

Strona projektu SaMURaI



Metryka projektu

- **Kierownik:** Krzysztof Kaczor
- **Tytuł:**
 - PL: *Metoda ujednoliconej wymiany reguł na poziomie semantycznym w systemach z bazą wiedzy.*
 - EN: *Semantic Method for Unified Rules Interoperability in Knowledge-Based Systems.*
- **Czas realizacji:** 2012-09-05 - 2015-01-04

Finansowanie

Projekt jest finansowany z funduszy Narodowego Centrum Nauki w ramach 3. konkursu PRELUDIM nr UMO-2011/03/N/ST6/00886.

Cel projektu

Głównym celem badań jest opracowanie nowej metody wymiany wiedzy regułowej. Obecnie istniejące narzędzia do modelowania systemów regułowych (takie jak CLIPS, Jess, Drools, OpenRules) dostarczają różnych języków przy pomocy których reguły mogą zostać zdefiniowane. Stanowią one jednak głównie specjalizowane języki programowania. Faktem jest, że posiadają one dobrze określoną składnię, jednak semantyka zakodowanej wiedzy rzadko kiedy posiada interpretację logiczną. Proponowane badania mają na celu opracowanie formalizmu dla wspomnianych języków regułowych, oraz dostarczenie formalnej metody wymiany wiedzy pomiędzy tymi językami.

Rezultatem projektu będzie formalna metoda pozwalająca na semantycznie spójną wymianę wiedzy regułowej. Zaletami proponowanej metody będą:

1. Jednolita logiczna interpretacja wiedzy zakodowanej przy pomocy języków regułowych (CLIPS, Jess, Drools, OpenRules). Może ona zostać osiągnięta poprzez ujednoliconą formalizację semantyki reguł. Formalizm ten może być wykorzystany jako podstawa do dalszych badań mających na celu stworzenie formalnej metody porównywania języków.
2. Semantycznie spójna metoda wymiany wiedzy pomiędzy językami regułowymi. Dzięki opisowi bazującemu na logice, możliwa będzie jednoznaczna translacja wiedzy z jednego formatu na inny przy jednoczesnym zachowaniu tej samej interpretacji logicznej. Opis ten pozwoli na formalne udowodnienie poprawności translacji.
3. Zwiększone możliwości utrzymania baz wiedzy. Dzięki opracowanej metodzie, reguły będą mogły być rozlokowane w wielu bazach wiedzy a następnie importowane i wykorzystywane w razie potrzeby. Pozwoli to na użycie większej liczby narzędzi do projektowania, analizowania i weryfikowania reguł.

Osiągnięcia

Poniższa lista przedstawia najważniejsze osiągnięcia wypracowane w ramach Projektu. Wszystkie z nich stanowią oryginalny wkład w dziedzinę i rozwiązanie postawionego we wniosku problemu semantycznej wymiany wiedzy regułowej:

1. Analiza i identyfikacja elementów języków regułowych decydujących o ich sile wyrazu – przeprowadzona analiza była koniecznym etapem prac gdyż na jej podstawie można było oszacować semantyczną złożoność sformalizowanego modelu reprezentacji reguł. W tym celu przeanalizowano cztery języki regułowe: CLIPS, Jess, Drools oraz OpenRules oraz możliwości logiki ALSV(FD) (Attributive Logic with Set Values over Finite Domain) pod kątem opisanie semantyki tych języków. Prace te zostały opisane w artykule Enabling Collaborative Modeling of Rule Bases by Semantic Interoperability Methods złożonego do czasopisma Future Generation Computer Systems posiadającego ImpactFactor 2.639.
2. Rozszerzenie logiki atrybutowej – na potrzeby prac w Projekcie do logiki atrybutowej ALSV(FD) wprowadzono istotne rozszerzenia. Najważniejsze z nich polegają na dodaniu do logiki systemu typów wspierającego dziedziczenie oraz polimorficzne tworzenie instancji reguł.
3. Wielopoziomowe podejście do wymiany wiedzy regułowej – nowym i bardzo istotnym elementem zaproponowanej metody jest próba uchwycenia semantyki bazy wiedzy jako całości a nie tylko jako semantyki poszczególnych jej elementów. Dlatego, semantyka wiedzy została opisana na dwóch poziomach: na poziomie poszczególnych elementów oraz na poziomie powiązań pomiędzy nimi (struktury bazy wiedzy). Koncepcję tego podejścia opisano w artykule Semantically-Driven Rule Interoperability – Concept Proposal.
4. Sformalizowany model reprezentacji wiedzy regułowej – na podstawie przeprowadzonej analizy języków regułowych opracowano sformalizowany model reprezentacji wiedzy regułowej. Model ten precyzyjnie określa semantykę poszczególnych elementów danego języka jak również wprowadza semantykę operacyjną definiującą znaczenie struktury bazy wiedzy wyrażonej w tym języku.
5. Sformalizowana metoda wymiany wiedzy – opracowany model stanowił kluczowy element do zdefiniowania metody wymiany wiedzy. Dzięki precyzyjnie określonej interpretacji bazy wiedzy, opracowana metoda dostarcza definicji wymiany semantyki na wszystkich rozważanych poziomach. W artykule Encapsulation-Driven Approach to Interchange of Knowledge Base Structure opisano koncepcje wymiany wiedzy na poziomie struktury bazy wiedzy.
6. Praktyczna ewaluacja metody – w celu ewaluacji metody zamodelowano kilka nietrywialnych przykładów systemów decyzyjnych we wszystkich rozważanych językach oraz dokonano ich translacji zgodnie z opracowaną metodą za pomocą narzędzia SUBITO. Przeprowadzona ewaluacja potwierdza poprawne działanie metody. Rezultaty ewaluacji praktycznej jak również informacje o implementacji narzędzia zostały zawarte w artykule Practical Approach to Interoperability in Production Rule Bases with SUBITO, który zostanie opublikowany w renomowanej serii Lecture Notes in Artificial Intelligence wydawanej przez wydawnictwo Springer.

Oprócz powyższych osiągnięć, prace zrealizowane dzięki Projektowi zaowocowały złożeniem i obroną **z wyróżnieniem** pracy doktorskiej Kierownika Projektu Krzysztofa Kaczora pt. *Knowledge Formalization Methods for Semantic Interoperability in Rule Bases* kierowanej i promowanej przez opiekuna Naukowego Projektu dr. hab. inż. Grzegorza J. Nalepę. W rozprawie można znaleźć wszystkie szczegóły na temat powyższych osiągnięć wraz opisem zidentyfikowanych ograniczeń zaproponowanego podejścia i planowanymi tematami przyszłych prac.

Uzyskane wyniki

Ostatecznym wynikiem prac wykonanych w Projekcie jest sformalizowana metoda pozwalająca na semantycznie spójną wymianę wiedzy regułowej pomiędzy reprezentacjami wspieranymi przez narzędzia CLIPS, Jess, Drools czy też OpenRules. Opracowana metoda umożliwia efektywną wymianę wiedzy regułowej poprzez translację która działa zarówno na poziomie syntaktycznym jak i semantycznym. Oznacza to, że podczas tłumaczenia wykonywana jest translacja nie tylko składni języka regułowego ale brane są także pod uwagę kwestie związane z semantyką zgromadzonej wiedzy. O ile zadanie tłumaczenia składni jest relatywnie łatwe do wykonania o tyle tłumaczenie semantyki nie jest problemem trywialnym. W celu lepszego usystematyzowania kwestii związanych z tym problemem, w Projekcie wprowadzono rozróżnienie semantyki bazy wiedzy na dwa odrębne poziomy abstrakcji. Pierwszy poziom jest związany z interpretacją reguł oraz ich poszczególnych elementów składowych i stanowi semantykę statyczną bazy wiedzy. Natomiast drugi poziom jest związany z interpretacją struktury bazy wiedzy i określa semantykę operacyjną (dynamiczną). Uwzględnienie wszystkich trzech poziomów abstrakcji (składni, semantyki statycznej i semantyki operacyjnej) bazy wiedzy pozwala na efektywne tłumaczenie wiedzy przez które, w Projekcie, rozumie się takie tłumaczenie w którym zarówno oryginalna jak i przetłumaczona baza wiedzy pozwalają na generowanie analogicznych wniosków dla analogicznych stanów początkowych. W celu wykazania poprawności proponowanego działania metody, w końcowej fazie Projektu przeprowadzona została teoretyczna i praktyczna ewaluacja zdefiniowanej metody.

Ewaluacja teoretyczna polegała na porównaniu semantyki operacyjnej reguł znajdujących się w bazach wiedzy, które zostały otrzymane jako rezultat translacji wykonanej z wykorzystaniem opracowanej metody. Analiza ta miała na celu wykazanie, że zaproponowana translacja struktury bazy wiedzy zachowuje jej semantykę i prowadzi do tego samego stanu końcowego systemu co struktura w bazie oryginalnej. W celu przeprowadzenia tej analizy wprowadzono odpowiednią notację przy pomocy której opisano semantykę operacyjną wszystkich zidentyfikowanych elementów determinujących strukturę bazy wiedzy. Przeprowadzone na jej podstawie rozważania pokazują, że zaproponowana translacja zachowuje semantykę, a co za tym idzie, stan końcowy osiągany przed i po translacji jest taki sam.

Druga część ewaluacji miała wymiar praktyczny. W ramach tej ewaluacji wykonano tłumaczenie bazy wiedzy przykładowych systemów regułowych, wyrażonych za pomocą sformalizowanego modelu, do wybranych języków i porównano rezultaty otrzymane po ich uruchomieniu. Poniżej przedstawiono przykładowe rezultaty dla modelu PLI – jednego z rozważanych systemów decyzyjnych, którego model wymaga zdefiniowania nietrywialnej struktury bazy wiedzy. System ten definiuje sposób obliczania wartości ubezpieczenia samochodowego OC w zależności od takich czynników jak: okres bezwypadkowej jazdy, kategoria kierowcy, pojemność samochodu, liczba poduszek powietrznych zamontowanych w samochodzie, itp. Jako rezultat uruchomienia wnioskowania w bazach wiedzy, otrzymanych na drodze translacji wykonanej zgodnie z opracowaną metodą, w każdym przypadku otrzymano taką samą wartość ubezpieczenia OC. Poniższe przykłady pokazują rezultaty produkowane w trakcie wnioskowania w bazach wiedzy zapisanych językach CLIPS oraz Drools:

- Rezultat uruchomienia wnioskowania w bazie wiedzy zamodelowanej języku CLIPS:

```
-----  
SUBIT0: Execution of ploc-clips.clp using CLIPS...  
-----
```

```
1 The base charge is set to 753  
2 Driver class has been increased from 1 to 2
```

```

3 90% of base charge because of the driver class (2)
4 +30% because of the small driver experience (2) that is less than 3
years
5 -10% because of the single payment
6 -10% because driver continues the previous agreement
7 +15% because of the car age (12) that is older than 10 years
8 Modifying insurance value 753 by 15% of basecharge 753
9 Modifying insurance value 865.95 by -10% of basecharge 753
10 Modifying insurance value 790.65 by 30% of basecharge 753
11 Modifying insurance value 1016.55 by -10% of basecharge 753
12 Modifying insurance value 941.25 by -10% of basecharge 753
13 The final amount to pay is equal to 865.95
real 0m0.026s
user 0m0.024s
sys 0m0.000s

```

```

-----
SUBIT0: Execution of ploc-clips.clp using CLIPS...done
-----

```

- Rezultat uruchomienia wnioskowania w bazie wiedzy zamodelowanej języku Drools:

```

-----
SUBIT0: Execution of ploc-drools.drl using Drools...
-----

```

```

1 Driver class has been increased from 1 to 2
2 The base charge is set to 753
3 -10% because of the single payment
4 -10% because driver continues the previous agreement
5 +15% because of the car age (12) that is older than 10 years
6 +30% because of the small driver experience (2) that is less than 3
years
7 90% of base charge because of the driver class (2)
8 Modifying insurance value 753.0 by -10.0% of basecharge 753.0
9 Modifying insurance value 677.7 by 30.0% of basecharge 753.0
10 Modifying insurance value 903.6 by 15.0% of basecharge 753.0
11 Modifying insurance value 1016.5500000000001 by -10.0% of basecharge
753.0
12 Modifying insurance value 941.2500000000001 by -10.0% of basecharge
753.0
13 The final amount to pay is equal to 865.9500000000002
real 0m2.792s
user 0m4.596s
sys 0m0.160s

```

```

-----
SUBIT0: Execution of ploc-drools.drl using Drools...done
-----

```

W przypadku tych dwóch modeli mamy doczynienia z inną sekwencją uruchamianych reguł powstałą w skutek innej kolejności dopasowywania faktów do reguł i tworzenia ich aktywacji (instancji) przez silnik wnioskujący. Z tego powodu można zaobserwować różną kolejność komunikatów pojawiających się w liniach 3 - 7. Niemniej jednak w tym przypadku zmieniona kolejność nie jest wynikiem błędu

translacji i jest dozwolona ponieważ nie występują zależności pomiędzy tymi regułami. Dlatego też nie wpływa ona na finalną wartość ceny ubezpieczenia.

Dodatkowo w ramach praktycznej ewaluacji, przeprowadzono także translację oryginalnej bazy wiedzy bez uwzględniania jej struktury, czyli w taki sposób jak mogłaby zostać dokonana przez aktualnie istniejące metody takie jak KIF, RIF, RuleML czy też R2ML. Wnioskowanie uruchomione w otrzymanej w ten sposób bazie produkuje konkluzje niezgodne z konkluzjami otrzymanymi w oryginalnej bazie wiedzy co świadczy o tym, że struktura nie może być pomijana w trakcie translacji i wskazuje na słuszność założeń przyjętych w pracy. W poniższym przykładzie łatwo zauważyć, że atrybut class jest wielokrotnie zwiększany i finalnie otrzymuje wartość 9 (zamiast prawidłowej 2). Niepoprawna wartość tego atrybutu propaguje się przez cały proces wnioskowania i sprawia, że wyliczona wartość ubezpieczenia jest również niepoprawna (489.45 zamiast 865.95).

```
1 Driver class has been increased from 1 to 2
2 Driver class has been increased from 2 to 3
3 Driver class has been increased from 3 to 4
4 Driver class has been increased from 4 to 5
5 Driver class has been increased from 5 to 6
6 Driver class has been increased from 6 to 7
7 Driver class has been increased from 7 to 8
8 Driver class has been increased from 8 to 9
9 The base charge is set to 753
10 -10% because of the single payment
11 -10% because driver continues the previous agreement
12 +15% because of the car age (12) that is older than 10 years
13 +30% because of the small driver experience (2) that is less than 3 years
14 40% of base charge because of the driver class (9)
15 Modifying insurance value 753.0 by -60.0% of basecharge 753.0
16 Modifying insurance value 301.2 by 30.0% of basecharge 753.0
17 Modifying insurance value 527.1 by 15.0% of basecharge 753.0
18 Modifying insurance value 640.05000000000001 by -10.0% of basecharge 753.0
19 Modifying insurance value 564.75000000000001 by -10.0% of basecharge 753.0
20 The final amount to pay is equal to 489.45000000000001
```

Ewaluacja praktyczna została przeprowadzona przy pomocy narzędzia SUBITO (Samurai Batch Integration Tool), które powstało w celu praktycznego wsparcia zaproponowanej metody i jest jednym z rezultatów Projektu. Głównym zadaniem tego narzędzia jest wykonanie translacji bazy wiedzy wyrażonej przy pomocy modelu formalnego do rozważanych języków regułowych a następnie ich uruchomienie w środowisku dedykowanym dla danej reprezentacji. Narzędzie wykorzystuje XML-owy format MCS (Model Concrete Syntax) zawierający odpowiedniki wszystkich elementów modelu formalnego, którego składnia została zdefiniowana przy pomocy XSD (XML Schema Definition). Kompletną definicję tego języka można znaleźć poniżej.

SUBITO Framework

- [XML Schema języka opisu modelu formalnego](#)

Przykłady

1. [PLOC](#)
2. [Load Processor](#)
3. [UserV](#)

Publikacje

1. **Krzysztof KACZOR, Grzegorz J. NALEPA** - *A critical overview of methods for rules interoperability : position paper*
Automatyka = Automatics; ISSN 1429-3447. — Tytuł poprz.: Automatyka : półrocznik Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie. — 2013 vol. 17 no. 1, s. 53-58. — Bibliogr. s. 57-58. - [tekst](#)
2. **Krzysztof KACZOR, Grzegorz J. NALEPA** - *Encapsulation-driven approach to interchange of knowledge base structure*
Lecture Notes on Software Engineering ; ISSN 2301-3559. — 2016 vol. 4 no. 1, s. 66-72. — Bibliogr. s. 72, Abstr. <http://dx.doi.org/10.7763/LNSE.2016.V4.225>
3. **Krzysztof KACZOR, Grzegorz J. NALEPA** - *Semantically-driven rule interoperability - concept proposal*
Artificial Intelligence and Soft Computing : 12th International Conference, ICAISC 2013 : Zakopane, Poland, June 9-13, 2013 : proceedings, Pt. 2 / eds. Leszek Rutkowski [et al.]. — Berlin ; Heidelberg : Springer-Verlag, cop. 2013. — (Lecture Notes in Computer Science ; ISSN 0302-9743 ; 7895. Lecture Notes in Artificial Intelligence). — ISBN: 978-3-642-38609-1 ; e-ISBN: 978-3-642-38610-7. — S. 511-522. — Bibliogr. s. 521-522, Abstr.
4. **Krzysztof KACZOR, Grzegorz J. NALEPA** - *Semantically-driven rule interoperability - concept proposal*
Artificial Intelligence and Soft Computing : 12th International Conference, ICAISC 2013 : Zakopane, Poland, June 9-13, 2013 : proceedings, Pt. 2 / eds. Leszek Rutkowski [et al.]. — Berlin ; Heidelberg : Springer-Verlag, cop. 2013. — (Lecture Notes in Computer Science ; ISSN 0302-9743 ; 7895. Lecture Notes in Artificial Intelligence). — ISBN: 978-3-642-38609-1 ; e-ISBN: 978-3-642-38610-7. — S. 511-522. — Bibliogr. s. 521-522, Abstr.
5. **Krzysztof KACZOR** - *Practical Approach to Interoperability in Production Rule Bases with SUBITO*
Artificial Intelligence and Soft Computing : 14th International Conference, ICAISC 2015 : Zakopane, Poland, June 14-18, 2015 : proceedings/ eds. Leszek Rutkowski [et al.]. — Berlin ; Heidelberg : Springer-Verlag. Acceptor for publication.
6. **Krzysztof KACZOR, Grzegorz J. NALEPA** - *Enabling collaborative modeling of rule bases by semantic interoperability methods*
Future Generation Computer Systems; Elsevier. IF 2,639. Submitted.
7. **Krzysztof KACZOR, Grzegorz J. NALEPA** - *Formalization of production rule representation model for multilevel approach to rule interoperability*
Data & Knowledge Engineering; Elsevier. IF 1.115. Submitted.

From:

<https://geist.re/> - **GEIST Research Group**

Permanent link:

<https://geist.re/pub:projects:samurai:startpl>

Last update: **2016/01/05 22:29**

