Rukovanje pravilima u poslovnim veb aplikacijama

# Uvod

# Sistemi bazirani na pravilima i poslovne veb aplikacije

## O sistemima baziranim na pravilima i ekspertskim sistemima

U računarstvu, sistemi bazirani na pravilima služe za skladištenje i manipulaciju znanja, kako bi se interpretirale informacije na način koristan za problem koji aplikacija rešava. Često se koristi u aplikacijama koje se oslanjaju na računarsku inteligenciju, kao i u istraživanju računarske inteligencije. Termin *sistem baziran na znanju* se uglavnom koristi za sisteme koji uključuju pravila definisana od strane ljudi koji se smatraju ekspertima u domenu problema. Klasičan primer sistema baziranog na znanju je domenski specifičan *ekspertski sistem* koji koristi pravila da dođe do zaključaka ili izbora. Primera radi, ovakvi sistemi mogu da pomognu doktorima da dođu do tačne dijagnoze, na osnovu unesenih simptoma, ili da se odaberu taktički najoptimalniji potezi u nekoj igri, na primer šahu. Ovakvi sistemi se takođe mogu koristiti u leksičoj analizi, ili u procesiranju prirodnog jezika.

Pod ekspertskim sistemima se podrazumevaju sistemi bazirani na znanju, koji pri rezonovanju simuliraju sposobnost donošenja odluka karakterističnih za ljudskog eksperta. Oni su dizajnirani da rešavaju kompleksne probleme rezonovanjem kroz bazu znanja uz pomoć pravila, koja su uglavnom predstavljena kroz if-then strukturu. Ekspertski sistemi se sastoje iz dva podsistema: baze znanja i mehanizma za zaključivanje (rezonera). Baza znanja obuhvata pravila i činjenice (fakte), odnosno, adekvatno reprezentovano znanje relevantno za domen problema, a mehanizam za zaključivanje primenjuje pravila na postojeće činjenice, da bi došao do zaključaka, i iz njih izvukao nove činjenice.

Sistemi bazirani na pravilima se prvi put formalno pojavljuju 1965. godine na univerzitetu Stanford, gde su istraživači, predvođeni Edvardom Fajgenbaumom, koji se smatra ocem ekspertskih sistema, pokušali da stvore inteligentan sistem opšte svrhe, koji bi bio sposaban da rešava probleme svakog tipa. Očekivanja su bila visoka, ali je istraživanje završeno neuspehom, jer je ovaj pokušaj bio previše ambiciozan. Kasnije su shvatili da bi bilo bolje da se ograniče na jedan specifičan domen problema, pa su krajem 60ih godina pokušali da identifikuju domene u kojima je ekspertsko znanje visoko cenjeno i kompleksno, kao što je npr dijagnostika bolesti (Mycin, 1972), identifikacija nepoznatih organskih molekula (Dendral, 1969), ili geološka analiza tla i minerala (Prospector, 1979). Ovo dovodi do prvih velikih uspeha računarske inteligencije.

 Prvi sistem baziran na pravilima koji je imao kapacitet da bude veliki komercijalni uspeh je bio SID (Synthesis of Integral Design), razvijen 1982. godine. Pisan u LISP-u, SID je generisao 93% logičkih kola za superračunar VAX-9000. Ulazni parametri su bila pravila koja su ručno pisana od strane eksperata, a SID je proširivao pravila i generisao rutine za logičku sintezu, koje su bile višestruko veće od samih izvornih pravila. Kombinacija generisanih pravila je rezultovala dizajnom koji je nadmašio sposobnosti samih eksperata, i u većini slučajeva bolje funkcionisala od pravila pisanih ručno. Nakon završetka VAX-9000 superračunara, projekat SID je završen i nije više nikad korišćen.

*Edvard Fajgenbaum*

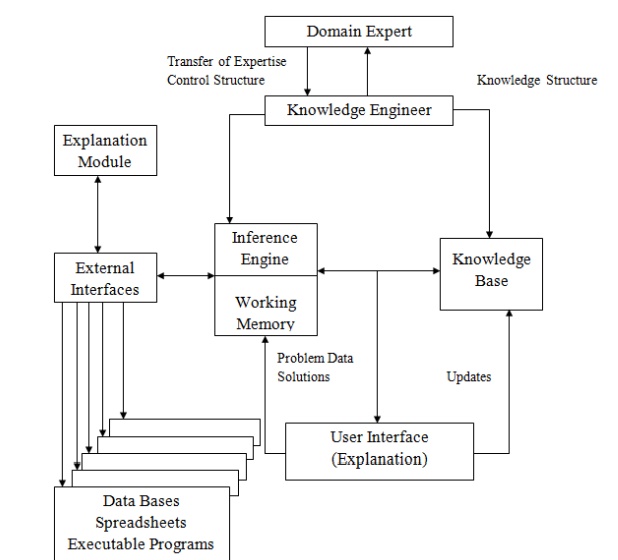
Moderni sistemi bazirani na pravilima i dalje primenjuju ideje ustanovljene od strane pionira, ali su često integrisani unutar velikih, kompleksnih sistemima. Ograničenja na koja su rani istraživači nailazili su ih naterala da potraže nove metode pristupa problemu, da bi se što vernije replicirao ljudski proces donošenja odluka. Često su ti pristupi bazirani na novijim metodama u računarskoj inteligenciji, kao što je *mašinsko učenje* i *data mining.* Moderni sistemi efikasnije pripajaju novo znanje i samim tim se efikasnije ažuriraju. Mogu da izvode generalizacije iz postojeće baze znanje i da obrađuju velike količine kompleksnih podataka – iz čega nastaje *big data* koncept. Ponekad se ovakvi sistemi nazivaju “inteligentni sistemi”.

Važno je napomenuti da sistemi bazirani na pravilima NISU veštačka inteligencija. Bitna karakteristika veštačke inteligencije jeste da je ona u stanju da sama odluči koju akciju da preduzme; može da uči i da se adaptira. Sistemi bazirani na pravilima samo izvršavaju pravila tačno onako kako su definisana od strane ljudi. Dakle, sistem ne radi za sebe i ne pravi “inteligentne“ odluke, niti se sam ažurira i ne uči iz svojih grešaka.

## Struktura sistema baziranih na pravilima

Tipičan sistem baziran na pravilima sadrži 4 komponente:

* Listu pravila, ili bazu pravila, koja predstavlja specifičnu vrstu baze znanja
  + Pravila su obično u formatu IF-THEN: postoje uslov (tzv. leva strana pravila) i akcija (tzv. desna strana pravila). Ako je uslov zadovoljen, izvrši akciju
  + Cilj je da se automatizuju procesi, uz razbijanje tih istih procesa na korake
* Modul za zaključivanje (rezoner), koji vrši zaključivanje na osnovu ulaznih podataka i baze pravila. Zaključivanje se vrši u 3 koraka:
  + Podudaranje – provera da li su uslovi pravila zadovoljeni, nakon čega se kreira uređena lista elemenata radne memorije koji zadovoljavaju levu stranu pravila
  + Razrešenje konflikta, odnosno određivanje redosleda u kom će se desne strane pravila koja zadovoljavaju uslove izvršiti
  + Akcija, odnosno izvršenje desnih strana pravila u redosledu određenom tokom faze razrešenja konflikta
* Modul za obrazlaganje, koji objašnjava rezonovanje sistema krajnjem korisniku. Kredibilitet sistema se uspostavlja onog momenta kada može da objasni *kako i zašto* je došao do nekog zaključka
  + Kako – da bi obrazložio kako je došao do zaključka, modul za obrazlaganje prati lanac pravila koja su izvršena tokom interakcije sa korisnikom, i deskriptivno objašnjava kako je dedukovao određenu činjenicu i zašto jeste ili nije koristio određena pravila.
  + Zašto – modul za obrazlaganje mora da objasni zašto je određena informacija potrebna da bi se kompletirao korak u rezonovanju
* Privremenu radnu memoriju, koja sadrži činjenice (fakte)
* Korisnički interfejs ili neku drugu vrstu povezivanja sa spoljnim svetom, koja služi za slanje i primanje ulaznih/izlaznih signala i komunikaciju sa korisnikom



*Struktura sistema baziranog na pravilima*

## Inženjerstvo znanja

Bitan proces u razvoju sistema baziranih na pravilima je inženjerstvo znanja. Ključni ljudi koji su uključeni u razvijanje sistema su inženjer znanja, ekspert za domen i krajnji korisnik. Nakon što inženjer znanja prikupi dovoljno generalnih informacija o domenu problema, on treba da se konsultuje sa domenskim ekspertom kako bi se rešili postojeći problemi oko dizajna, i nakon toga sistem je spreman za razvijanje – bira se način reprezentacije znanja, dizajn korisničkog interfejsa, itd. Po završetku dizajna, inženjer znanja pravi prototip sistema, koji treba da bude u stanju da reši probleme u malom poddomenu problema. Nakon što je prototip završen, inženjer znanja i domenski ekspert testiraju i dorađuju znanje davajući sistemu sve više problema iz domena koje treba da reši, i usput koriguju nedostatke.

## Strategije zaključivanja

Modul za zaključivanje može da vrši zaključivanje putem 2 strategije: ulančavanje unapred i ulančavanje unazad.

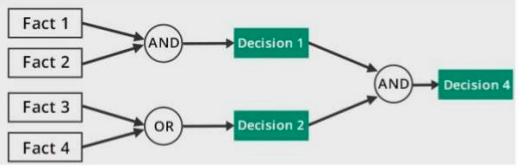
Kod ulančavanja unapred, počinje se od poznatih činjenica, i pomera se unapred primenom pravila na te činjenice, kako bismo došli do novih činjenica, i tako sve dok ne stignemo do cilja (zaključka). Zbog ovoga se ova strategija naziva i *zaključivanje vođeno podacima (data-driven),* ili *pristup* *odozdo na gore.* Izvršavanjem desnih strana pravila za one fakte čije su leve strane zadovoljene, dolazimo do zaključaka koje dalje možemo koristiti da bismo došli do novih zaključaka. Prolazi se kroz sva moguća pravila i koristi se strategija pretrage poznata kao *prvi u širinu*. Pretraga prvi u širinu – pretraga počinje od korenskog elementa, i pretražuje sve susedne elemente na grafu, odnosno elemente koji su na istom nivou dubine, pre nego što krene na sledeći nivo.Koristi se za zadatke kao što su planiranje, dijagnostika, klasifikacija, monitoring dizajn procesa...

Kod ulančavanja unazad, počinje se od cilja (zaključka) i ide se u nazad, odnosno izvršavaju se pravila kako bi se dobile činjenice koje podržavaju zaključak. Zbog ovoga se ova strategija naziva i *zaključivanje vođeno ciljem (goal-driven),* ili *pristup* *odozgo na dole.* Iz cilja se izvlače podciljevi dok se ne pronađu činjenice koje podržavaju sve podciljeve, a samim tim i cilj. Ne prolazi se kroz sva pravila, nego samo kroz ona neophodna da se dobiju zadovoljavajuće činjenice i koristi se strategija pretrage poznata kao *prvi u dubinu*. Pretraga prvi u dubinu – pretraga počinje od korenskog elementa, pomera se dalje za jedan nivo i ide što dublje može pre nego što se vrati za jedan nivo unazad i pretraži susedni čvor. Koristi se za zadatke kao što su dijagnostika, preskripcija, debagovanje aplikacija...





*Pretraga prvi u širinu Pretraga prvi u dubinu*



*Ulančavanje unapred*



*Ulančavanje unazad*

## Programski jezici bazirani na pravilima

* AWK – za procesuiranje teksta, ekstraksciju podataka i alat za izveštavanje
* CLIPS – najrasprostranjeniji alat za ekspertske sisteme, objektno orijentisan
* Constraint Handling Rules – deklarativni jezik namenjen rešavanju kombinatornih problema, takođe ima primenu u gramatičkoj indukciji, abduktivnom rezonovanju, multiagentskim sistemima, procesuiranju prirodnog jezika, prostorno-vremenskom rezonovanju, itd.
* Drools
* GOAL agent programming language – koristi se za programiranje kognitivnih agenata
* Jess – deklarativni jezik namenjen za Java platformu, nadskup CLIPS-a, ima primenu u automatizaciji ekspertskih sistema
* OPS5 – prvi jezik korišćen u uspešnom ekspertskom sistemu, za konfigurisanje VAX računara
* Prolog – logički programski jezik, sa korenima u logici prvog reda, danas jedan od najpopularnijih jezika ovog tipa. Ima široku primenu u automatizaciji planiranja, ekspertskim sistemima, dokazivanju teorema, procesuiranju prirodnog jezika, itd.
* ToonTalk – programski sistem namenjen za decu, prezentovan u formi animiranih likova
* Mathematica – široko rasprostranjen moderan računarski sistem koji podržava neuronske mreže, mašinsko učenje, procesuiranje slika, geometriju, vizualizacije, itd.
* Wolfram – jezik koji podržava velik broj paradigmi, sa akcentom na funkcionalnom programiranju, simboličkom računanju i programiranje bazirano na pravilima

## Prednosti i mane sistema baziranih na pravilima

Neke od glavnih prednosti sistema baziranih na pravilima su sledeće:

* Pravila se izražavaju deklarativno – omogućuju nam da kažemo šta želimo da se uradi, a ne kako da se uradi. Dakle, pravila su mnogo čitljivija od koda i često se rešenja težih problema mogu lakše predstaviti pravilima
* Razdvajanje poslovne logike od ostatka koda – poslovna logika je često podložna promenama, i ako je izražena u pravilima, lakša je za menjanje i održavanje. Na ovaj način, sva poslovna logika se nalazi na jednom mestu
* Brzina i skalabilnost – sistemi bazirani na pravilima često su praćeni algoritmima koji efikasno uparuju podatke i pravila (npr Rete algoritam, Leaps algoritam, itd.)
* Centralizacija znanja – koristići pravila kreiramo repozitorijum znanja koji nam služi kao izvor istinitosti za poslovnu logiku – idealno su pravila toliko čitljiva da mogu da služe i kao dokumentacija
* Sposobnost objašnjavanja – pravila nas obaveštavaju o svim odlukama, kao i zašto su one donete
* Razumljivost – sintaksa pravila je bliža prirodnom jeziku, i samim tim je razumljivija ljudima koji nemaju tehničke sposobnosti programera
* Mogućnost da se menjaju pravila bez ponovnog startovanja aplikacije

Nedostaci sistema baziranih na pravilima:

* Alati koji barataju pravilima su često resursno zahtevni
* Narušavanje objektno orijentisanog koncepta na kom počivaju mnoge aplikacije, logika se prenosi u pravila umesto da ostaje unutar objekata
* Sa programerske strane, može biti teško testirati i refaktorisati deklarativno definisana pravila i integrisati ih sa kodom koji je pisan na drugačiji način
* Sa povećanjem broja pravila, postaje sve teže održavati ih – mora se voditi računa o tome koja pravila su međusobno zavisna i koje objekte menjaju, o prioritetu izvršavanja pravila, itd
* “Eksplozija pravila“ – u pokušaju da se reši problem, često dolazi do prevelikog broja pravila koja se preklapaju i koja mogu biti kontradiktorna; sistem postaje kompleksan i nerazumljiv
* Sintaksa pravila se mora naučiti i to zahteva vreme – što programeri često nisu voljni da rade

Kada koristiti sisteme bazirane na pravilima?

* Kada ne postoji zadovoljavajuće tradicionalno programersko rešenje problema
* Kada je poslovna logika podložna čestim promenama
* Kada domenski eksperti nemaju adekvatno programersko znanje – lakše ih je podučiti sintaksi pravila nego sintaksi programskog jezika

Kada ne koristiti sisteme bazirane na pravilima?

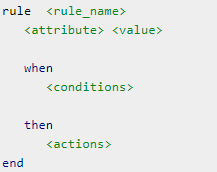
* Ukoliko je mali projekat, sa manje od 20 pravila, korišćenje pravila je nepotrebno
* Ukoliko je poslovna logika dobro definisana i statična, nepodložna promenama
* Ukoliko su pravila previše prosta, i uključuju rad nad malim brojem objekata
* Ukoliko je primarni fokus na performansu, u tom slučaju je optimalnije ostaviti logiku u samom kodu i ne koristiti pravila
* Ukoliko problem nije prikladan za pisanje u pravilima, ili je teško da se izrazi pravilima

## Programski jezik Drools

Drools je integraciona platforma za poslovnu logiku pisana u Javi, podržana od strane Jboss-a i Red Hat-a koja implementira Rete algoritam, o kom će više reči biti kasnije. Predstavlja kolekciju alata koji nam omogućavaju da rezonujemo nad logikom i podacima u poslovnim procesima. Drools možemo podeliti u 2 dela:

* autoring, odnosno proces pisanja pravila
* vreme izvršavanja (runtime), odnosno proces pravljenja objekata u radnoj memoriji i regulisanja aktivacije pravila

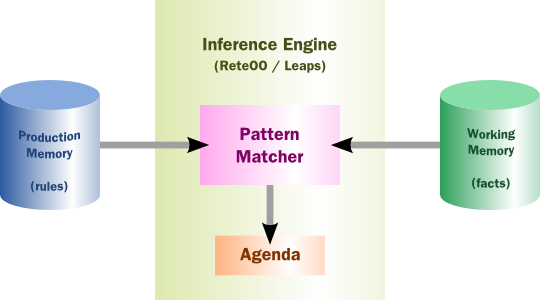
Kao i kod većine jezika baziranih na pravilima, pravila u Drools-u su u formatu if-then – ako je neki uslov zadovoljen, izvrši akciju.

Proces uparivanja postojećih činjenica sa pravilima se naziva *podudaranje šablona* (*pattern matching*), i njega izvršava modul za zaključivanje. Algoritam za sprovođenje ovog procesa se u slučaju Drools-a zove ReteOO algoritam – poboljšana i optimizovana verzija Rete algoritma za objektno orijentisane sisteme. Drools radi po principu hibridnog ulančavanja – može da radi i ulančavanje unapred i unazad. Drools se lako integriše sa postojećim razvojnim okruženjima kao što su Eclipse, IntelliJ i Visual Studio Code, koja nude podršku za debagovanje, validaciju i pisanje pravila.

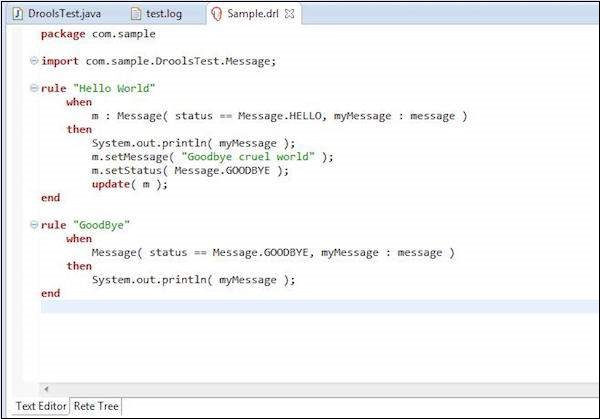
*Pseudokod jednog pravila*

Terminologija u Drools-u:

* Pravila – fajlovi sa ekstenzijom .drl u kojima navodimo uslove i akcije koje treba da se izvrše ako su uslovi zadovoljeni (when ‘a’ then ’b’)
* Činjenice – podaci nad kojima se pravila izvršavaju. U Javi, to su obično POJO objekti
* Sesija – glavna komponenta koja služi za okidanje pravila. Činjenice se ubacuju u sesiju i kada su uslovi ispunjeni, izvršavaju se desne strane pravila. Dva tipa sesije: stateless i stateful.
  + Stateless – stanje sesije između poziva se ne održava, za svaki poziv se pravi nova sesija
  + Stateful – pamti stanje između poziva
* Aktivacija – akcija, desna strana pravila
* Agenda – logički koncept, mesto na kom aktivacije čekaju da budu izvršene, odnosno mesto na kom se nalaze sve aktivacije čiji su uslovi zadovoljeni



*Struktura Drools-a*



*Primer jednog pravila u Drools-u*

## Rete algoritam

Ovaj algoritam je razvio američki naučnik Charles L. Forgy kasnih 70ih godina, ali je prošlo dosta vremena pre nego što je ovaj algoritam postao popularan. Pojavom tehnologije poslovnih pravila, algoritam dobija na popularnosti. Algoritam je prošao kroz dosta izmena od inicijalne verzije, do konačne verzije 1982. godine koju je Forgy predao kao doktorat. Naziv Rete potiče od latinske reči za mrežu – graf algoritma u velikoj meri zaista podseća na mrežu. Suština algoritma jeste da uklanja potrebu za suvišnim kalkulacijama – on pamti kako celokupna, tako i parcijalna zadovoljenja uslova.

# Arhitektura sistema

# Implementacija

# Zaključak

# Literatura

<https://en.wikipedia.org/wiki/Rule-based_system>

<https://en.wikipedia.org/wiki/Expert_system>

<https://www.thinkautomation.com/eli5/what-is-a-rule-based-system-what-is-it-not/>

<http://www.cs.stir.ac.uk/courses/ITNP60/lectures/2%20Decision%20Support/1%20-%20Rule%20Based%20Systems.pdf>

<https://www.brainkart.com/article/Rule-Based-Architecture-of-an-Expert-System_8931/>

<https://link.springer.com/article/10.1007/s41066-016-0021-6>

<https://www.javatpoint.com/forward-chaining-and-backward-chaining-in-ai>

<https://www.javatpoint.com/difference-between-backward-chaining-and-forward-chaining>

<https://en.wikipedia.org/wiki/Depth-first_search>

https://en.wikipedia.org/wiki/Breadth-first\_search

<https://en.wikipedia.org/wiki/List_of_programming_languages_by_type#Rule-based_languages>

<http://www.jbug.jp/trans/jboss-rules3.0.2/ja/html/ch01s02.html>

<https://stackoverflow.com/questions/250403/rules-engine-pros-and-cons>

<https://jessrules.com/guidelines.shtml>

<https://www.studytonight.com/drools/not-use-rule-engine>

<https://www.tutorialspoint.com/drools/drools_introduction.htm>

<https://www.tutorialspoint.com/drools/drools_frequently_used_terms.htm>

<https://www.sparklinglogic.com/rete-algorithm-demystified-part-1/>

<https://www.sparklinglogic.com/rete-algorithm-demystified-part-2/>