyrimo objektas yra funkcinio-reaktyvaus programavimo bei įvykių kaupimo principai.

1.2. Darbo tikslai ir uždaviniai

Darbo tikslas yra išnagrinėti funkcinio-reaktyvaus programavimo pritaikymo galimybes įvykių kaupimo sistemose.

Siekiant šio tikslo, turi būti išspręsti šie uždaviniai:

- įrodyti, kad funkcinį-reaktyvų programavimą įmanoma taikyti įvykių kaupimo sistemose;
- sukurti konkretizuotą kalbą (angl. domain specific language), apjungiančią funkcinio-reaktyvaus programavimo bei įvykių kaupimo principus;
- aprašyti konkretizuotuos kalbos kūrimo metodiką, apibrėžti gautų rezultatų apribojimus, suformuluoti iškilusias problemas bei paaiškinti jų priežastis.

1.3. Tyrimo aktualumas

Funkcinis reaktyvus programavimas integruoja laiko tėkmę bei sudėtinius įvykius į funkcinį programavimą. Šis principas suteikia elegantišką būdą išreikšti skaičiavimus interaktyvių animacijų, robotikos, kompiuterinio vaizdavimo, vartotojo sąsajos ir modeliavimo srityse [Cza12, p. 4]. Pagrindinės funkcinio reaktyvaus programavimo sąvokos:

- signalai arba elgsena reikšmės, besikeičiančios bėgant laikui;
- įvykiai momentinių reikšmių kolekcijos arba laiko-reikšmės poros.

Funkcinis-reaktyvus programavimas įgalina apsirašyti elgseną deklaratyviai naudojant imperatyvios programavimo kalbos struktūras [EH97, p.1]. Elgsena ir įvykiai

gali būti komponuojami kartu, išreikšti vienas per kitą. Funkcinis reaktyvus programavimas apibrėžia kaip signalai arba elgsena reaguoja į įvykius. [Ams, p. 1] Šį principą galima iliustruoti pavyzdžiu. Tarkime turime Excel lapą, kuriame yra trys laukai: darbuotojo pradirbtos valandos, valandinis užmokestis bei formulė, kuri paskaičiuoja konkretų darbuotojo atlyginimą. Darbuotojui pradirbus daugiau valandų, atnaujinamas pradirbtų valandų skaičius. Kartu atsinaujina ir pačios formulės reikšmė, tai yra konkretus užmokestis pakinta. Šiuo atveju įvykus reikšmės atnaujinimo įvykiui, nuo jos priklausomos formulės taipogi atsinaujina arba tam tikras įvykis iššaukia elgseną sistemoje.

Įvykių kaupimo principo esmė – objektas yra atvaizduojamas kaip įvykių seka. Kaip pavyzdį tai galima parodyti remiantis banko sąskaita. Tarkime vartotojas, banko klientas, turi 100 litų sąskaitos balansą. Tarkime vartotojas nusipirko prekę už 20 litų, tada įnešė į savo sąskaitą 15 litų ir galiausiai nusipirko tam tikrą paslaugą už 30 litų. Akivaizdu, jog turint šią įvykių seką, galima atvaizduoti dabartinę objekto būseną - tai yra 65 litai vartotojo sąskaitoje. Įvykių kaupimo principas užtikrina, jog visi būsenos pasikeitimai yra saugomi įvykių žurnale kaip įvykų seka [Ver13b]. Įvykių kaupimo principui yra būdinga, jog įvykių negalimą ištrinti bei atnaujinti, duomenys yra nekeičiami, dėl to įvykių žurnalas yra sistemos gyvavimo istorija (tiesos šaltinis). Tačiau toks modelis turi ir trūkumų. Jis nėra pritaikytas patogiam užklausų rašymui. Iš įvykių srautų yra kuriamos projekcijos, skirtos konkretiems sistemoms poreikiams, pavyzdžiui: paieškai, klasifikacijai ar ataskaitų ruošimui.

Pritaikius funkcinį-reaktyvų programavimą įvykių kaupimo principu paremtose sistemose būtų galima modeliuoti ne tik momentinius įvykius, tačiau turėti ir jų istoriją. Yra poreikis sukurti konkretizuotą kalbą (angl. domain specific language), kuri įgalintų paslėpti įvykių žurnalą (arba duomenų saugyklą). Pastarosios naudotojas galėtų orientuotis į pačią sprendžiamos srities problemą, nekreipdamas dėmesio į žemesnio lygio realizacijos detales. Šiuo atveju viename faile būtų galima deklaratyviai (ką kažkuri programos dalis turi daryti) apsirašyti elgseną, nutikus įvykiui, kartu su imperatyviomis(instrukcijos, kurios aprašo, kaip programos dalys atlieka savo užduotis) struktūromis.

1.4. Pritaikymo pavyzdys

Tarkime turime domeno sritį - bankininkystė. Turime įvykių srautą - vartotojų sukūrimas. Naudojant įsivaizduojamą Scala API galima sukurti vartotojų paieškos puslapį pagal vardą ir pavardę.

```
// stream model
case class CustomerCreate (val name: String, val surname: String, val personalNum:
   String)
val es = EventSourceConnection(url)
val createStream = Stream(es, "customerCreate")
class case CustomerModel(val name: String, val surname: String, val personalNum:
   String) extends ViewableModel
trait CustomerArgModel extends Arg2Model[String,String]{
 val name: Option[String]
 val surname: Option[String]
}
//args are passed on form view/post
val customerView = View(args: CustomerArgModel).foldLeft(
 (acc, event) => event match {
   case CustomerCreate(name, surname, personalNum) =>
     for {
       newName < - name
       newSurname < - surname
       newPersNum <- personalNum
       if (args.name == name && args.surname == surname)
     } yield CustomerModel(newName, newSur, newPersNum)
 }
// getting data for all Kucinskai
val specificData = customerView(None, Some("Kucinskas")):
   Option[List[CustomerModel]]
```

Šiuo atveju veiksmai peržiūrint duomenis yra sumaišomi kartu su veiksmais gaunant įvykius. Verta pastebėti, jog lokali duomenų saugykla nebuvo paminėta arba apibrėžta. Pastaroji gali būti sugeneruota bei valdoma automatiškai.

Antruoju pavyzdiniu atveju turimas vartotojo balanso(įplaukiančios/išplaukiančios lėšos) įvykių srautas ir norima gauti vartotojo, kurio asmens kodas yra 39008226547, einamosios savaitės sąskaitos balansą.

Šiuo atveju *event* kintamasis yra įvykių srauto monada (terminas vartojamas funkciniame programavime, kilęs iš kategorijų teorijos ir turi savas taisykles), o *filtered* kintamasis - duomenų saugyklos monada. Bendruoju atveju skirtingos monados tarpusavyje nesiderina [KW92]. Dėl to reikia išsiaiškinti - ar yra įmanoma ir kaip šias monadas suderinti.

1.5. Tyrimo metodika

Darbo analitinėje dalyje bus naudojami tradiciniai bibliotekinio tyrimo metodai. Darbo tikslui pasiekti tiriamojoje dalyje bus pasirinkta konkreti funkcinė programavimo kalba (pvz.: Haskell, Scala) bei aprašoma kūrimo metodika.

Papildomai I semestro metu ketinama užbaigti Erik Meijer, Martin Odersky, Roland Kuhn dėstomą reaktyvaus programavimo kursą (Coursera). Esant galimybėms žadama apsilankyti į funkcinį programavimą orientuotose konferencijose Scala Days (http://scaladays.org/), Scalar (http://scalar-conf.com/) ar kitose.

1.6. Laukiami rezultatai

Magistrinio darbo metu planuojama išnagrinėti funkcinio-reaktyvaus programavimo ir įvykio kaupimo principus, įrodyti, jog šie principai gali būti panaudoti kartu bei suderinti, sukurti konkretizuotą kalbą (angl. domain specific language), apjungiančią šiuos principus, bei aprašyti kūrimo eigos metodiką, apibrėžti gautus rezultatus, suformuluoti apribojimus, iškilusias problemas bei paaiškinti jų priežastis.

2. Literatūros analizė

2.1. Funkcinis-reaktyvus programavimas

Daugiausia remtąsi [Ams].

2.1.1. Įvadas

Funkcinis-reaktyvus programavimas, toliau vadinamas tiesiog FRP, yra būdas modeliuoti reaktyvų - besikeičiančius laiko tekmėje bei reaguojančius į išorinį stimulą - elgesį visiškai funkcinėse programavimo kalbose. FRP leidžia deklaratyviu ir paprastu būdu modeliuoti sistemas, kurios turi reaguoti į duomenis bėgant laikui.

2.1.2. Pagrindinis tiklas

Pagrindinis funkcinio-reaktyvaus programavimo tikslas:

- saugus programavimas kompiliatorius turi kiek įmanoma patikrinti programų korektiškumą;
- efektyvus programavimas programos turėtų veikti realiu laiku, todėl efektyvios ir optimizuotos operacijos yra būtinos;
- komponavimas FRP leidžia kurti programas iš smulkesnių programų, o ne orientuotą į problemą, vientisą kodą.

2.1.3. Sąvokos

Pagrindinės FRP sąvokos yra:

- signalai arba elgsena besikeičiančios laike reikšmės;
- įvykiai kolekcija momentinių reikšmių arba laiko-reikšmės poros.

FRP pasiekia reaktyvumą naudodamas konstrukcijas, kurios tiksliai apibrėžia kaip signalai arba elgsena pasikeičia reaguodami į įvykius. Tai yra pagrindinis būdas išreiškiant bei realizuojant elgseną. Kitu būdu, elgsena gali būti laiko semantinės funkcijos¹, kuriose laikas yra pakeičiamas kilus įvykiui.

¹http://msdl.cs.mcgill.ca/people/tfeng/docs/as/node5.html

2.1.4. Semantika

2.2. Įvykių kaupimas

Šiame skyriuje yra aprašomos žinios apie įvykių kaupimą, pliusus ir minusus, įvykių srautus bei įvykių kaupimą funkciniame programavime remiantis Vaughn Vernon surinkta ir aprašyta informacija [Ver13a].

2.2.1. Įvadas

Kartais verslui svarbu fiksuoti objekto pasikeitimus domeno modelyje². Šiuos pasikeitimus galima stebėti skirtingais būdais. Įprastai yra pasirenkama stebėti kai esybė³ yra:

- sukurta,
- paskutinį kartą modifikuota
- bei kas atliko modifikaciją.

Tačiau šis būdas nepateikia jokios informacijos apie vienkartinius pasikeitimus.

Atsiradus poreikiu stebėti pasikeitimus detaliau, verslas reikalauja dar daugiau metaduomenenų⁴, ko pasekoje tokie faktai kaip individualios operacijos laiko tekmėje bei jų įvykdymo laikas tampa svarbūs. Šie poreikiai verčia įvesti audito žurnalą fiksuoti labai tikslias panaudojimo atvejų metrikas, tačiau pastarasis būdas turi apribojimų. Jis gali atskleisti dalį informacijos apie tai kas nutiko sistemoje, leisti rasti bei ištaisyti dalį riktų bei klaidų programinėje įrangoje, bet audito žurnalas neleidžia patikrinti domeno objekto būsenos prieš ir po tam tikrų pasikeitimų. O jeigu būtų galima išgauti daugiau informacijos iš pasikeitimų stebėjimo?

Visi programinės įrangos kūrėjai susiduria su labai tiksliu pasikeitimų stebėjimu. Įprastas ir populiarus pavyzdys yra išeities kodo saugyklos, tokios kaip CVS⁵, Subversion⁶, Git⁷ arba Mercurial⁸. Visos šios pataisų valdymo sistemos leidžia stebėti

²http://en.wikipedia.org/wiki/Domain model

³http://en.wikipedia.org/wiki/Entity

⁴http://en.wikipedia.org/wiki/Metadata

⁵http://www.nongnu.org/cvs/

⁶http://subversion.apache.org/

⁷http://git-scm.com/

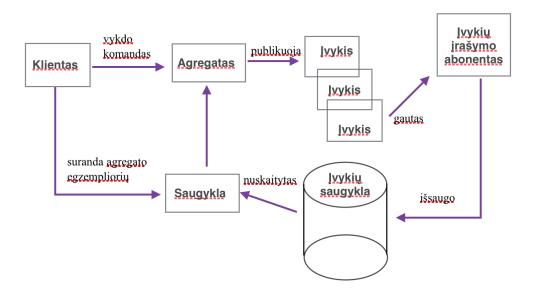
⁸http://mercurial.selenic.com/

pirminių failų pasikeitimus. Įrankiai leidžia peržiūrėti išeities kodo artefaktus nuo pačios pirmosios pataisos iki paskutinės. Kai visi išeities failai yra nusiunčiami į pataisų kontrolės sistemą, ši gali stebėti pasikeitimus viso programinės įrangos kūrimo gyvavimo ciklo metu.

Jeigu šis principas būtų pritaikytas vienai esybei, tada vienam agregatui⁹ bei galiausiai kiekvienam modelio agregatui, galima suprasti kokią naudą atneša sistemos objektų pasikeitimų stebėjimas:

- Kas būtent nutiko modelyje, jog agregato egzempliorius buvo sukurtas?
- Kas nutiko agregato egzemplioriui bėgant laikui? (Operacijų požiūriu)

Turint visų atliktų operacijų istoriją, galima palaikyti laikinus modelius. Toks kaitos stebėjimas yra įvykių kaupimo principas. 1 diagramoje pateika šio principo aukšto lygio reprezentacija. Agregatai publikuoja įvykius, kurie yra išsaugomi įvykių saugykloje ir naudojami sekti modelio būsenos pasikeitimus. Verta paminėti, jog įvykiai reprezentuoja tam tikrą būsenos pasikeitimą bėgant laikui, todėl jie nėra atnaujinami arba ištrinami. Saugykla nuskaito įvykius iš įvykių saugyklos ir pritaiko juos vieną po kito taip atkurdama agregato būseną.



1 pav. Įvykių kaupimo aukšo lygio reprezentacija

⁹http://martinfowler.com/bliki/DDD_Aggregate.html

2.2.2. Momentinė kopija

Ilgame periode sistemoje susikaupia daugybė įvykių. Atkuriant agregato būseną reikia atkartoti šimtus, tūkstančius ar net milijonus įvykių. Tai tampa šio modelio silpnąja puse, nes įvykių atkartojimas užtrunka vis ilgiau sistemai plečiantis.

Tačiau šio duomenų kamsčio galima išvengti naudojant agregato būsenos momentines kopijas. Tam tikrame įvykių saugyklos istorijos taške yra padaroma agregato būsenos kopija. Serializuota agregato būsena yra įrašoma į įvykių saugyklą. Nuo to momento, agregatas yra atkuriamas pirmiausia naudojantis naujausia jo būsenos momentine kopija ir tik po to atkartojami visi naujesni įvykiai.

Momentinės kopijos nėra atkuriamos atsitiktinai. Jos gali būti kuriamos kas apibrėžtą skaičių įvykių. Šis skaičius turėtų būti parinktas analizuojant domeno sritį bei sistemą ir radus optimalų variantą. Tikėtinai gali būti 50 arba 100 įvykių tarp momentinių kopijų.

2.2.3. Įvykių kaupimo privalumai ir trūkumai

Kaip saugojimo mechanizmas, įvykių kaupimas stipriai skiriasi ir pakeičia ORM¹⁰ įrankį. Kadangi įvykiai dažnai įrašomi kaip dvejetainės reprezentacijos, jie negali būti optimaliai naudojami užklausoms atlikti. Faktiškai įvykių kaupimu pagrįstoms saugykloms tereikia vienos operacijos - gauti įrašus pagal unikalią agregato tapatybę. To pasekoje užklausom daryti reikia kito kelio. Dažniausiai tam pasirenkamas CQRS¹¹ principas.

Įvykių kaupimas verčia kitaip mąstyti apie domeno modelį. Įvykių istorija gali padėti surasti bei ištaisyti sistemos defektus bei klaidas. Derinimas naudojant istoriją visų veiksmų, kurie nutiko sistemoje, turi didžiulį pranašumą. Įvykių kaupimas gali vesti prie didelio našumo domeno modelių, tai yra palaikyti ypač didelį skaičių operacijų per sekundę. Pavyzdžiui, įrašymas į vieną duomenų saugyklos lentelę yra ypač greitas. Negana to, tai leidžia CQRS užklausų modelį išplėsti horizontaliai, nes duomenų šaltinio atnaujinimai įvykdomi fone, kai įvykių saugykla yra atnaujinama naujais įvykiais.

¹⁰http://www.orm.net/

¹¹http://martinfowler.com/bliki/CQRS.html

2.2.4. Įvykių srautas

Pagal [Bas], įvykių srautas yra eilė pagal laiką surikiuotų įvykių, pavyzdžiui akcijų rinkos srautas.

Įvykių srautas kaip duomenų srauto tipas formaliai atrodo kaip pora (s, t), kur s yra seka surikiuotų sąrašo įvykių, o t yra seka laiko intervalų ir kiekvienas intervalas yra netuščias.

Tokio duomenų srauto pavyzdžiai gali būti:

- akcijų kursas,
- paspaudimų srautas,
- tinklo srautas,
- GPS¹² duomenys.

Įvykių srauto apdorojimas pagal atsitikimo laiką turi privalumų:

- įvykių apdorojimo algoritmai naudoja mažai atminties, nes jiems nereikia prisiminti daug įvykių;
- algoritmai gali būti labai greiti;
- gavus įvykį, skaičiavimai atliekami iškart, todėl galima perduoti rezultatą kitam skaičiavimui ir pamiršti įvykį.

Įvykių srauto apdorojimas labiau akcentuoja didelio našumo duomenų gavimą ir matematinių algoritmų pritaikymą įvykių duomenims. Taip pat įvykių srautai iprastai pritaikomi konkrečiai sistemai ar organizacijai.

2.2.5. Įvykių kaupimas funkciniame programavime

Vaughn Vernon pateikia keletą pastebėjimų apie įvykių kaupimą funkciniame programavime, kurie gali būti naudingi atliekant projektinius sprendimus bei eksperimentinį tyrimą:

¹²http://en.wikipedia.org/wiki/Global Positioning System

Agregatas projektuojamas kaip nekintantis būsenos įrašas kartu su funkcijomis, kurios keičia būseną. Šios funkcijos paprasčiausiai priima būsenos įrašą ir įvykių argumentus ir gražina naują būsenos įrašą kaip rezultatą. Tokios funkcijos atrodo:

Func<Busena, Ivykis, Busena>

- Dabartinė agregato būsena gali būti apibrėžta kaip suskleidimas į kairę visų praeities įvykių, kurie yra perduodami būseną keičiančiai funkcijai.
- Agregato metodai gali būti išreikšti kaip kolekcija funkcijų be būsenos.
- Įvykių saugykla gali būti suvokiama bei naudojama kaip funkcinė duomenų bazė, nes ji perduoda argumentus funkcijoms, kurios keičia agregato būseną.
 Momentinės kopijos įvykių saugykloje primena įsiminimą atmintyje¹³ funkciniame programavime.

2.3. Monados

2.3.1. Terminas

Funkciniame programavime monada yra struktūra, kuri atspindi skaičiavimus, apibrėžtus kaip seka žingsnių. Tipas kartu su monados struktūra apibrėžia ką reiškia vykdyti operacijas viena po kitos bei naudoti to pačio tipo įdėtines funkcijas. Tai leidžia kurti komandų grandines, kurios pažingsniui apdoroja informaciją, kur kiekvienas veiksmas yra dekoruojamas naujomis apdorojimo taisyklėmis, kurias apibrėžia monada [?].

taisykles?

2.4. Išvados

Literatūros analizės metu remiantis kitų autorių patirtimi:

- išnagrinėtas funkcinis-reaktyvus programavimas,
- išnagrinėtas įvykių kaupimo principas,

¹³http://en.wikipedia.org/wiki/Memoization

- išnagrinėti įvykių srautai,
- susipažinta su įvykių kaupimu funkciniame programavime,
- susipažinta su monadomis,
- įvaldyta sąvokų sistema, susijusi su nagrinėjama tematika.

3. Sąvokų apibrėžimai

- Agregatas (angl. aggregate) DDD modelis, rinkinys domeno objektų, kurie gali būti laikomi kaip visuma.
- Derinimas (angl. debugging) riktų bei klaidų paieška programinėje įrangoje bei jų taisymas.
- Esybė (angl. entity) kažkas, kas egzistuoja pats savaime, faktiškai arba hipotetiškai.
- Horizontalus išplečiamumas (angl. horizontal scaling) galimybė sujungti daugybę techninės ar programinės įrangos esybių taip, jog jos dirbtų kaip visuma. Pavyzdžiui, galima pridėti keletą serverių pasinaudojant grupavimu arba apkrovos paskirstymu taip pagerinant sistemos našumą bei prieinamumą.
- Metaduomenys (angl. metadata) duomenys apie kitus duomenis.

4. Santrumpos ir paaiškinimai

- CQRS (angl. Command Query Responsibility Segregation) komandų-užklausų atsakomybių atskyrimas.
- DDD (angl. Domain-Driven Design) būdas kurti programinę įrangą, skirtą spręsti sudėtingus uždavinius, bei apjungti realizaciją kartu su augančiu domeno modeliu.
- NoSQL duomenų bazė, skirta architektūriniams modeliams, kuriems nereikia palaikyti stiprios darnos principo, kuris naudojamas reliacinėse duomenų bazėse. Tai įgalina horizontalų išplečiamumą bei aukštesnį prieinamumą.
- ORM (angl. Object-Relational Mapping) programavimo technika duomenų konvertavimui tarp nesuderinamų sistemų tipų, naudojama objektinio programavimo kalbose. Pavyzdys: JAVA programavimo kalba bei Oracle duomenų bazė.

Literatūros sąrašas

- [Ams] Edward Amsden. A survey of functional reactive programming. Rochester Institute of Technology.
- [Bas] Tim Bass. Mythbusters: Event Stream Processing Versus Complex Event Processing. page 1. Invited talk.
- [Cza12] Evan Czaplicki. Elm: Concurrent frp for functional guis. Master's thesis, 30 March 2012.
- [EH97] Conal Elliott and Paul Hudak. Functional reactive animation. In *International Conference on Functional Programming*, 1997.
- [KW92] David J. King and Philip Wadler. Combining monads. In John Launchbury and Patrick M. Sansom, editors, Functional Programming, Workshops in Computing, pages 134–143. Springer, 1992.
- [Ver13a] V. Vernon. *Implementing Domain-Driven Design*. Pearson Education, 2013.
- [Ver13b] Vaughn Vernon. Implementing Domain-Driven Design. Addison-Wesley Professional, February 16, 2013.