Rozdział 5

RBS(2U) – system regułowy z niepewnością drugiego rzędu

5.1 Wstęp

Każdy system ekspercki, także system regułowy z niepewnością, ma swojego odbiorcę. W nielicznych wypadkach (np. regulator bezpośredniego działania) jest nim maszyna, działająca na podstawie odpowiedzi generowanych przez ten system w trakcie prowadzenia wnioskowania. Odpowiedzi takie mają charakter bezwzględnych zaleceń dotyczących zmiany lub utrzymania na niezmienionym poziomie parametrów pracy maszyny. Sposób formułowania odpowiedzi, ściśle związany z reprezentacją wiedzy w systemie, powinien być wtedy dostosowany do potrzeb maszyny.

W większości wypadków, odbiorcą i użytkownikiem systemu eksperckiego jest człowiek. O jakości takiego systemu decydują czynniki mierzalne, o których jest mowa podrozdziałach 2.7 i 4.5, lecz także niemierzalne, które wynikają z poziomu zadowolenia jego użytkowników. Odpowiedzi systemu powinny być zredagowane na tyle prosto, by nie stwarzać trudności interpretacyjnych naiwnemu użytkownikowi, i – równocześnie – na tyle precyzyjnie, by zaspokoić wymagania wyrafinowanego użytkownika. W wypadku systemu regułowego z niepewnością, wyrafinowany użytkownik będzie oczekiwał w odpowiedzi nie tylko zbioru możliwych hipotez, lub też weryfikacji/falsyfikacji zadanej hipotezy, lecz także określenia poziomu wiarygodności, z jakim są one formułowane. To zaś oznacza, że do reprezentacji wiedzy należy koniecznie zastosować metodę z dwupunktową estymacją pewności. Z drugiej strony, pożądane jest by ta pewność była zadana w sposób granularny: oddzielnie dla konkluzji zawartej w hipotezie, i oddzielnie dla poziomu wiarygodności hipotezy jako całości. W takiej sytuacji naiwny użytkownik będzie mógł, na życzenie, zrezygnować z dostępu do informacji o poziomie wiarygodności.

Spośród metod reprezentacji niepewności omawianych w podrozdziale 3.7, żądane kryteria spełnia nie-aksjomatyczny system wnioskowania NARS według Wanga, z regułami w formacie (4.8). W regule Wanga $P \rightarrow C\ \langle p_1;\ p_2 \rangle$ dwupunktowa estymacja pewności jest zaimplementowana za pomocą parametrów: częstotliwości (p_1), określającego prawdopodobieństwo zajścia konkluzji C przy pewnym zajściu przesłanki P, oraz zaufania (p_2), określającego stopień wiarygodności parametru p_1. Oba parametry mają interpretację częstotliwościową i przyjmują wartości z przedziałów, odpowiednio $\langle 0; 1\rangle$ i $(0; 1)$. Sposób implementacji funkcji propagacji niepewności f_{2i} oraz f_{3i} w systemie NARS gwarantuje zróżnicowanie wpływu poszczególnych reguł na wynik końcowy wnioskowania. Jeśli reguła $P \rightarrow C\ \langle p_1;\ p_2\rangle$ wpływa na pewną konkluzję końcową $P_k \rightarrow C_k\ \langle p_{k1};\ p_{k2}\rangle$ procesu wnioskowania, to jej wpływ na częstotliwość p_{k1} konkluzji jest tym większy, im wyższe jest jej zaufanie p_2.

System NARS posiada jednak i poważne ograniczenia, które zmniejszają jego praktyczną użyteczność (patrz: podrozdz. 4.4.3). Najważniejsze z nich dotyczą budowy reguły oraz specyfiki działania silnika wnioskującego.

W regule Wanga zarówno przesłanka, jak i konkluzja mają postać formuł atomowych prostych (niezanegowanych). Taka reguła może wyrażać tylko relację dwuargumentową. Tymczasem, większość relacji pomiędzy obiektami, stanami i zjawiskami obserwowalnymi w świecie rzeczywistym ma charakter wieloargumentowy. Do ich wyrażania potrzebne są reguły z przesłankami złożonymi, sformułowanymi w postaci koniunkcji formuł atomowych. Takie reguły odkrywa się w procesie uczenia maszynowego, z danych zgromadzonych w repozytoriach o schemacie atrybutowym płaskim (relacyjne bazy danych) lub strukturalnym (obiektowe bazy danych). Użycie w nich parametru zaufania ($0 < p_2 < 1$) byłoby w pełni uzasadnione.

Relacje dwuargumentowe (i reguły z przesłanką atomową) dobrze nadają się do modelowania zależności semantycznych, na czele z klasycznymi relacjami hiperonimii i meronimii. Tych relacji nie odkrywa się jednak z danych, lecz zadaje aksjomatycznie, z maksymalną możliwą pewnością, np.

$$\text{Rodzaj\_studiow} = \text{uniw\_II\_st} \rightarrow \text{Tytuł\_zawodowy} = \text{mgr}\ \langle 0.98; 1\rangle).$$

W takiej sytuacji parametr zaufania ($p_2 = 1$) staje się nadmiarowy.

Podobnie jak w innych systemach regułowych, także i w systemie NARS wszystkie reguły w bazie wiedzy systemu muszą wyrażać zależności dodatnie monotoniczne. Spełnienie tego warunku jest niezbędne dla prawidłowego przebiegu wnioskowania i prawidłowej propagacji niepewności w systemie. Przy tym wymogu, ograniczenie w regułach Wanga postaci formuł atomowych do prostych (niezanegowanych) jeszcze bardziej zawęża zbiór zależności, które da się wyrazić przy ich użyciu. Przykładowo, nie da się skonstruować reguły Wanga:

$$\text{Typ\_uczelni} = \text{uniwersytet} \rightarrow \text{Liczba\_kierunkow} = (\leq 40)\ \langle 0.35; 1\rangle,$$

gdzie symboliczny zapis (≤ 40) oznacza liczbę całkowitą nie większą niż 40 – bo zależność pomiędzy przesłanką a konkluzją byłaby w niej zależnością ujemną monotoniczną. W wypadku dopuszczenia operatora negacji, można by natomiast zaproponować poprawną regułę:

$\text{Typ\_uczelni} = \text{uniwersytet} \rightarrow \text{Liczba\_kierunkow} = \neg(\leq 40)\ \langle 0.65; 1 \rangle$.

Silnik wnioskujący w systemie NARS działa przy założeniu zmiennej zawartości bazy wiedzy. Przyjmuje się, że w trakcie prowadzenia wnioskowania baza ta może być uzupełniana o dowolne reguły, których treść wynika z nowo udostępnionej ewidencji. W szczególności, możliwa jest sytuacja, w której do bazy wiedzy zawierającej pewną regułę $P \rightarrow C\ \langle p_{11}; p_{12} \rangle$ dodaje się inną regułę $P \rightarrow C\ \langle p_{21}; p_{22} \rangle$, z tymi samymi przesłanką i konkluzją. Takie działanie jest uzasadnione w wypadku, gdy system działa w warunkach dynamicznego otoczenia i w trybie rzeczywistym wyciąga wnioski na temat jego zmieniających się parametrów. W opisanej sytuacji należałoby jednak usunąć z bazy wiedzy regułę pierwotną $P \rightarrow C\ \langle p_{11}; p_{12} \rangle$. Następnie, w trosce o zachowanie wiarygodności systemu, w ślad za regułą $P \rightarrow C\ \langle p_{11}; p_{12} \rangle$ należałoby usunąć/zmodyfikować w bazie danych systemu wszystkie te fakty, na których postać wpłynęła ta reguła, następnie fakty zależne od tych poprzednich itd., aż do momentu osiągnięcia stanu stabilnego sytemu. Zważywszy nieustaloną (w systemie regułowym) kolejność uaktywniania reguł, realizacja tego procesu musiałaby się opierać na algorytmie heurystycznym, a jego złożoność miałaby charakter wykładniczy. Oczywistym jest fakt, że proces ten byłby niemożliwy do przeprowadzenia w reżimie czasu rzeczywistego.

Aby rozwiązać powyższy problem, Wang wzbogacił regułę i fakt o dwa dodatkowe parametry, zwane pilnością i trwałością, które determinują początkowy priorytet reguły/faktu i jej/jego trwałość. Oba parametry mają charakter względny i przyjmują wartości z przedziału $\langle 0; 1 \rangle$. Przy tym, o ile częstotliwość i zaufanie stanowią uniwersalne wskaźniki pozyskanej wiedzy, o tyle pilność i trwałość określa się na potrzeby konkretnego wnioskowania, bez związku z jakością tej wiedzy. Dzięki tym dodatkowym parametrom, podsystem utrzymywania wiarygodności staje się zbędny. Łatwo jednak zauważyć, że przy braku właściwości lokalnej zgodności, te dodatkowe parametry będą wpływać nie tylko na kolejność uaktywniania reguł, lecz także na wynik końcowy wnioskowania. Z tego powodu, przedstawione rozwiązanie – choć bardzo wydajne – wydaje się ryzykowne: efektywność działania takiego systemu zależy w znacznym stopniu od samego użytkownika.

Z powyższych rozważań wynika, że system NARS, podobnie jak większość systemów regułowych z niepewnością, jest raczej predestynowany do działania w warunkach statycznych. Do oceny statycznych właściwości obiektu/stanu/sytuacji wystarcza jednak baza wiedzy o stałej, niezmieniającej się w czasie zawartości. Zależnie od potrzeb, można ją aktualizować synchronicznie lub asynchronicznie w okresach bezczynności systemu. Przy takim użyciu systemu, subiektywne parametry pilności i trwałości reguł stałyby się zbędne, a jakość odpowiedzi systemu zależałaby wyłącznie od jakości jego bazy wiedzy i działania silnika wnioskującego.

Zgodnie ze wzorem 4.11, jakość bazy wiedzy systemu regułowego z niepewności zależy od liczby reguł w bazie wiedzy i od średniego poziomu wiarygodności tych reguł. Choć w obu wypadkach ta zależność ma charakter dodatni monotoniczny, odpowiednio logarytmiczny i liniowy, to jednak powiększanie bazy wiedzy nie może się odbywać w sposób bezwzględny. W szczególności, ze względu na wymóg niesprzeczności, baza wiedzy nie może zawierać dwóch różnych reguł z tą samą przesłanką i tą samą konkluzją, np. $P \rightarrow C \ \langle p_{11};\ p_{12}\rangle$ i $P \rightarrow C \ \langle p_{21};\ p_{22}\rangle$. Takie reguły należy scalić w jedną na etapie projektowania bazy. Jeśli połączenie zostanie wykonane przy użyciu reguły rewizji (3.9), to parametr zaufania w wypadkowej regule $P \rightarrow C \ \langle p_1;\ p_2\rangle$ będzie miał wartość większą od wartości parametrów zaufania w obu regułach cząstkowych $(p_2 > p_{12}, p_2 > p_{22})$.

Zapewnienie niesprzeczności bazy wiedzy systemu nie jest równoznaczne z nadaniem systemowi właściwości lokalnej zgodności. Po rezygnacji z posługiwania się subiektywnymi parametrami reguł – pilnością i trwałością, na jakość odpowiedzi systemu wpływałby nowy algorytm rozstrzygania konfliktów w agendzie. W celu zwiększenia stopnia poprawności odpowiedzi, warto oprzeć jego działanie na regułach cechujących się najwyższą jakością, wyrażaną w tym wypadku za pomocą parametru zaufania.

5.2 Reguła 2U - składnia i semantyka

W pracy [101] zaproponowano model systemu regułowego z niepewnością oparty na formacie reguły z niepewnością drugiego rzędu, zwanej w skrócie regułą 2U (ang. 2-uncertain rule). Stanowi on wynik poszukiwań takiej metody reprezentacji wiedzy niepewnej i takiego algorytmu wnioskowania przybliżonego, które w połączeniu zagwarantują:

– konstrukcję systemu cechującego się wysoką jakością,
– możliwość pozyskiwania wiedzy do bazy wiedzy tego systemu w sposób półautomatyczny, z danych zapisanych w popularnym formacie atrybutowym.

W powyższym modelu zaadaptowano niektóre z rozwiązań zastosowanych przez Wanga w jego systemie NARS. Wprowadzone modyfikacje miały na celu zniesienie pewnych ograniczeń tego systemu, zarówno tych dotyczących efektywności i wydajności działania systemu, jak i tych, związanych z obszarem jego zastosowań.

Niech $\mathcal{D}_u$ oznacza dziedzinę problemową, dla której należy zaprojektować system regułowy z niepewnością, $\mathcal{T}_{\mathcal{D}u}$ – wybraną konceptualizację tej dziedziny, a $C_{\mathcal{T}\mathcal{D}u} = \{\mathrm{C_{u1}}, \mathrm{C_{u2}}, \ldots, \mathrm{C}_{\mathrm{u}n}\}$ – skończony zbiór kategorii użytych w konceptualizacji $\mathcal{T}_{\mathcal{D}u}$. Powyższym kategoriom nadano unikalne nazwy, zwane atrybutami zbiorowymi, pochodzące ze zbioru $\mathbf{A_S} = \{\mathrm{A_{s1}}, \mathrm{A_{s2}}, \ldots, \mathrm{A}_{\mathrm{s}n}\}$.

Definicja 5.1. Schematem danych w dziedzinie $\mathcal{D}_u$, zgodnym z konceptualizacją $\mathcal{T}_{\mathcal{D}u}$, nazywa się dowolny niepusty podzbiór $\mathbf{S_S} = \{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$ zbioru atrybutów zbiorowych $\mathbf{A_S}$.

Definicja 5.2. Formułą atomową z atrybutem zbiorowym A_{si} nazywa się dowolne wyrażenie w postaci:

$$\mathrm{A}_{si} \in \{v_1, v_2, \dots, v_k\}, \tag{5.1}$$

$$\neg(\mathrm{A}_{si} \in \{v_1, v_2, \dots, v_k\}), \tag{5.2}$$

$$\{v_1, v_2, \dots, v_k\} \subseteq \mathrm{A}_{si}, \tag{5.3}$$

$$\text{lub} \quad \neg(\{v_1, v_2, \dots, v_k\} \subseteq \mathrm{A}_{si}), \tag{5.4}$$

gdzie $0 \le k \le |\mathrm{C}_{\mathrm{u}i}|$, a v_1, v_2, …, v_k oznaczają parami różne wartości – przykłady kategorii $\mathrm{C}_{\mathrm{u}i}$. Dla ujednolicenia zapisu wyrażeń 5.1 – 5.4, przyjęto dla nich wspólną notację:

$$\mathrm{A}_{si} \blacksquare \{v_1, v_2, \dots, v_k\} q \tag{5.5}$$

w której $\blacksquare$ oznacza symbol relacyjny $=$ lub $\neq$, a q – kwalifikator $\odot$ lub $\oplus$. Powyższy kwalifikator nadaje zbiorowi $\{v_1, v_2, \dots, v_k\}$ interpretację, odpowiednio koniunkcyjną lub dysjunkcyjną. Znaczenie formuły (5.5) dla poszczególnych kombinacji wartości $\blacksquare$ i q jest przy tym następujące: $\oplus$ i $=$ – wyrażenie 5.1, $\oplus$ i $\neq$ – wyrażenie 5.2, $\odot$ i $=$ – wyrażenie 5.3, $\odot$ i $\neq$ – wyrażenie 5.4. Zbiór wartości z kwalifikatorem $\{v_1, v_2, \dots, v_k\} q$ nazywa się w dalszym ciągu zbiorem kwalifikowanym i stosuje dla niego oznaczenia vq, vq_1, vq_2, itd.

Definicja 5.3. Regułą 2U opartą na schemacie danych $\mathbf{S_S} = \{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$ nazywa się dowolne wyrażenie w postaci:

$$\text{it is declared with } \mathrm{grf}(p_r): P_1, P_2, \dots, P_d \ \rightarrow C \text{ with } \mathrm{irf}(p_c) \tag{5.6}$$

gdzie $0 \le d \le m$, P_1, P_2, …, P_d, C oznaczają formuły atomowe z parami różnymi atrybutami ze zbioru $\{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$, irf oraz grf – nazwy współczynników wiarygodności, odpowiednio wewnętrznego i zewnętrznego, a p_c i p_r – wartości współczynników, odpowiednio irf i grf, pochodzące z przedziału $(0;\ 1\rangle$. W dalszym tekście, regułę 2U zapisuje się w skróconej postaci:

$$\mathrm{grf}(p_r): P_1, P_2, \dots, P_d \ \rightarrow C \text{ with } \mathrm{irf}(p_c). \tag{5.7}$$

Reguła wyraża dodatnią monotoniczną zależność między przesłanką $P_1 \wedge P_2 \wedge \dots \wedge P_d$ a konkluzją C. Użyte w regule współczynniki wiarygodności irf i grf oznaczają, odpowiednio: stopień pewności konkluzji C przy założeniu 100% pewności przesłanki $P_1 \wedge P_2 \wedge \dots \wedge P_d$, oraz stopień pewności (wiarygodności) reguły jako całości w kontekście dostępnej ewidencji – zbioru danych z dziedziny $\mathcal{D}_u$, zbudowanych według schematu $\mathbf{S_S} = \{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$.

Współczynnik irf służy w regule 2U do oznaczenia poziomu niepewności pierwszego rzędu, a współczynnik grf – poziomu niepewności drugiego rzędu. Współczynniki irf i grf stanowią odpowiedniki częstotliwości i zaufania w regule Wanga.

Definicja 5.4. Faktem 2U opartym na schemacie danych $\mathbf{S_S} = \{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$ nazywa się dowolną formułę atomową z atrybutem zbiorowym ze zbioru $\{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$. Dla ujednolicenia rozważań fakt ten można wyrazić jako regułę 2U:

$$\mathrm{grf}(p_r)\colon \mathrm{true} \rightarrow \mathrm{A}_{sij} \blacksquare \{v_1, v_2, \dots, v_k\} q \text{ with } \mathrm{irf}(p_c) \tag{5.8}$$

ze stałą logiczną true w roli przesłanki i rozważaną formułą atomową w roli konkluzji, gdzie $1 \leq j \leq m$, a pozostałe symbole zachowują znaczenia z definicji 5.2 i 5.3.

Zakłada się, że reguły 2U będą pozyskiwane z danych zgromadzonych na drodze doświadczalnej i zapisanych w formacie atrybutowym (patrz podrozdz. 6.1.1).

Reguła 2U, choć posługuje się do oznaczenia stopnia pewności podobnymi parametrami jak reguła Wanga, ma większą od niej siłę wyrazu. Zapewnia ją:
- używanie zbiorów kwalifikowanych w miejsce wartości pojedynczych,
- używanie w formułach dwóch operatorów relacyjnych $\in$ i $\subseteq$ w miejsce jednego równościowego = ,
- dopuszczenie w formułach operatora negacji $\neg$,
- użycie w przesłance reguły koniunkcji formuł atomowych w miejsce pojedynczej formuły atomowej.

Posługiwanie się w formułach atomowych zbiorami kwalifikowanymi powoduje, że te formuły stają się złożone. Kwalifikator zbioru $\odot$ oznacza, że rozważany atrybut przyjmuje równolegle wszystkie wartości z kwalifikowanego zbioru (stwierdzenie równoważne koniunkcji formuł elementarnych), a kwalifikator zbioru $\oplus$ – że przyjmuje on jedną z wartości wymienionych w kwalifikowanym zbiorze (stwierdzenie równoważne dysjunkcji formuł elementarnych). Podczas gdy formuła zapisana z użyciem kwalifikatora $\oplus$ przypomina hipotezę z teorii Dempstera-Shafera, to formuła zapisana z użyciem kwalifikatora $\odot$ – ma charakter nowatorski. Szerokie uzasadnienie potrzeby posługiwania się oboma kwalifikatorami można znaleźć w rozdziale 8, prezentującym możliwości budowy medycznych systemów eksperckich w oparciu o reguły 2U.

Konsekwencją użycia zbiorów kwalifikowanych jest postać zbioru operatorów relacyjnych $\{\in, \subseteq\}$, stanowiącego podzbiór właściwy odpowiadającego mu zbioru z logiki ALSV(FD) (ang. Attributive Logic with Set Values) [103]. Pozostałych operatorów relacyjnych z logiki ALSV nie wzięto pod uwagę ze względu na ich słabą aplikowalność (np. operator =) lub ze względu na trudności implementacyjne (np. operator $\sim$) .

O operatorze negacji i jego wpływie na siłę wyrazu reguły była mowa w podrozdziale 5.1. Z drugiej strony, szczególną rolę tego operatora należy rozpatrywać w kontekście założeń przyjmowanych odnośnie do zasad wnioskowania w systemie RBS(2U) (patrz podrozdz. 5.4).

W zamyśle, także parametr wiarygodności grf ma głębsze znaczenie od parametru zaufania w regule Wanga. Przyjmuje się, że grf może uwzględniać różne aspekty ewidencji, z której pozyskano regułę, oprócz wielkości ewidencji – także stopień jej homogeniczności, czy źródło pochodzenia. Zgodnie z planem, parametr ten ma odgrywać ważną rolę podczas wnioskowania w systemach z regułami 2U. W szczególności, w procesach wnioskowania w przód może służyć do rozstrzygania konfliktów w agendzie.

5.3 Baza wiedzy i baza danych systemu RBS(2U)

5.3.1 Projektowanie bazy wiedzy

Zbiór reguł 2U pozyskanych z ewidencji atrybutowej o schemacie $\mathbf{S_S}$ stanowi punkt wyjścia do konstrukcji bazy wiedzy systemu regułowego RBS(2U) dedykowanego dziedzinie $\mathcal{D}_{\mathcal{U}}$. Aby z niego uzyskać bazę wiedzy spełniającą wysokie wymagania jakościowe, należy:

– zweryfikować dla każdej reguły charakter zależności łączącej jej przesłankę i konkluzję: jeśli zależność nie ma charakteru monotonicznego, to taką regułę należy usunąć ze zbioru; jeśli ma charakter monotoniczny ujemny, to regułę należy „odwrócić” poprzez zanegowanie konkluzji i odpowiednią zmianę parametrów liczbowych; jeśli ma charakter monotoniczny dodatni, to regułę należy pozostawić w zbiorze w niezmienionej postaci;
– usunąć ze zbioru pary reguł sprzecznych; ewentualnie, po konsultacji z ekspertem – reguły uznane za poprawne przywrócić do zbioru;
– usunąć ze zbioru reguły, dla których można wskazać w zbiorze reguły subsumujące.

Weryfikacja charakteru zależności pomiędzy przesłanką a konkluzją reguły jest problemem ogólnym, dotyczącym każdego regułowego systemu z niepewnością. Podstawowa trudność weryfikacji polega na tym, że należy ją prowadzić globalnie, na podstawie równoczesnej obserwacji całej bazy wiedzy. Szczegółowy przebieg weryfikacji zależy od formatu reguły z niepewnością, czyli od liczby i znaczenia parametrów liczbowych definiujących stopień pewności.

Definicja 5.5. Niech $\leq_{\mathrm{AS}}$ oznacza relację porządku częściowego zdefiniowaną na zbiorze formuł atomowych z atrybutami zbiorowymi ze zbioru $\mathbf{A_S}$. Formuła P_i pozostaje w relacji porządku częściowego $\leq_{\mathrm{AS}}$ z formułą P_j wtedy i tylko wtedy, gdy:

– $P_i \overset{\text{def}}{=} (\mathrm{A}_{si} \in \{v_{i1}, v_{i2}, \dots, v_{iki}\})$ i $P_j \overset{\text{def}}{=} \left(\mathrm{A}_{sj} \in \left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\}\right)$ i
$\mathrm{A}_{si} = \mathrm{A}_{sj}$ i $\left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\} \subseteq \{v_{i1}, v_{i2}, \dots, v_{iki}\}$, lub

– $P_i \overset{\text{def}}{=} \neg(\mathrm{A}_{si} \in \{v_{i1}, v_{i2}, \dots, v_{iki}\})$ i $P_j \overset{\text{def}}{=} \neg\left(\mathrm{A}_{sj} \in \left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\}\right)$ i
$\mathrm{A}_{si} = \mathrm{A}_{sj}$ i $\{v_{i1}, v_{i2}, \dots, v_{iki}\} \subseteq \left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\}$, lub

– $P_i \overset{\text{def}}{=} (\{v_{i1}, v_{i2}, \dots, v_{iki}\} \subseteq \mathrm{A}_{si})$ i $P_j \overset{\text{def}}{=} \left(\left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\} \subseteq \mathrm{A}_{sj}\right)$ i
$\mathrm{A}_{si} = \mathrm{A}_{sj}$ i $\{v_{i1}, v_{i2}, \dots, v_{iki}\} \subseteq \left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\}$, lub

– $P_i \overset{\text{def}}{=} (\{v_{i1}, v_{i2}, \dots, v_{iki}\} \subseteq \mathrm{A}_{si})$ i $P_j \overset{\text{def}}{=} \left(\left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\} \subseteq \mathrm{A}_{sj}\right)$ i
$\mathrm{A}_{si} = \mathrm{A}_{sj}$ i $\left\{v_{j1}, v_{j2}, \dots, v_{jkj}\right\} \subseteq \{v_{i1}, v_{i2}, \dots, v_{iki}\}$, lub

– $P_i \overset{\text{def}}{=} (\mathrm{A}_{si} \in \{v_{i1}\})$ i $P_j \overset{\text{def}}{=} \left(\left\{v_{j1}\right\} \subseteq \mathrm{A}_{sj}\right)$ i
$\mathrm{A}_{si} = \mathrm{A}_{sj}$ i $v_{i1} = v_{j1}$.

Definicja 5.6. Niech R_i i R_j oznaczają dwie reguły 2U oparte na schemacie danych $\mathbf{S_S} = \{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$, zredagowane w postaci:

$$R_i \overset{\text{def}}{=} \mathrm{grf}(p_{ri}) \colon P_{i1}, P_{i2}, \dots, P_{idi} \to C_i \text{ with } \mathrm{irf}(p_{ci}) \tag{5.9}$$

$$\text{i } R_j \overset{\text{def}}{=} \mathrm{grf}\left(p_{rj}\right) \colon P_{j1}, P_{j2}, \dots, P_{jdj} \to C_j \text{ with } \mathrm{irf}\left(p_{cj}\right) \tag{5.10}$$

i spełniające zależność:

$$\bigwedge\nolimits_{1 \le k \le di} \bigvee\nolimits_{1 \le l \le dj} \left(P_{ik} \le_{\mathrm{AS}} P_{jl}\right) \wedge \left(C_i = C_j\right) \tag{5.11}$$

$$\text{lub } \bigwedge\nolimits_{1 \le l \le dj} \bigvee\nolimits_{1 \le k \le di} \left(P_{jl} \le_{\mathrm{AS}} P_{ik}\right) \wedge \left(C_i = C_j\right) \tag{5.12}$$

Jeśli przy tym zachodzi relacja: $p_{ci} > p_{cj}$ w wypadku 5.11, lub $p_{cj} > p_{ci}$ w wypadku 5.12, to przynajmniej jedna z reguł R_i i R_j jest zbudowana nieprawidłowo – nie wyraża zależności dodatniej monotonicznej.
W zapisach reguł, P_{ik} $(1 \le k \le d_i)$, C_i, P_{jl} $(1 \le j \le d_j)$ i C_j oznaczają formuły atomowe z atrybutami zbiorowymi ze zbioru $\mathbf{S_S}$, p_{ci} i p_{cj} – wartości współczynników irf, a p_{ri} i p_{rj} – wartości współczynników grf.

W celu zweryfikowania poprawności budowy reguły R_i należy wziąć pod uwagę wszystkie te reguły R_j z wyjściowego zbioru reguł, które pozostają z nią w zależności 5.11 lub 5.12. Ostateczną decyzję można podjąć na podstawie procentowego udziału zależności prawidłowych w ogólnej liczbie wszystkich tych zależności.

Kolejnym istotnym krokiem na drodze do budowy wysokiej jakości bazy wiedzy jest detekcja i usuwanie par reguł sprzecznych. Realizacja tego procesu ma charakter obowiązkowy.

Definicja 5.7. Niech R_i i R_j oznaczają dwie reguły 2U oparte na schemacie danych $\mathbf{S_S} = \{A_{si1}, A_{si2}, \dots, A_{sim}\}$ i zredagowane w postaci, odpowiednio 5.9 i 5.10. Reguły R_i i R_j pozostają w sprzeczności wtedy i tylko wtedy, gdy:

– zbiory formuł atomowych $\{P_{i1}, P_{i2}, \dots, P_{idi}\}$ i $\{P_{j1}, P_{j2}, \dots, P_{jdj}\}$ są identyczne oraz

– $(C_i \leq_{\mathrm{AS}} C_j$ i $p_{ci} < p_{cj})$ lub $(C_j \leq_{\mathrm{AS}} C_i$ i $p_{cj} < p_{ci})$.

Ostatnią ważną czynnością projektową jest oczyszczanie bazy wiedzy z reguł nadmiarowych. Choć nie jest to czynność obligatoryjna, to jednak warto ją przeprowadzić z uwagi na wydajność późniejszych procesów wnioskowania.

Definicja 5.8. Niech R_i i R_j oznaczają dwie reguły 2U oparte na schemacie danych $\mathbf{S_S} = \{A_{si1}, A_{si2}, \dots, A_{sim}\}$ i zredagowane w postaci, odpowiednio 5.9 i 5.10. Reguła R_i subsumuje regułę R_j wtedy i tylko wtedy, gdy:

– zbiory formuł atomowych $\{P_{i1}, P_{i2}, \dots, P_{idi}\}$ i $\{P_{j1}, P_{j2}, \dots, P_{jdj}\}$ nie są identyczne,

– $\bigwedge_{1 \leq k \leq di} \bigvee_{1 \leq l \leq dj} \left(P_{ik} \leq_{\mathrm{AS}} P_{jl}\right)$,

– $C_i = C_j$ i $p_{ci} = p_{cj}$ oraz

– $p_{ri} \geq p_{rj}$.

Po zweryfikowaniu zależności między przesłankami i konkluzjami reguł oraz usunięciu par reguł sprzecznych i reguł nadmiarowych, bazę wiedzy systemu RBS(2U) można poddać ocenie za pomocą metryki jakości 4.11. Funkcja avg przyjmuje w tym wypadku postać średniej arytmetycznej z wartości współczynników grf wszystkich reguł.

Przykładowy algorytm konstrukcji bazy wiedzy systemu RBS(2U) ze zbiorczych danych atrybutowych zamieszczono w podrozdziale 6.2.

5.3.2 Obsługa bazy danych

Baza danych systemu RBS(2U) jest zbiorem faktów w postaci 5.8. Nowy stan bazy danych jest ustalany na początku każdego procesu wnioskowania, tak progresywnego, jak i regresywnego. W wypadku wnioskowania progresywnego, stan ten będzie podlegał obowiązkowym aktualizacjom w trakcie prowadzenia wnioskowania i opcjonalnej zmianie normalizacyjnej po jego zakończeniu.

Zgodnie z zaleceniem sformułowanym w podrozdziale 4.4, w bazie danych systemu RBS(2U) będą przechowywane tylko fakty z konkluzjami w formie prostej, czyli niezanegowanej. Za spełnienie tego wymogu odpowiadają algorytmy wnioskowania progresywnego i regresywnego, w tym – definicje funkcji propagacji niepewności (patrz: podrozdział 5.4). Pomimo takiej gwarancji, stan bazy danych w procesie wnioskowania progresywnego powinien być przez cały czas monitorowany przez podsystem utrzymywania wiarygodności. Ta potrzeba wynika ze specyfiki formuł atomowych w faktach/regułach 2U, używających wartości zbiorowych w miejsce tradycyjnych wartości pojedynczych. Podobnie jak nie należy dopuszczać do występowania w bazie wiedzy par reguł sprzecznych, tak też nie należy dopuszczać do równoczesnego występowania w bazie danych par faktów sprzecznych.

Definicja 5.9. Niech F_i i F_j oznaczają dwa fakty 2U oparte na schemacie danych $\mathbf{S_S} = \{\mathrm{A}_{si1}, \mathrm{A}_{si2}, \dots, \mathrm{A}_{sim}\}$ i zredagowane w postaci:

$$F_i \stackrel{\text{def}}{=} \mathrm{grf}(p_{ri}): true \rightarrow C_i \text{ with } \mathrm{irf}(p_{ci}) \tag{5.13}$$

$$\text{i } F_j \stackrel{\text{def}}{=} \mathrm{grf}(p_{rj}): true \rightarrow C_j \text{ with } \mathrm{irf}(p_{cj}) . \tag{5.14}$$

Fakty F_i i F_j pozostają w sprzeczności wtedy i tylko wtedy, gdy:

– istnieje taki ciąg formuł atomowych $C_{k1}, \dots, C_{kl}$, które pozostają w relacjach: $C_i \leq_{\mathrm{AS}} C_{k1} \leq_{\mathrm{AS}} \dots \leq_{\mathrm{AS}} C_{kl} \leq_{\mathrm{AS}} C_j$ i zachodzi zależność: $p_{ci} < p_{cj}$, lub
– istnieje taki ciąg formuł atomowych $C_{k1}, \dots, C_{kl}$, które pozostają w relacjach: $C_j \leq_{\mathrm{AS}} C_{k1} \leq_{\mathrm{AS}} \dots \leq_{\mathrm{AS}} C_{kl} \leq_{\mathrm{AS}} C_i$ i zachodzi zależność: $p_{ci} > p_{cj}$.

Wykrycie takiej sprzeczności powinno uruchamiać natychmiastowy proces jej eliminacji. Może on być wykonywany automatycznie lub pół-automatycznie, przy współudziale użytkownika. Podstawą rozstrzygnięć w procesie automatycznym może być wartość współczynnika grf: z dwóch znajdujących się w bazie danych i pozostających w sprzeczności faktów 2U należy raczej usunąć ten, który ma mniejszą wartość tego współczynnika.

5.4 Wnioskowanie w systemie RBS(2U)

5.4.1 Założenia

Wnioskowanie w systemie RBS(2U) opiera się na kilku istotnych założeniach. Przede wszystkim, przyjmuje się, że będzie ono prowadzone przy założeniu otwartego świata OWA.

Kolejne założenia są związane z interpretacją parametrów liczbowych używanych w regule/fakcie 2U: współczynnika wiarygodności irf i współczynnika wiarygodności grf. W istocie, parametry te odnoszą się do różnych właściwości ewidencji, z której pozyskano regułę. Dlatego nic nie stoi na przeszkodzie, by silnik wnioskujący systemu używał ich niezależnie od siebie, gdy zachodzi taka potrzeba. Taką odrębną, ważną rolę przewidziano w systemie dla współczynnika grf. W procesie wnioskowania w przód będzie on sterował procesem konstrukcji agendy, lecz także – brał aktywny udział przy rozstrzyganiu konfliktów w agendzie.

Ze względu na planowaną, pół-automatyczną akwizycję wiedzy z ewidencji danych należy zakładać, że p_c i p_r będą wyrażać oszacowanie, a nie dokładne wartości współczynników irf i grf reguły w postaci 5.7. Z drugiej strony, te wartości muszą pochodzić z pewnych ustalonych zakresów $\langle Bel_c; Pls_c\rangle$ i $\langle Bel_r ; Pls_r\rangle$ (patrz podrozdz. 3.5.2). Oczywiście, takim samym ograniczeniom podlegają wartości współczynników we wszystkich faktach 2U, które będą generowane lub weryfikowane w trakcie prowadzonego wnioskowania. To zaś oznacza konieczność takiego zdefiniowania funkcji propagacji niepewności, które gwarantuje spełnienie tych ograniczeń.

Aby zrealizować podane założenia przyjęto, że w procesie wnioskowania progresywnego:

- kwalifikacja do agendy reguły $\mathrm{grf}(p_{ri})\colon P_1, \dots, P_n \ \rightarrow C$ with $\mathrm{irf}(p_{ci})$ będzie się odbywać przy spełnieniu warunków:
  - wartość p_{ri} jest nie niższa od wymaganej wartości progowej τ_1,
  - dla każdej zawartej w przesłance $P_1, \dots, P_n$ formuły atomowej P_i $(1 \leq i \leq n)$ w formie prostej (niezanegowanej), w bazie danych znajduje się fakt $\mathrm{grf}(p_{rj})\colon \mathrm{true} \rightarrow P_j$ with $\mathrm{irf}(p_{cj})$ taki, że:
    - wartość p_{rj} jest nie niższa od wymaganej wartości progowej τ_1,
    - można wskazać ciąg formuł atomowych w formie prostej P_{k1}, …, P_{kl} spełniających relacje: $P_i \leq_{\mathrm{AS}} P_{k1} \leq_{\mathrm{AS}} \dots P_{kl} \leq_{\mathrm{AS}} P_j$,
    - wartość p_{cj} jest nie niższa od wymaganej wartości progowej τ_2;
  - dla każdej zawartej w przesłance $P_1, \dots, P_n$ formuły atomowej P_i $(1 \leq i \leq n)$ w formie zanegowanej $(\neg P_i')$, w bazie danych znajduje się fakt $\mathrm{grf}(p_{rj})\colon \mathrm{true} \rightarrow P_j$ with $\mathrm{irf}(p_{cj})$ taki, że:
    - wartość p_{rj} jest nie niższa od wymaganej wartości progowej τ_1,
    - można wskazać ciąg formuł atomowych w formie prostej P_{k1}, …, P_{kl} spełniających relacje: $P_j \leq_{\mathrm{AS}} P_{k1} \leq_{\mathrm{AS}} \dots \leq_{\mathrm{AS}} P_{kl} \leq_{\mathrm{AS}} P_i'$,
    - wartość p_{cj} jest nie wyższa od wymaganej wartości progowej τ_2;
- konflikt w agendzie będzie rozstrzygany na rzecz reguły o najwyższej wartości współczynnika grf; jeśli zastosowanie kryterium grf nie będzie wystarczające do ustalenia kolejności reguł w agendzie, to zostanie wzięta pod uwagę dodatkowa cecha reguły, np. jej pozycja porządkowa w bazie wiedzy;
- przy uaktywnianiu reguły $\mathrm{grf}(p_{ri})\colon P_1, \dots, P_i, \dots, P_n \ \rightarrow \ C$ with $\mathrm{irf}(p_{ci})$, spośród wszystkich faktów w bazie danych spełniających wymagania formuły atomowej P_i $(1 \leq i \leq n)$ w formie prostej (niezanegowanej), w procesie propagacji niepewności zostanie uwzględniony fakt z najwyższą wartością współczynnika irf;

– przy uaktywnianiu reguły grf(p_{ri}): $P_1, \dots, \neg P_i', \dots, P_n \ \rightarrow C$ with irf(p_{ci}), spośród wszystkich faktów w bazie danych spełniających wymagania formuły atomowej P_i' $(1 \leq i \leq n)$ w formie prostej (niezanegowanej), w procesie propagacji niepewności zostanie uwzględniony fakt z najniższą wartością współczynnika irf.

Konkludując, na danym etapie wnioskowania progresywnego będzie uaktywniana najbardziej wiarygodna spośród tych reguł, które spełniają minimalne żądania dotyczące istnienia i szczegółowej postaci faktów odpowiadających formułom atomowym z przesłanki reguły. Dla każdej formuły atomowej, spośród wszystkich odpowiadających jej faktów, w procesie propagacji niepewności będą uwzględniany ten, które jest do niej najbardziej „podobny".

Skutek uaktywnienia reguły będzie zależał od stanu bazy danych na danym etapie procesu wnioskowania. Przyjąwszy, że reguła jest w postaci 5.7, należy rozpatrzyć następujące trzy przypadki:

– fakt z konkluzją C (negacją konkluzji C) znajduje się w bazie danych i jest oznaczony jako aksjomat,
– nie ma w bazie danych faktu z konkluzją C (negacją konkluzji C),
– fakt z konkluzją C (negacją konkluzji C) znajduje się w bazie danych i nie ma żadnego oznaczenia (nie jest aksjomatem, lecz został wygenerowany w poprzednich krokach procesu wnioskowania).

W wymienionych sytuacjach silnik wnioskujący zadziała jak następuje:

– nie podejmie żadnego działania – fakty aksjomatyczne to dane wejściowe dla procesu wnioskowania; w związku z tym, nie powinny podlegać żadnym modyfikacjom;
– wstawi odpowiedni, nie-aksjomatyczny fakt z konkluzją C (negacją konkluzji C) do bazy danych,
– zastąpi w bazie danych nie-aksjomatyczny fakt z konkluzją C (negacją konkluzji C) nowym faktem z tą samą konkluzją.

W drugim i trzecim wypadku, konstrukcja faktu z konkluzją C (negacją konkluzji C) będzie się odbywać przy wykorzystaniu odpowiednio zaimplementowanych funkcji propagacji niepewności f_{1i}, f_{2i}, f_{3i}, f_{4i} $(1 \leq i \leq 2)$ oraz ustalonej wartości t (patrz podrozdz. 5.4.2). Wszystkie planowane zmiany w bazie danych powinny być poddane ocenie systemu utrzymywania wiarygodności.

Wnioskowanie regresywne będzie przebiegać zgodnie z uniwersalnym schematem przyjętym dla systemów regułowych z niepewnością. Szczegółowe założenia tyczące tego procesu będą zrealizowane poprzez odpowiednią implementację funkcji propagacji niepewności f_{5i}, f_{60}, f_{6jk}, f_{7i} $(0 \leq i \leq 2,\ 1 \leq j \leq n,\ 1 \leq k \leq 2)$ (patrz podrozdz. 5.4.3). Spośród możliwych operatorów relacyjnych, w pytaniach będą używane jedynie $\leq$ i $\geq$.

5.4.2 Propagacja niepewności w procesie wnioskowania progresywnego

Podana w poprzednim podrozdziale charakterystyka działania silnika wnioskującego w systemie RBS(2U) wymaga uzupełnienia o definicje funkcji propagacji niepewności. Oto definicje funkcji używanych w procesie wnioskowania progresywnego.

Propagacja niepewności przesłanek składowych do przesłanki złożonej. W systemie RBS(2U) dopuszczono przesłanki złożone, sformułowane w postaci koniunkcyjnej. Aby określić stopień pewności całej przesłanki F_v: $\mathrm{grf}(p_{rv})\colon true \rightarrow P_1 \wedge P_2 \wedge \ldots \wedge P_n$ with $\mathrm{irf}(p_{cv})$ na podstawie znanych stopni pewności n faktów składowych F_i: $\mathrm{grf}(p_{ri})\colon true \rightarrow P_i$ with $\mathrm{irf}(p_{ci})$ $(1 \le i \le n)$, należy skorzystać z poniższych funkcji f_{11} i f_{12}:

$$f_{11}, f_{12}\colon ([0;1] \times [0;1])^n \rightarrow [0;1]\ ;$$

$$f_{11}(p_{c1}, p_{r1}, p_{c2}, p_{r2}, \ldots, p_{cn}, p_{rn}) = \min\{\mathrm{p}_{c1}, \mathrm{p}_{c2}, \ldots, \mathrm{p}_{cn}\}\ ;$$

$$f_{12}(p_{c1}, p_{r1}, p_{c2}, p_{r2}, \ldots, p_{cn}, p_{rn}) = \min\{\mathrm{p}_{r1}, \mathrm{p}_{r2}, \ldots, \mathrm{p}_{rn}\}.$$

Z powyższych wzorów wynika, że współczynniki irf i grf przesłanki złożonej, reprezentującej wirtualny fakt P_v, oblicza się jako minimum z wartości współczynników składowych, odpowiednio irf i grf. Zatem, każda z funkcji f_{11} i f_{12} ma w istocie n jednorodnych parametrów. Obliczeniom tych funkcji nie towarzyszy ani modyfikacja, ani tworzenie rzeczywistego faktu 2U – analizowana przesłanka pełni rolę faktu wirtualnego.

Zgodnie z założeniem, wartość p_{cv} współczynnika grf wirtualnego faktu F_v powinna pochodzić z zakresu $\langle Bel_{cv}; Pls_{cv}\rangle$ takiego, że: $Bel_{cv} = = \max\{Bel_{c1} + \ldots + Bel_{cn} - (n-1), 0\}$ i $Pls_{cv} = \min\{Pls_{c1}, \ldots, Pls_{cn}\}$, gdzie Bel_{ci}, Pls_{ci} $(1 \le i \le n)$ oznaczają, odpowiednio, globalny stopień przekonania i stopień możliwości współczynnika irf i-tego faktu składowego F_i. Podobnie, wartość p_{rv} współczynnika grf faktu wirtualnego F_v powinna pochodzić z zakresu $\langle Bel_{rv}; Pls_{rv}\rangle$ takiego, że: $Bel_{rv} = \min\{Bel_{r1}, \ldots, Bel_{rn}\}$ i $Pls_{rv} = \min\{Pls_{r1}, \ldots, Pls_{rn}\}$, gdzie Bel_{ri}, Pls_{ri} $(1 \le i \le n)$ oznaczają, odpowiednio, globalny stopień przekonania i stopień możliwości współczynnika grf i-tego faktu składowego F_v.

Zgodnie z tym samym założeniem, prawdziwe są nierówności: $Bel_{ci} \le p_{ci} \le Pls_{ci}$ oraz $Bel_{ri} \le p_{ri} \le Pls_{ri}$. W takiej sytuacji, można łatwo pokazać, że: $p_{cv} \ge Bel_{cv}$. Skoro $p_{cv} = \min\{\mathrm{p}_{c1}, \mathrm{p}_{c2}, \ldots, \mathrm{p}_{cn}\}$, to $p_{cv} = p_{ci}$ dla pewnego i z przedziału $\langle 1; n\rangle$. Z tego wynika, że: $p_{cv} = p_{ci} \ge Bel_{ci} \ge Bel_{ci} + x$, gdzie $x \le 0$. Kontynuując, $p_{cv} \ge Bel_{ci} + \big((Bel_{c1} - 1) + \cdots + (Bel_{ci-1} - 1) + + (Bel_{ci+1} - 1) + \cdots + (Bel_{cn} - 1)\big) = Bel_{c1} + \cdots + Bel_{cn} - (n-1)$. Z drugiej strony, skoro $p_{cv} = \min\{\mathrm{p}_{c1}, \mathrm{p}_{c2}, \ldots, \mathrm{p}_{cn}\}$, to zachodzi: $p_{cv} \ge 0$. Podsumowując, $p_{cv} \geq \max\{Bel_{ci} + \cdots + Bel_{cn} - (n-1), 0\} = Bel_{cv}$. Prawdziwość nierówności $p_{cv} \leq Pls_{cv}$, a także nierówności $Bel_{rv} \leq p_{rv} \leq Pls_{rv}$ można łatwo wykazać z definicji funkcji f_{11} i f_{12}.

Propagacja niepewności przesłanki reguły – przez regułę – do nowej hipotezy. Podstawowym działaniem wykonywanym w procesie wnioskowania w przód jest uaktywnianie reguł. Jeśli uaktywniana reguła R ma postać $\mathrm{grf}(p_{rr}): P \rightarrow C \text{ with } \mathrm{irf}(p_{cr})$, a oba współczynniki wiarygodności faktu F: $\mathrm{grf}(p_{rp}): true \rightarrow P \text{ with } \mathrm{irf}(p_{cp})$ przyjmują wartość 1, to w wyniku uzyskuje się fakt cząstkowy F_i: $\mathrm{grf}(p_{ri}): true \rightarrow C \text{ with } \mathrm{irf}(p_{ci})$. W większości wypadków, wartości współczynników p_{cp} i p_{rp} są jednak różne od 0. Dlatego, przy uaktywnianiu reguły zachodzi potrzeba skorzystania z poniższych funkcji f_{21} i f_{22}:

$$f_{21}, f_{22}: ([0;1] \times [0;1])^2 \rightarrow [0;1] ;$$

$$f_{21}(p_{cr}, p_{rr}, p_{cp}, p_{rp}) = p_{cr} \cdot p_{cp} ;$$

$$f_{22}(p_{cr}, p_{rr}, p_{cp}, p_{rp}) = \min\{p_{rr}, p_{rp}\} .$$

W istocie, każda z funkcji f_{21} i f_{22} zależy nie od czterech, ale od dwóch jednorodnych parametrów. Zależnie od stanu bazy danych, utworzony fakt cząstkowy F_i ignoruje się, wstawia w uzyskanej (zanegowanej) postaci do bazy danych, lub dołącza do występującego tam faktu z konkluzją C (negacją konkluzji C). Łatwo zauważyć, że wraz ze wzrostem długości łańcucha wnioskowania, maleją współczynniki irf i grf kolejnych generowanych faktów cząstkowych.

Przyjmuje się, że wartość p_{ci} współczynnika irf faktu F_i powinna pochodzić z zakresu $\langle Bel_{ci}; Pls_{ci} \rangle$ takiego, że: $Bel_{ci} = Bel_{cr} \cdot Bel_{cp}$ i $Pls_{ci} = Pls_{cr} \cdot Pls_{cp}$, gdzie Bel_{cr}, Bel_{cp} oznaczają globalny stopień przekonania współczynnika irf, odpowiednio reguły R i faktu F, a Pls_{rr}, Pls_{rp} – stopień możliwości współczynnika grf, odpowiednio reguły R i faktu F. Przy obowiązujących założeniach $Bel_{cr} \leq p_{cr} \leq Pls_{cr}$ oraz $Bel_{cp} \leq p_{cp} \leq Pls_{cp}$, można łatwo dowieść spełnienia tego wymogu, czyli: $Bel_{ci} \leq p_{ci} \leq Pls_{ci}$. Podobne rozumowanie da się przeprowadzić dla współczynnika grf. W tym wypadku zakłada się, że: $Bel_{ri} = \min\{Bel_{rr}, Bel_{rp}\}$ oraz $Pls_{ri} = \min\{Pls_{rr}, Pls_{rp}\}$. Jeśli przy tym, zgodnie z założeniem, prawdziwe są nierówności: $p_{rr} \geq Bel_{rr}$ oraz $p_{rp} \geq Bel_{rp}$, to: $p_{ri} = \min\{p_{rr}, p_{rp}\} \geq \min\{Bel_{rr}, Bel_{rp}\} = Bel_{ri}$. W podobny sposób można dowieść nierówności $p_{ci} \leq Pls_{ci}$.

Zaproponowane definicje funkcji f_{21} i f_{22} pozostają w zgodności z kluczowym założeniem na temat dodatniej monotonicznej zależności między przesłanką reguły a jej konkluzją.

Propagacja niepewności hipotez wielokrotnych do hipotezy zbiorczej. Jeśli na pewnym etapie procesu wnioskowania w przód został wygenerowany fakt cząstkowy F_j: $\mathrm{grf}(p_{rj})$: $true \to C$ with $\mathrm{irf}(p_{cj})$ z konkluzją C taką, że w bazie danych znajduje się nie-aksjomatyczny fakt F_i: $\mathrm{grf}(p_{ri})$: $true \to C$ with $\mathrm{irf}(p_{ci})$ z tą samą konkluzją, to silnik wnioskujący zastąpi w bazie danych fakt F_i – nowym faktem F_k z konkluzją C: $\mathrm{grf}(p_{rk})$: $true \to C$ with $\mathrm{irf}(p_{ck})$ (zgodnie z przyjętym założeniem, w opisanej sytuacji konkluzja C musi być zredagowana w formie prostej, czyli niezanegowanej). Powyższą akcję silnik przeprowadzi przy wykorzystaniu następujących funkcji f_{31} i f_{32}:

f_{31}, f_{32}: $([0;1] \times [0;1])^2 \to [0;1]$;

$f_{31}(p_{ci}, p_{ri}, p_{cj}, p_{rj}) = (1 - w_{n+1}) \cdot p_{ci} + w_{n+1} \cdot p_{cj}$;

$f_{32}(p_{ci}, p_{ri}, p_{cj}, p_{rj}) = \max\{p_{ri}, p_{rj}\}$,

gdzie $n + 1$ ($n \geq 1$) oznacza numer porządkowy faktu cząstkowego z konkluzją C wygenerowanego w bieżącym kroku procesu wnioskowania (fakt F_j), a w_{n+1} – wagę wyjściową tego faktu, obliczaną ze wzoru gwarantującego spełnienie zależności:

– $w_{n+1} \geq w_n \geq \cdots \geq w_1$ oraz
– $w_{r(n+1)} + w_{rn} + \cdots + w_{r1} = 1$,

gdzie w_{ri} ($1 \leq i \leq +1$) oznacza urealnioną wagę i-tego faktu cząstkowego z konkluzją C, przy (n+1) wygenerowanych faktach cząstkowych z tą konkluzją (urealnienie zachodzi w następstwie wykonania funkcji f_{31}). Przykładową postać tego wzoru zaproponowała Szymkowiak w pracy [111]:

$$w_n = \begin{cases} 1, & dla\ n = 1 \\ \frac{w_{n-1}}{t + w_{n-1}}, & dla\ n \geq 2 \end{cases} \tag{5.15}$$

gdzie t oznacza stałą reprezentującą stosunek $\frac{w_{r(n+1)}}{w_{rn}}$ ($n \geq 1$) urealnionych wag dwóch kolejno uzyskanych faktów cząstkowych z tą samą konkluzją. Przykładowo, dla $t = 1.5$, kolejne wagi wyjściowe wynoszą: $w_1 = 1$, $w_2 = \frac{1}{2.5} = 0.4$, $w_3 = \frac{0.4}{1.9} \cong 0.21$, $w_4 \cong \frac{0.21}{1.71} \cong 0.12$, itd. Po wygenerowaniu czwartego faktu cząstkowego, następuje urealnienie wag do wartości: $w_{r4} \cong 0.12$, $w_{r3} \cong 0.21 \cdot$ $\cdot (1 - 0.12) \cong 0.18$, $w_{r2} \cong 0.4 \cdot (1 - 0.21) \cdot (1 - 0.12) \cong 0.28$, $w_{r2} \cong 1 \cdot$ $(1 - 0.4) \cdot (1 - 0.21) \cdot (1 - 0.12) \cong 0.42$, przy zachowanym stosunku $t \cong$ $\cong \frac{w_{r2}}{w_{r1}} \cong \frac{w_{r3}}{w_{r2}} \cong \frac{w_{r4}}{w_{r3}} \cong 1.5$.

Urealnione wagi faktów cząstkowych stają się wagami ich współczynników irf, a przez to wpływają na wartość współczynnika irf faktu zbiorczego. Wagi te nie mają natomiast wpływu na wartość współczynnika grf faktu zbiorczego.

Zgodnie z założeniem, wartość p_{ck} współczynnika irf faktu F_k powinna pochodzić z zakresu $\langle Bel_{ck}; Pls_{ck} \rangle$ takiego, że: $Bel_{ck} = \min\{Bel_{ci}, Bel_{cj}\}$

i $Pls_{ck} = \max\{Pls_{ci}, Pls_{cj}\}$, gdzie Bel_{ci}, Pls_{ci} oznaczają, odpowiednio globalny stopień przekonania i stopień możliwości współczynnika irf faktu F_i, a Bel_{cj}, Pls_{cj} – globalny stopień przekonania i stopień możliwości współczynnika irf faktu F_j. Z definicji funkcji f_{31} wynika, że $p_{ck} =$ $= (1 - w_{n+1}) \cdot p_{ci} + w_{n+1} \cdot p_{cj} \geq \min\{p_{ci}, p_{cj}\} \geq \min\{Bel_{ci}, Bel_{cj}\} = Bel_{ck}$ oraz $p_{ck} = (1 - w_{n+1}) \cdot p_{ci} + w_{n+1} \cdot p_{cj} \leq \max\{p_{ci}, p_{cj}\} \leq \max\{Pls_{ci}, Pls_{cj}\} =$ $= Pls_{ck}$. W takim razie, prawdziwa jest podwójna nierówność: $Bel_{ck} \leq p_{ck} \leq Pls_{ck}$.

Z kolei, wartość p_{rk} współczynnika grf faktu F_k musi pochodzić z zakresu $\langle Bel_{rk}; Pls_{rk} \rangle$ takiego, że: $Bel_{rk} = \max\{Bel_{ri}, Bel_{rj}\}$ i $Pls_{rk} =$ $= \max\{Pls_{ri}, Pls_{rj}\}$, gdzie Bel_{ri}, Pls_{ri} oznaczają, odpowiednio globalny stopień przekonania i stopień możliwości współczynnika grf faktu F_i, a Bel_{rj}, Pls_{rj} – globalny stopień przekonania i stopień możliwości współczynnika grf faktu F_j. Przy obowiązujących założeniach $Bel_{ri} \leq p_{ri} \leq Pls_{ri}$ oraz $Bel_{rj} \leq p_{rj} \leq Pls_{rj}$, można dowieść nierówności $Bel_{rk} \leq p_{rk} \leq Pls_{rk}$ w podobny sposób, jak to przeprowadzono dla funkcji f_{22} propagacji niepewności przesłanki reguły wewnątrz reguły.

Zastosowana w systemie RBS(2U) metoda obliczania współczynnika wiarygodności irf hipotezy wielokrotnej różni się zasadniczo od metody obliczania analogicznego współczynnika (CF) w systemie MYCIN. W systemie RBS(2U) waga poszczególnych hipotez wielokrotnych jest zróżnicowana i zależy od porządku generacji tych hipotez. Zakłada się przy tym, że hipoteza ma tym większą wartość, im wcześniej ją wygenerowano ($t \geq 1$). Z kolei, w systemie MYCIN wszystkie hipotezy wielokrotne mają równorzędne znaczenie ($t = 1$) i jednakowy wpływ na wartość końcową współczynnika pewności. Założenie poczynione dla systemu RBS(2U) bazuje na spostrzeżeniu, że silnik wnioskujący w pierwszej kolejności uaktywnia te reguły, które w największym stopniu zależą od faktów aksjomatycznych. Dalsze reguły mają charakter wtórny: zależą od wniosków pośrednich, wygenerowanych na wcześniejszych etapach procesu wnioskowania.

Metoda obliczania współczynnika wiarygodności irf w systemie RBS(2U) wykazuje natomiast duże podobieństwo do metody obliczania współczynnika częstotliwości f w nie-aksjomatycznym systemie wnioskowania NARS. Wartość współczynnika f hipotezy zbiorczej zależy od współczynników f jej hipotez wielokrotnych, lecz także od współczynników zaufania c tych hipotez: im większa jest wartość zaufania hipotezy wielokrotnej, tym większy jest wpływ częstotliwości tej hipotezy na częstotliwość hipotezy zbiorczej. O ile jednak w systemie NARS uwzględnia się w tych obliczeniach bezwzględne wartości współczynników zaufania poszczególnych hipotez wielokrotnych, o tyle w systemie RBS(2U) bierze się pod uwagę tylko uporządkowanie współczynników grf tych hipotez przez relację $\geq$. Powyższa różnica wynika z odmiennych interpretacji znaczenia współczynnika zaufania f w regule Wanga i współczynnika wiarygodności zewnętrznej grf w regule 2U. Pierwszy współczynnik ma interpretację częstotliwościową: jego wielkość można wyliczyć na podstawie samej wielkości ewidencji. Drugi współczynnik ma interpretację złożoną: przyjmuje się, że wiarygodność reguły zależy od wielkości ewidencji, z której tę regułę pozyskano, lecz także, między innymi, od stopnia homogeniczności tej ewidencji oraz od źródła jej pochodzenia. Wpływ wymienionych czynników jest trudny do zmierzenia, dlatego wartości grf nie da się obliczyć – można ją co najwyżej oszacować. Zakłada się przy tym, że błąd oszacowania nie ma wpływu na porządek $\geq$, w jakim pozostają wiarygodności poszczególnych reguł.

Z tego samego założenia wynika przyjęta w systemie RBS(2U) postać funkcji f_{32}. Ze wszystkich współczynników grf hipotez wielokrotnych, przy szacowaniu wielkości współczynnika grf hipotezy zbiorczej bierze się pod uwagę tylko jeden – najwyższy. W ten sposób hipoteza zbiorcza przechowuje pamięć o „najmocniejszym” wywodzie konkluzji. Pozostałe dopuszczone wywody, spełniające wymogi progów τ_1 i τ_2, wpłyną na wartość współczynnika irf tego faktu.

Propagacja niepewności hipotezy do hipotezy przeciwnej. Z definicji 5.3 wynika, że w regule 2U dopuszcza się użycie formuły atomowej z negacją (5.2 lub 5.4), zarówno po stronie przesłanki, jak i konkluzji. Z drugiej strony, konkluzje wszystkich faktów przechowywanych w bazie systemu RBS(2U) są zredagowane w formie prostej, czyli niezanegowanej. Z tego powodu, każda operacja na regule 2U zawierającej w przesłance formułę zanegowaną musi być poprzedzona „odwróceniem” tej formuły. Proces taki odbywa się w locie i nie powoduje żadnej zmiany w stanie bazy danych (nowo uzyskany fakt ma charakter wirtualny). Podobny proces zachodzi po uaktywnieniu reguły 2U z konkluzją w postaci formuły zanegowanej. Zanim zostanie podjęta właściwa obsługa faktu cząstkowego z tą konkluzją, wcześniej nastąpi jej odwrócenie. Aby określić stopień pewności hipotezy F_i': grf(p_{ri}'): $true \rightarrow \neg C$ with irf(p_{ci}') uzyskanej w procesie zmiany hipotezy F_i: grf(p_{ri}): $true \rightarrow C$ with irf(p_{ci}) na przeciwną, należy skorzystać z funkcji propagacji niepewności f_{41} i f_{42}, zdefiniowanych jak następuje:

$$f_{41}, f_{42}: [0;1] \times [0;1] \rightarrow [0;1] \;;$$

$$f_{41}(p_{ci}, p_{ri}) = 1 - p_{ci} \;;$$

$$f_{42}(p_{ci}, p_{ri}) = p_{ri} \;.$$

Z założenia wynika, że $Bel_{ci} \leq p_{ci} \leq Pls_{ci}$. W takim razie, ze względu na postać funkcji f_{41}, prawdziwa jest podwójna nierówność: $(1 - Pls_{ci}) \leq p_{ci}' \leq$ $\leq (1 - Bel_{ci})$ oraz, ze względu na zachodzenie oczywistych równości: $Bel_{ci}' = (1 - Pls_{ci})$ i $Pls_{ci}' = (1 - Bel_{ci})$, prawdziwa jest także podwójna nierówność: $Bel_{ci}' \leq p_{ci}' \leq Pls_{ci}'$. Z drugiej strony, ze względu na zachodzenie równości: $Bel_{ri}' = Bel_{ri}$ i $Pls_{ri}' = Pls_{ri}$, prawdziwe są nierówności: $Bel_{ri}' \leq p_{ri}' = p_{ri} \leq Pls_{ci}'$.

Należy w tym miejscu podkreślić, że funkcja propagacji niepewności f_{41} może być stosowana tylko w odniesieniu do faktów 2U. Gdyby dopuścić odwracanie dowolnych reguł 2U, doszłoby do naruszenia wymogu zależności dodatniej monotonicznej między przesłanką a konkluzją reguły. Z dwóch reguł: $\mathrm{grf}(p_{r1}): P \rightarrow C$ with $\mathrm{irf}(p_{c1})$ i $\mathrm{grf}(p_{r2}): P \rightarrow \neg C$ with $\mathrm{irf}(p_{c2})$, gdzie P oznacza przesłankę różną od true, przynajmniej jedna nie spełnia tego wymogu. W wypadku przesłanki równej true, reguła staje się faktem, a zależność między przesłanką a konkluzją nabiera charakteru trwałego: przesłanka true będzie zawsze spełniona w 100%.

5.4.3 Propagacja niepewności w procesie wnioskowania regresywnego

Choć wnioskowanie regresywne przebiega w systemie RBS(2U) według ogólnie przyjętych zasad (patrz podrozdz. 4.3.2), to używane w tym trybie funkcje propagacji niepewności mają charakter oryginalny. W dalszym ciągu podano definicje i krótkie opisy tych funkcji. Należy przypomnieć, że w procesie wnioskowania regresywnego nie powstają żadne nowe fakty rzeczywiste, czy wirtualne. Z tego powodu, nie trzeba tutaj dowodzić spełnienia jakichkolwiek ograniczeń, tak jak to miało miejsce w wypadku funkcji używanych w trybie wnioskowania progresywnego.

Wsteczna propagacja niepewności hipotezy – przez regułę – do nowej hipotezy. Podstawową czynnością wykonywaną w procesie wnioskowania wstecz jest reformułowanie hipotezy wyjściowej. Jeśli to postępowanie prowadzi ostatecznie do aksjomatu/aksjomatów znajdujących się w bazie danych, to następuje weryfikacja hipotezy, w przeciwnym razie – jej falsyfikacja. Jeśli na pewnym etapie tego procesu pytanie ma postać Q_i: $(\mathrm{grf}(p_{ri}): C \text{ with } \mathrm{irf}(p_{ci}) \leftarrow \mathrm{true}, op_i)$, a w bazie wiedzy znajduje się reguła R w postaci: $\mathrm{grf}(p_r): P_1, P_2, \dots, P_n \rightarrow C$ with $\mathrm{irf}(p_c)$ taka, że: $p_c \neq 0$, $p_c\ op_i\ p_{ci}$ i $p_r \geq p_{ri}$, to można to pytanie przekształcić do nowej postaci Q_j: $(\mathrm{grf}(p_{rj}): P_1, P_2, \dots, P_n \text{ with } \mathrm{irf}(p_{cj}) \leftarrow \mathrm{true}, op_j)$, gdzie op_j oznacza operator wyznaczony przy użyciu funkcji f_{50}, a p_{cj} i p_{rj} oznaczają współczynniki wiarygodności obliczone przy użyciu następujących funkcji, odpowiednio f_{51} i f_{52}:

$f_{50}: \{\leq, \geq\} \rightarrow \{\leq, \geq\}$;

$f_{51}, f_{52}: ([0;1] \times [0;1])^2 \rightarrow [0;1]$;

$f_{50}(\leq) = \leq$; $f_{50}(\geq) = \geq$;

$f_{51}(p_{ci}, p_{ri}, p_c, p_r) = \frac{p_{ci}}{p_c}$;

$f_{52}(p_{ci}, p_{ri}, p_c, p_r) = p_{ri}$.

Kierunek wnioskowania regresywnego, prowadzącego od hipotezy do aksjomatów, jest odwrotny do kierunku wnioskowania progresywnego. Stąd, funkcja f_{51} wykonuje działanie odwrotne do jej progresywnego odpowiednika f_{21}, a funkcja f_{52} – działanie odwrotne do jej odpowiednika f_{22}. Założenia odnośnie do postaci reguły R są konieczne dla zagwarantowania poprawności definicji f_{51} i f_{52}.

Wsteczna propagacja niepewności hipotezy złożonej do hipotez składowych. Reguła użyta na pewnym etapie weryfikacji hipotezy wejściowej może mieć przesłankę złożoną, sformułowaną w postaci koniunkcyjnej. Aby określić wymagany stopień pewności hipotez składowych, zawartych w pytaniach Q_i: $(\text{grf}(p_{ri}): P_i \leftarrow true \text{ with } \text{irf}(p_{ci}), op_i)$ $(1 \leq i \leq n)$, na podstawie znanego, wymaganego stopnia pewności pełnej hipotezy, zawartej w pytaniu Q_v: $(\text{grf}(p_{rv}): P_1 \land P_2 \land \ldots \land P_n \text{ with } \text{irf}(p_{cv}) \leftarrow true,\ op_v)$, należy zastosować poniższe funkcje f_{6i1} i f_{6i2} $(1 \leq i \leq n)$:

$$f_{60}: \{\leq, \geq\} \rightarrow \{\leq, \geq\}\ ;$$

$$f_{6i1}, f_{6i2}: [0; 1] \times [0; 1] \rightarrow [0; 1]\ ;$$

$$f_{60}(\leq) = \leq\ ;\ f_{60}(\geq) = \geq\ ;$$

$$f_{6i1}(p_{cv}, p_{rv}) = p_{cv}\ ;$$

$$f_{6i2}(p_{cv}, p_{rv}) = p_{rv}\ .$$

Z definicji tych wynika, że zarówno współczynniki irf, jak i współczynniki grf hipotez składowych, które wymagają weryfikacji dla udowodnienia pewnej hipotezy złożonej, są takie same jak w tej hipotezie złożonej. Tym samym, każda z funkcji f_{6i1} i f_{6i2} $(1 \leq i \leq n)$ jest w istocie funkcją jednoargumentową.

Wsteczna propagacja niepewności pytania do pytania przeciwnego. Podobnie jak w procesie wnioskowania progresywnego zachodzi potrzeba odwracania hipotez, tak podczas wnioskowania regresywnego zachodzi potrzeba odwracania pytań. Proces formułowania pytania Q': $(\text{grf}(p'_r): P' \leftarrow true \text{ with } \text{irf}(p'_c), op')$, przeciwnego do pytania Q: $(\text{grf}(p_r): P \leftarrow true \text{ with } \text{irf}(p_c), op)$, odbywa się przy użyciu następujących funkcji propagacji niepewności f_{7i} $(1 \leq i \leq n)$:

$$f_{70}: \{\leq, \geq\} \rightarrow \{\leq, \geq\}\ ;$$

$$f_{71}, f_{72}: [0; 1] \times [0; 1] \rightarrow [0; 1]\ ;$$

$$f_{70}(\leq) = \geq\ ;\ f_{60}(\geq) = \leq\ ;$$

$$f_{71}(p_c, p_r) = 1 - p_c\ ;$$

$$f_{72}(p_c, p_v) = p_r\ .$$

Spośród przytoczonych funkcji, szczególną uwagę zwraca f_{70}: jest to jedyna funkcja wstecznej propagacji niepewności, która dokonuje zmiany operatora relacyjnego na przeciwny. Podobnie jak poprzednio, funkcje f_{71} i f_{72} są w istocie funkcjami jednoargumentowymi.

5.4.4 Przykłady wnioskowania

Zamieszczone poniżej przykłady ilustrują sposób działania silnika wnioskującego w systemie RBS(2U). W wyliczeniach wartości współczynników irf i grf zastosowano zaokrąglenia arytmetyczne do dwóch miejsc po przecinku.

Przykład 5.1. W bazie wiedzy systemu RBS(2U) znajdują się następujące reguły R_1 – R_5, reprezentujące wiedzę na temat zależności występujących pomiędzy:
– wydziałem i rodzajem studiów studenta Politechniki Poznańskiej (PP) (Wydz_Rodz) a kierunkiem jego studiów na PP (Kier) – R_1, R_2;
– kierunkiem studiów, rokiem studiów (Rok) i płcią (Plec) studenta PP – a posiadanym przez niego sprzętem komputerowym (Sprzet) – R_3, R_4, R_5:

$$R_1\text{: grf(0.75): Wydz\_Rodz} = \{\text{WInf\_st, WEl\_st}\} \oplus \\ \rightarrow \text{Kier} = \{\text{Inf\_st}\} \odot \text{ with irf(0.55)}$$

$$R_2\text{: grf(0.9): Wydz\_Rodz} = \{\text{WInf\_st}\} \odot \\ \rightarrow \text{Kier} = \{\text{Inf\_st}\} \odot \text{ with irf(0.8)}$$

$$R_3\text{: grf(0.7): Kier} = \{\text{Inf\_st}\} \odot \\ \rightarrow \text{Sprzet} = \{\text{K\_st, Lap}\} \odot \text{ with irf(0.6)}$$

$$R_4\text{: grf(0.9): Kier} = \{\text{Inf\_st}\} \odot, \\ \text{Rok} = \neg\{1,2\} \oplus \\ \rightarrow \text{Sprzet} = \{\text{K\_st, Lap}\} \odot \text{ with irf(0.8)}$$

$$R_5\text{: grf(0.85): Kier} = \{\text{Inf\_st}\} \odot, \\ \text{Rok} = \neg\{1,2\} \oplus, \\ \text{Plec} = \neg\{\text{K}\} \odot, \\ \rightarrow \text{Sprzet} = \{\text{K\_st, Lap}\} \odot \text{ with irf(0.85)}$$

gdzie literał WInf_st oznacza Wydział_Informatyki_studia_stacjonarne; WEl_st – Wydział_Elektryczny_studia_stacjonarne; Inf_st – Informatykę_stacjonarną; K_st – komputer_stacjonarny; Lap – laptop; 1, 2 – odpowiednio, pierwszy i drugi rok studiów; K – kobietę. Użyte w regułach współczynniki wiarygodności mają jedynie charakter poglądowy.

Konstrukcję zbioru reguł $R_1 - R_5$ należy uznać za poprawną, ponieważ:
– nie istnieje zadana przesłanka świadcząca o nieprawidłowej budowie którejkolwiek z reguł $R_1 - R_5$,
– nie ma w zbiorze takiej pary reguł, które pozostawałyby we wzajemnej sprzeczności,

– nie ma w zbiorze reguły subsumowanej przez inną regułę z tego zbioru.

Jeśli w opisanej sytuacji, przy stanie bazy danych:

F_1^*: grf(1.0): true
→ Wydz_Rodz = {WInf_st} ⊙ with irf(1.0)

F_2^*: grf(0.95): true
→ Rok = {1,2} ⊕ with irf(0.0) ,

przy parametrach progowych $\tau_1 = 0.7$ i $\tau_2 = 0.7$ oraz stosunku wag współczynników $t = 1.8$ zostanie zainicjowane wnioskowanie w trybie progresywnym, to zawartość bazy danych będzie się zmieniać jak następuje:

Krok1:

F_1^*: grf(1.0): true
→ Wydz_Rodz = {WInf_st} ⊙ with irf(1.0)

F_2^*: grf(0.95): true
→ Rok = {1,2} ⊕ with irf(0.0)

F_3: grf(0.9): true
→ Kier = {Inf_st} ⊙ with irf(0.8)

Ten stan bazy danych jest wynikiem uaktywnienia reguły R_2, wybranej z dwóch zakwalifikowanych do agendy reguł R_1 i R_2. Obie reguły spełniały wymagania progowe τ_1 ($\min\{0.75, 1.0\} \geq 0.7$, $\min\{0.9, 1.0\} \geq 0.7$) i τ_2 ($1.0 \geq 0.7$). Wybór reguły R_2 wynikał z relacji pomiędzy współczynnikami grf obu reguł: $0.9 > 0.75$. Uaktywnienie reguły R_2 wiązało się z użyciem funkcji propagacji f_{21} i f_{22}. Wartość p_{c3} współczynnika irf nowo uzyskanego faktu F_3 została obliczona z zależności: $p_{c3} = 0.8 \cdot 1.0 = 0.8$, a wartość p_{r3} współczynnika grf tego faktu – z zależności: $p_{r3} = \min\{0.9, 1.0\} = 0.9$.

W zapisie faktów F_1^* i F_2^* użyto symbolu $*$ na oznaczenie tego, że są one aksjomatami i nie podlegają żadnym modyfikacjom w całym procesie wnioskowania. Odmienny charakter ma fakt F_3, który stanie się w Kroku3 podmiotem działania funkcji propagacji niepewności f_{31} i f_{32}.

Krok 2:

F_1^*: grf(1.0): true
→ Wydz_Rodz = {WInf_st} ⊙ with irf(1.0)

F_2^*: grf(0.95): true
→ Rok = {1,2} ⊕ with irf(0.0)

F_3: grf(0.9): true
→ Kier = {Inf_st} ⊙ with irf(0.8)

F_4: grf(0.9): true
→ Sprzet = {K_st, Lap} ⊙ with irf(0.64)

Nowo umieszczony w bazie danych fakt F_4 znalazł się tam w wyniku uaktywnienia reguły R_4, wybranej z agendy spośród reguł R_1 i R_4. Kwalifikacja R_4 do agendy odbyła się po uprzednim odwróceniu faktu F_2 do postaci F_2': $\text{grf}(0.95)\text{: true} \rightarrow \text{Rok} = \neg\{1,2\} \oplus \text{with irf}(1.0)$ (z zastosowaniem funkcji propagacji f_{41} i f_{42}), a następnie skonstruowaniu faktu wirtualnego F_v: $\text{grf}(0.9)\text{: true} \rightarrow (\text{Kier} = \{\text{Inf\_st}\} \odot \wedge \text{Rok} = \neg\{1,2\} \oplus) \text{with irf}(0.8)$ (z zastosowaniem funkcji propagacji f_{11} i f_{12}). Wartości p_{c4} i p_{r4} współczynników, odpowiednio irf i grf faktu F_4 wyznaczono ze wzorów: $p_{c4} = 0.8 \cdot 0.8 = 0.64$ i $p_{r4} = \min\{0.9, 0.9\} = 0.9$.

Krok3:

$F_1^*\text{: grf}(1.0)\text{: true}$
$\rightarrow \text{Wydz\_Rodz} = \{\text{WInf\_st}\} \odot \text{with irf}(1.0)$

$F_2^*\text{: grf}(0.95)\text{: true}$
$\rightarrow \text{Rok} = \{1,2\} \oplus \text{with irf}(0.0)$

$F_3\text{: grf}(0.9)\text{: true}$
$\rightarrow \text{Kier} = \{\text{Inf\_st}\} \odot \text{with irf}(0.71)$

$F_4\text{: grf}(0.9)\text{: true}$
$\rightarrow \text{Sprzet} = \{\text{K\_st}, \text{Lap}\} \odot \text{with irf}(0.6)$

W kolejnym kroku wnioskowania została uaktywniona reguła R_1. Reguła ta spowodowała wygenerowanie faktu cząstkowego F_{3_1}: $\text{grf}(0.75)\text{: true} \rightarrow \text{Kier} = \{\text{Inf\_st}\} \odot \text{with irf}(0.55)$, a następnie zastąpienie dotychczasowego faktu F_3 – nowym (zbiorczym) faktem F_3, w postaci: $\text{grf}(0.9)\text{: true} \rightarrow \text{Kier} = \{\text{Inf\_st}\} \odot \text{with irf}(0.71)$. Wartości p_{c3} i p_{r3} współczynników, odpowiednio irf i grf zmodyfikowanego faktu F_3 zostały obliczone przy użyciu funkcji propagacji f_{31} i f_{32}, jako: $p_{c3} = \left(1 - \frac{1}{1.8+1}\right) \cdot 0.8 + \frac{1}{1.8+1} \cdot 0.55 \cong 0.71$ oraz $p_{r3} = \max\{0.9, 0.75\} = 0.9$.

Krok 4:

$F_1^*\text{: grf}(1.0)\text{: true}$
$\rightarrow \text{Wydz\_Rodz} = \{\text{WInf\_st}\} \odot \text{with irf}(1.0)$

$F_2^*\text{: grf}(0.95)\text{: true}$
$\rightarrow \text{Rok} = \{1,2\} \oplus \text{with irf}(0.0)$

$F_3\text{: grf}(0.9)\text{: true}$
$\rightarrow \text{Kier} = \{\text{Inf\_st}\} \odot \text{with irf}(0.71)$

$F_4\text{: grf}(0.9)\text{: true}$
$\rightarrow \text{Sprzet} = \{\text{K\_st}, \text{Lap}\} \odot \text{with irf}(0.54)$

Jako ostatnia w procesie wnioskowania została uaktywniona reguła R_3. Także i ona spowodowała wygenerowanie faktu cząstkowego – F_{4_1}, w postaci: grf(0.7): true → Sprzet = {K_st, Lap} ⊙ with irf(0.43), a po nim nowego (zbiorczego) faktu F_4: grf(0.9): true → Sprzet = {K_st, Lap} ⊙ with irf(0.6), który zastąpił w bazie danych fakt F_4 z kroku 3. Współczynniki tego nowego faktu zostały wyznaczone w rezultacie obliczeń: $p_{c4} = \left(1 - \frac{1}{1.8+1}\right) \cdot 0.6 + + \frac{1}{1.8+1} \cdot 0.43 \cong 0.54$ oraz $p_{r4} = \max\{0.9, 0.7\} = 0.9$.

Wszystkie reguły kwalifikowane w poszczególnych krokach wnioskowania do agendy spełniały, co łatwo sprawdzić, wymagania progowe τ_1 i τ_2. Przy obowiązującym w systemie RBS(2U) założeniu otwartego świata OWA, reguła R_5 nie znalazła się w agendzie ze względu na brak w bazie danych faktu z konkluzją (Plec = {K} ⊙).

Przykład 5.2. Jeśli baza danych będzie miała identyczną zawartość początkową jak w przykładzie 5.1, to – przy tej samej zawartości bazy wiedzy (R_1 – R_5), tych samych wartościach progowych $\tau_1 = 0.7$ i $\tau_2 = 0.7$ oraz nowym, radykalnie zmniejszonym stosunku wag współczynników $t = 1.2$ – proces wnioskowania będzie miał następujący przebieg:

– Krok1 i Krok2 zakończą się identycznym rezultatem, jak w trakcie poprzedniego wnioskowania;
– Krok3 spowoduje zmianę stanu bazy danych na poniższy, różny od osiągniętego w trakcie poprzedniego wnioskowania:

  Krok3:

  F_1^*: grf(1.0): true
  → Wydz_Rodz = {WInf_st} ⊙ with irf(1.0)

  F_2^*: grf(0.95): true
  → Rok = {1,2} ⊕ with irf(0.0)

  F_3: grf(0.9): true
  → Kier = {Inf_st} ⊙ with irf(0.69)

  F_4: grf(0.9): true
  → Sprzet = {K_st, Lap} ⊙ with irf(0.6),

  gdzie wartości p_{c3} i p_{r3} współczynników, odpowiednio irf i grf zmodyfikowanego faktu F_3 wynikają ze wzorów: $p_{c3} = \left(1 - \frac{1}{1.2+1}\right) \cdot 0.8 + \frac{1}{1.2+1} \cdot 0.55 \cong \cong 0.69$ oraz $p_{r3} = \max\{0.9, 0.75\} = 0.9$;
– proces wnioskowania dobiegnie końca ze względu na brak reguł kwalifikujących się do agendy. W nowych warunkach, reguła R_3 nie spełnia wymogu $\tau_2 \geq 0.7$.

Pobieżna analiza odpowiedzi uzyskanych w procesach wnioskowania z przykładów 5.1 i 5.2 prowadzi do wniosku, że przebiegiem wnioskowania progresywnego w systemie RBS(2U) można sterować za pomocą każdego z parametrów τ_1, τ_2 i t. W szczególności, sama zmiana wartości t może wpływać na skrócenie/wydłużenie łańcuchów wnioskowań.

W ogólności, wraz z przebiegiem procesu wnioskowania progresywnego postępuje spadek wielkości współczynników (grf i irf) nowo generowanych faktów 2U. Spadek ten powoduje, że spełnienie wymagań progowych τ_1 i τ_2 staje się coraz trudniejsze, a w końcu niemożliwe, dając efekt w postaci pustej agendy. Z drugiej strony, przy dużych wartościach t, wpływ kolejno uzyskiwanych faktów cząstkowych na postać faktu zbiorczego może być tak nikły, że nie warto go brać pod uwagę. W takich okolicznościach można by rozważyć nowe kryterium zakończenia procesu wnioskowania progresywnego.

Przykład 5.3. Jeśli baza wiedzy będzie miała identyczną zawartość jak w przykładzie 5.1 (reguły $R_1 - R_5$), to przy nowej początkowej zawartości bazy danych:

$$F_1^*\!: \mathrm{grf}(1.0): \ \ \mathrm{true}$$
$$\rightarrow \mathrm{Wydz\_Rodz} = \{\mathrm{WEl\_st}\} \odot \text{ with } \mathrm{irf}(1.0)$$

$$F_2^*\!: \mathrm{grf}(0.95): \ \mathrm{true}$$
$$\rightarrow \mathrm{Rok} = \{1{,}2\} \oplus \text{ with } \mathrm{irf}(0.0)\ ,$$

oraz niezmienionych wartościach progowych $\tau_1 = 0.7$, $\tau_2 = 0.7$ i stosunku wag współczynników $t = 1.8$, proces wnioskowania progresywnego zakończy się na Kroku1, w którym zostanie wygenerowany nowy fakt F_3 w postaci:

$$F_3\!: \mathrm{grf}(0.75): \ \ \mathrm{true}$$
$$\rightarrow \mathrm{Kier} = \{\mathrm{Inf\_st}\} \odot \text{ with } \mathrm{irf}(0.55)$$

Powyższy fakt jest wynikiem uaktywnienia reguły R_1, jedynej znajdującej się w agendzie. W nowym stanie bazy danych żadna z reguł nie spełni wymogów kwalifikacji do agendy. Główną przeszkodę na tej drodze stanowi niska wiarygodność irf faktu F_3.

Podsumowując, o ile dla studenta Wydziału_Informatyki_studiów_stacjonarnych roku trzeciego lub wyższego dało się wygenerować hipotezę na temat posiadania przez niego komputera stacjonarnego i laptopa, o tyle nie można uzyskać analogicznej hipotezy dla studenta Wydziału_Elektrycznego_studiów_stacjonarnych. To zróżnicowanie jest w pełni uzasadnione: wiedza na temat posiadanego sprzętu komputerowego dotyczy tylko studentów kierunku Informatyka_stacjonarna (reguły $R_3 - R_5$), a ci są w mniejszości wśród studentów Wydziału_Elektrycznego_studiów_stacjonarnych (uśredniona wartość 0.55 współczynnika irf w regule R_1 odnosi się do studentów obu rozważanych wydziałów – Elektrycznego i Informatyki; to uśrednienie znajduje wyraz w stosunkowo niskiej wiarygodności grf tej reguły).

Chcąc koniecznie uzyskać hipotezę dotyczącą sprzętu komputerowego studenta Wydziału Elektrycznego, należy obniżyć wymagania progowe, np. do poziomów $\tau_1 = 0.6$, $\tau_2 = 0.5$. W takiej sytuacji, przy niezmienionym poziomie $t = 0.8$, system wygeneruje w odpowiedzi następujące fakty:

$$F_3\colon \mathrm{grf}(0.75)\colon\ \mathrm{true}$$
$$\rightarrow \mathrm{Kier} = \{\mathrm{Inf\_st}\} \odot \ \text{with irf}(0.55)$$
$$F_4\colon \mathrm{grf}(0.75)\colon\ \mathrm{true}$$
$$\rightarrow \mathrm{Sprzet} = \{\mathrm{K\_st}, \mathrm{Lap}\} \odot \ \text{with irf}(0.44)$$

Jak widać, jest między nimi żądana hipoteza (F_4); jej wiarygodność zewnętrzna ($\mathrm{p}_{r4} = 0.7$) jest jednak znacznie niższa od tej, z jaką wygenerowano analogiczną hipotezę dla studenta informatyki.

Przykład 5.4. Gdyby w bazie wiedzy systemu RBS(2U) umieścić reguły $R_1 - R_4$ w postaci identycznej jak w przykładzie 5.1 oraz regułę R_5 w nowej postaci:

$$\begin{aligned} R_5\colon \mathrm{grf}(0.85)\colon\ & \mathrm{Kier} = \{\mathrm{Inf\_st}\} \odot, \\ & \mathrm{Rok} = \neg\{1,2\} \oplus, \\ & \mathrm{Plec} = \neg\{\mathrm{K}\} \odot, \\ & \rightarrow \mathrm{Sprzet} = \{\mathrm{K\_st}, \mathrm{Lap}\} \odot \ \text{with irf}(0.75) \end{aligned}$$

ze zmienioną wartością współczynnika irf w stosunku do pierwowzoru, to wówczas jedną z reguł R_4 i R_5 należałoby uznać za zredagowaną niepoprawnie (niespełnienie warunku zależności dodatniej monotonicznej między przesłanką a konkluzją reguły).

Przykład 5.5. Gdyby w bazie wiedzy systemu RBS(2U) umieścić reguły $R_1 - R_5$ w postaci identycznej jak w przykładzie 5.1 i, dodatkowo, następującą regułę R_{3_1}:

$$R_{3\_1}\colon \mathrm{grf}(0.75)\colon\ \mathrm{Kier} = \{\mathrm{Inf\_st}\} \odot$$
$$\rightarrow \mathrm{Sprzet} = \{\mathrm{K\_st}\} \odot \ \text{with irf}(0.55)$$

to wówczas wystąpiłaby sprzeczność między regułami R_3 i R_{3_1} ($\mathrm{Sprzet} = \{\mathrm{K\_st}\} \odot \leq_{\mathrm{AS}} \mathrm{Sprzet} = \{\mathrm{K\_st}, \mathrm{Lap}\} \odot$ oraz $0.55 < 0.6$). Tę sprzeczność należałoby rozwiązać poprzez usunięcie z bazy przynajmniej jednej z tych reguł (np. R_3, ze względu na niższą wartość współczynnika grf).

Przykład 5.6. Gdyby w bazie wiedzy systemu RBS(2U) umieścić reguły $R_1 - R_4$ w postaci identycznej jak w przykładzie 5.1 oraz regułę R_5 w nowej postaci:

$$\begin{aligned} R_5\colon \mathrm{grf}(0.85)\colon\ & \mathrm{Kier} = \{\mathrm{Inf\_st}\} \odot, \\ & \mathrm{Rok} = \neg\{1,2\} \oplus, \\ & \mathrm{Plec} = \neg\{\mathrm{K}\} \odot, \\ & \rightarrow \mathrm{Sprzet} = \{\mathrm{K\_st}, \mathrm{Lap}\} \odot \ \text{with irf}(0.8) \end{aligned}$$

ze zmienioną wartością współczynnika irf w stosunku do pierwowzoru, to wówczas ta reguła, subsumowana w nowej sytuacji przez regułę R_4, stałaby się nadmiarowa.

Przykład 5.7. Jeśli w bazie wiedzy systemu RBS(2U) zostaną umieszczone reguły $R_1 - R_5$ w postaci identycznej jak w przykładzie 5.1, to przy nowej początkowej zawartości bazy danych:

$$F_1^*\colon \mathrm{grf}(1.0)\colon\ \mathrm{true} \\ \to \mathrm{Wydz\_Rodz} = \{\mathrm{WInf\_st}\} \odot \mathrm{with\ irf}(1.0)$$

$$F_2^*\colon \mathrm{grf}(0.95)\colon\ \mathrm{true} \\ \to \mathrm{Rok} = \{1{,}2\} \oplus \mathrm{with\ irf}(0.0)$$

$$F_5^*\colon \mathrm{grf}(1.0)\colon\ \mathrm{true} \\ \to \mathrm{Plec} = \{\mathrm{K}\} \odot \mathrm{with\ irf}(0.1)\ ,$$

oraz niezmienionych wartościach progowych $\tau_1 = 0.7$, $\tau_2 = 0.7$ i stosunku wag współczynników $t = 1.8$, proces wnioskowania będzie miał następujący przebieg:

– Krok1 i Krok2 zakończą się podobnym rezultatem, jak w trakcie wnioskowania z przykładu 5.1 (w bazie danych będzie się znajdował dodatkowy fakt F_5^*);
– w Kroku3 zostanie uaktywniona reguła R_5, która spowoduje wygenerowanie faktu cząstkowego F_{4_1}: $\mathrm{grf}(0.85)\colon \mathrm{true} \to \mathrm{Sprzet} = \{\mathrm{K\_st, Lap}\} \odot$ $\mathrm{with\ irf}(0.68)$, a następnie zastąpienie dotychczasowego faktu F_4 – nowym (zbiorczym) faktem F_4, w postaci: $\mathrm{grf}(0.9)\colon\ \mathrm{true} \to \mathrm{Sprzet} = \{\mathrm{K\_st, Lap}\} \odot$ $\mathrm{with\ irf}(0.65)$, w którym wartości p_{c4} i p_{r4} współczynników zmodyfikowanego faktu F_4 obliczono ze wzorów: $p_{c4} = \left(1 - \frac{1}{1.8+1}\right) \cdot 0.64 + \frac{1}{1.8+1} \cdot 0.68 \cong$ $\cong 0.65$ oraz $p_{r4} = \max\{0.9, 0.85\} = 0.9$;
– w Kroku4 zostanie uaktywniona reguła R_1; w wyniku tego uaktywnienia dotychczasowy fakt F_3 zostanie zastąpiony nowym, w postaci: $F_3\colon \mathrm{grf}(0.9)\colon \mathrm{true}\ \to \mathrm{Kier} = \{\mathrm{Inf\_st}\} \odot\ \mathrm{with\ irf}(0.71)$, tak jak to miało miejsce w Kroku3 wnioskowania z przykładu 5.1;
– w Kroku5 zostanie uaktywniona reguła R_3, która spowoduje wygenerowanie faktu cząstkowego F_{4_2}: $\mathrm{grf}(0.7)\colon\ \mathrm{true} \to \mathrm{Sprzet} = \{\mathrm{K\_st, Lap}\} \odot$ $\mathrm{with\ irf}(0.43)$, a następnie zastąpienie dotychczasowego faktu F_4 – nowym (zbiorczym) faktem F_4, w postaci: $\mathrm{grf}(0.9)\colon\ \mathrm{true} \to \mathrm{Sprzet} = \{\mathrm{K\_st, Lap}\} \odot$ $\mathrm{with\ irf}(0.57)$, w którym wartości p_{c4} i p_{r4} współczynników zmodyfikowanego faktu F_4 obliczono ze wzorów: $p_{c4} = \left(1 - \frac{\frac{1}{1.8+1}}{\frac{1}{1.8+1}+1.8}\right) \cdot 0.65 +$ $+ \frac{\frac{1}{1.8+1}}{1.8+\frac{1}{1.8+1}} \cdot 0.43 \cong 0.61$ oraz $p_{r4} = \max\{0.9, 0.7\} = 0.9$.

Reasumując, po zakończeniu procesu wnioskowania baza danych będzie miała następującą zawartość:

F_1^*: grf(1.0): true
→ Wydz_Rodz = {WInf_st} ⊙ with irf(1.0)

F_2^*: grf(0.95): true
→ Rok = {1,2} ⊕ with irf(0.0)

F_5^*: grf(1.0): true
→ Plec = {K} ⊙ with irf(0.1) ,

F_3: grf(0.9): true
→ Kier = {Inf_st} ⊙ with irf(0.71)

F_4: grf(0.9): true
→ Sprzet = {K_st, Lap} ⊙ with irf(0.61)

Znajdujący się w tej bazie fakt zbiorczy F_4 jest wynikiem trzykrotnej generacji konkluzji Sprzet = {K_st, Lap} ⊙, ze współczynnikami grf, kolejno: 0.9, 0.85 i 0.7, oraz współczynnikami irf, kolejno: 0.64, 0.68 i 0.43. Wpływ trzech ostatnich wartości na wartość końcową współczynnika irf był zróżnicowany i zależny od wag konkluzji cząstkowych, te zaś – od wartości stałej proporcji $t = 1.8$. Poszczególne wagi wyniosły:

$$w_{r1} = 1 \cdot \left(1 - \frac{1}{1.8+1}\right) \cdot \left(1 - \frac{\frac{1}{1.8+1}}{1.8 + \frac{1}{1.8+1}}\right) = \frac{324}{604} \cong 0.536;$$

$$w_{r2} = \frac{1}{1.8+1} \cdot \left(1 - \frac{\frac{1}{1.8+1}}{1.8 + \frac{1}{1.8+1}}\right) = \frac{180}{604} \cong 0.298;$$

$$w_{r3} = \frac{\frac{1}{1.8+1}}{1.8 + \frac{1}{1.8+1}} = \frac{100}{604} \cong 0.166.$$

Oczywiście, zachodzą zależności: $\frac{w_{r2}}{w_{r1}} = \frac{w_{r3}}{w_{r2}} = 1.8$ oraz $w_{r1} + w_{r2} + w_{r3} = 1$.

Przykład 5.8. Jeśli w bazie wiedzy systemu RBS(2U) zostaną umieszczone reguły $R_1 - R_5$ w postaci identycznej jak w przykładzie 5.1 i, dodatkowo, reguła R_6, w postaci:

R_6: grf(0.7): Wydz_Rodz = {WInf_st} ⊙
→ Sprzet = {Lap} ⊙ with irf(0.5),

to przy tej samej zawartości początkowej bazy danych (fakty F_1^* i F_2^*), niezmienionych wartościach progowych $\tau_1 = 0.7$, $\tau_2 = 0.7$ i stosunku wag współczynników $t = 1.8$, proces wnioskowania będzie miał następujący przebieg:

– Krok1, Krok2, Krok3 i Krok4 zakończą się identycznym rezultatem, jak w trakcie wnioskowania z przykładu 5.1;

– proces wnioskowania będzie kontynuowany, i w Kroku5 zostanie uaktywniona reguła R_6; w wyniku tego uaktywnienia zostanie wygenerowany nowy fakt F_5, w postaci:

$$F_5\text{: grf(0.7): true} \\ \rightarrow \text{Sprzet} = \{\text{Lap}\} \odot \text{ with irf(0.5)}\ .$$

który będzie pozostawał w sprzeczności ze znajdującym się tam faktem F_4:

$$F_4\text{: grf(0.9): true} \\ \rightarrow \text{Sprzet} = \{\text{Komp\_st, Lap}\} \odot \text{ with irf(0.54)}\ .$$

Powyższa sprzeczność powinna zostać wykryta i wyeliminowana przez podsystem utrzymywania wiarygodności. Można to uczynić, między innymi, poprzez usunięcie z bazy danych faktu F_5, jako tego, który cechuje się mniejszą wiarygodnością (0.7 < 0.9).

Przykład 5.9. W bazie wiedzy systemu RBS(2U) znajdują się reguły $R_1 - R_5$, a w bazie danych – fakty F_1^*, F_2^* i F_5^* w postaci identycznej jak w przykładzie 5.7. W opisanej sytuacji, uruchomienie procesu wnioskowania regresywnego dla pytania Q_1, w postaci:

$$Q_1\text{: (grf(0.85): Sprzet} = \{\text{K\_st, Lap}\} \odot \text{ with irf(0.75)} \leftarrow \text{ true, } \geq)$$

spowoduje wykonanie następujących czynności:

Krok1:

$$Q_2\text{: (grf(0.85): (Kier} = \{\text{Inf\_st}\} \odot), \\ (\text{Rok} = \neg\{1,2\} \oplus) \text{ with irf(0.94)} \leftarrow \text{ true, } \geq)$$

Poszukiwanie w bazie danych faktu z konkluzją nie słabszą niż konkluzja $\text{Sprzet} = \{\text{Komp\_st, Lap}\} \odot$ kończy się porażką. W tej sytuacji jest inicjowany proces przeszukiwania bazy wiedzy pod kątem obecności w niej reguł z wymaganą konkluzją. Pierwszą taką regułą jest R_3:

$$R_3\text{: grf(0.7): Kier} = \{\text{Inf\_st}\} \odot \\ \rightarrow \text{Sprzet} = \{\text{K\_st, Lap}\} \odot \text{ with irf(0.6).}$$

Niestety, jej współczynniki wiarygodności nie spełniają obligatoryjnego warunku: $(0.6 \neq 0) \wedge (0.6 \geq 0.75) \wedge (0.7 \geq 0.85)$ (patrz podrozdz. 5.4.3). Kolejna reguła z wymaganą konkluzją to R_4:

$$R_4\text{: grf(0.9): Kier} = \{\text{Inf\_st}\} \odot\ , \\ \text{Rok} = \neg\{1,2\} \oplus \\ \rightarrow \text{Sprzet} = \{\text{K\_st, Lap}\} \odot \text{ with irf(0.8)}$$

W tym wypadku warunek dotyczący współczynników wiarygodności reguły $((0.8 \neq 0) \wedge (0.8 \geq 0.75) \wedge (0.9 \geq 0.85))$ jest spełniony, więc następuje generacja pytania pomocniczego Q_2, z hipotezą opartą na przesłance tej reguły.

Generacja pytania Q_2 wiąże się z użyciem funkcji wstecznej propagacji niepewności f_{50}, f_{51} i f_{52}, odpowiadających za:
– ustawienie operatora pytania na $\geq$ – operator identyczny jak w pytaniu Q_1;
– obliczenie wartości p_{c2} współczynnika irf ze wzoru: $p_{c2} = \frac{0.75}{0.8} \cong 0.94$,
– ustawienie wartości p_{r2} współczynnika grf na: $p_{r2} = 0.85$.

Krok2:

$$\begin{Bmatrix} Q_{3\_1}: (\text{grf}(0.85)\text{: Kier} = \{\text{Inf\_st}\} \odot \text{ with irf}(0.94) \leftarrow \text{true}, \geq), \\ Q_{3\_2}: (\text{grf}(0.85)\text{: Rok} = \neg\{1,2\} \oplus \text{ with irf}(0.94) \leftarrow \text{true}, \geq) \end{Bmatrix}$$

Pytanie Q_2 jest transformowane do zbioru pytań pomocniczych $\{Q_{3\_1}, Q_{3\_2}\}$, z których każde reprezentuje jedną hipotezę składową hipotezy złożonej pochodzącej z Q_2. Niniejsza transformacja jest wykonywana przy wykorzystaniu funkcji wstecznej propagacji niepewności f_{6ij} $(0 \leq i, j \leq 2)$, odpowiadających za:
– f_{601}, f_{02} – ustawienie operatora pytań Q_{3_1} i Q_{3_2} na $\geq$ – operator identyczny jak w pytaniu Q_2;
– f_{611}, f_{12} – ustawienie wartości p_{c3_1} i p_{r3_1} współczynników irf i grf pytania Q_{3_1} na identyczne jak w pytaniu Q_2;
– f_{621}, f_{622} – ustawienie wartości p_{c3_2} i p_{r3_2} współczynników irf i grf pytania Q_{3_2} na identyczne jak w pytaniu Q_2.

Krok3:

$$\begin{aligned} Q_4\text{: (grf}(0.85)\text{: } & (\text{Kier} = \{\text{Inf\_st}\} \odot), \\ & (\text{Rok} = \neg\{1,2\} \oplus), \\ & (\text{Plec} = \neg\{\text{K}\} \odot) \text{ with irf}(0.88) \leftarrow \text{true}, \geq) \end{aligned}$$

Poszukiwanie w bazie danych faktu z konkluzją nie słabszą niż Kier = {Inf_st} ⊙ kończy się porażką. Niestety, porażką kończy się także poszukiwanie w bazie wiedzy odpowiedniej reguły z taką konkluzją (żadna z reguł R_1 i R_2 nie spełnia warunku dotyczącego współczynników wiarygodności). W tej sytuacji następuje w prowadzonym procesie abdukcyjnym nawrót do najbliższego punktu wnioskowania, w którym można dokonać wyboru alternatywnej reguły. Jest nim punkt wcześniejszego wyboru reguły R_4. Zamiast niej, silnik wnioskujący rozpatruje w tym kroku regułę R_5.

$$\begin{aligned} R_5\text{: grf}(0.85)\text{: } & \text{Kier} = \{\text{Inf\_st}\} \odot, \\ & \text{Rok} = \neg\{1,2\} \oplus, \\ & \text{Plec} = \neg\{\text{K}\} \odot, \\ \rightarrow & \text{ Sprzet} = \{\text{K\_st}, \text{Lap}\} \odot \text{ with irf}(0.85) \end{aligned}$$

Niniejsza reguła spełnia warunek dotyczący współczynników wiarygodności, więc ma miejsce generacja nowego pytania pomocniczego Q_4. Procedura wyznaczania operatora pytania Q_4 i współczynników wiarygodności irf i grf jego hipotezy ma podobny przebieg jak w Kroku1.

Krok4

$$\begin{cases} Q_{5\_1}\text{: (grf(0.85): Kier} = \{\text{Inf\_st}\} \odot \text{ with irf(0.88)} \leftarrow \text{true}, \geq), \\ Q_{5\_2}\text{: (grf(0.85): Rok} = \neg\{1,2\} \oplus \text{ with irf(0.88)} \leftarrow \text{true}, \geq), \\ Q_{5\_3}\text{: (grf(0.85): Plec} = \neg\{\text{K}\} \odot \text{ with irf(0.88)} \leftarrow \text{true}, \geq) \end{cases}$$

Ze względu na złożoną postać hipotezy zawartej w pytanie Q_4, następuje w tym punkcie transformacja Q_4 do zbioru pytań pomocniczych $\{Q_{5\_1}, Q_{5\_2}, Q_{5\_3}\}$. Ma ona analogiczny przebieg jak w Kroku2.

Krok5

Niestety, zarówno poszukiwanie w bazie danych faktu z konkluzją nie słabszą niż Kier = {Inf_st} ⊙, jak i poszukiwanie w bazie wiedzy odpowiedniej reguły z taką konkluzją kończy się porażką. W tej sytuacji, przy braku możliwości wykonania kolejnego nawrotu, proces wnioskowania kończy się udzieleniem odpowiedzi false na pytanie wyjściowe Q_1.

Przykład 5.10. W bazie wiedzy systemu RBS(2U) znajdują się następujące reguły $R_1 - R_5$, w których zaszyta jest wiedza na temat zależności pomiędzy:

– wynikami osiągniętymi przez studenta na ćwiczeń laboratoryjnych (Lab_SI), kolokwium (Kol_SI) i teście egzaminacyjnym (Test_SI) ze sztucznej inteligencji (SI) oraz
– poziomem obecności i aktywności studenta na wykładach z SI (Wykl_SI) a poziomem jego wiedzy z SI, wyrażonym faktem zdania/ niezdania egzaminu z SI (Egz_SI):

R_1: grf(1.0): Test_SI = {z} ⊙
→ Egz_SI = {+} ⊙ with irf(0.75)

R_2: grf(0.9): Egz_SI = {+} ⊙,
Lab_SI = ¬{z} ⊙
→ Ocena_posr = {sciaga} ⊙ with irf(0.9)

R_3: grf(0.7): Ocena_posr = {sciaga} ⊙,
Wykl_SI = ¬{obecn} ⊙,
→ Egz_SI = ¬{+} ⊙ with irf(0.95)

R_4: grf(0.9): Egz_SI = ¬{+} ⊙,
Wykl_SI = {aktywn} ⊙
→ Ocena_posr = {watpl} ⊙ with irf(1.0)

R_5: grf(0.7): Ocena_posr = {watpl} ⊙,
Kol_SI = ¬{z} ⊙,
→ Egz_SI = ¬{+} ⊙ with irf(0.8)

gdzie literał z oznacza fakt zaliczenia kolokwium/ ćwiczeń laboratoryjnych/ testu egzaminacyjnego, + – fakt zdania egzaminu, sciaga – przekonanie egzaminatora o użyciu przez studenta ściągi na teście egzaminacyjnym, watpl – wątpliwość egzaminatora odnośnie do poziomu wiedzy studenta, obecn – częstą obecność na wykładach z SI, aktywn – dużą aktywność na wykładach z SI.

Jeśli w powyższym systemie, przy stanie bazy danych:

F_1^*: grf(1.0): true
→ Test_SI = {z} ⊙ with irf(1.00)

F_2^*: grf(1.0): true
→ Kol_SI = {z} ⊙ with irf(0.4)

F_3^*: grf(1.0): true
→ Lab_SI = {z} ⊙ with irf(0.1)

F_4^*: grf(1.0): true
→ Wykl_SI = {obecn} ⊙ with irf(0.2)

F_5^*: grf(1.0): true
→ Wykl_SI = {aktywn} ⊙ with irf(0.6)

parametrach progowych $\tau_1 = 0.65$ i $\tau_2 = 0.5$ oraz stosunku wag współczynników $t = 0.5$ zostanie zainicjowane wnioskowanie w trybie progresywnym, to w kolejnych krokach wnioskowania baza danych będzie zawierać następujące fakty prócz $F_1^* - F_5^*$:

Krok1

F_6: grf(1.0): true
→ Egz_SI = {+} ⊙ with irf(0.75)

(nowy fakt F_6 – wynik uaktywnienia reguły R_1 i zastosowania funkcji propagacji f_{21} i f_{22})

Krok2

F_6: grf(1.0): true
→ Egz_SI = {+} ⊙ with irf(0.75)

F_7: grf(0.9): true
→ Ocena_posr = {sciaga} ⊙ with irf(0.68)

(nowy fakt F_7 – wynik zastosowania odpowiednich funkcji propagacji i uaktywnienia reguły R_2)

Krok3

F_6: grf(1.0): true
→ Egz_SI = {+} ⊙ with irf(0.48)

F_7: grf(0.9): true

→ Ocena_posr = {sciaga} $\odot$ with irf(0.68)

(nowa postać faktu F_6 – wynik zastosowania odpowiednich funkcji propagacji, uaktywnienia reguły R_3 oraz obliczenia współczynników hipotezy zbiorczej Egz_SI = {+} $\odot$; ze względu na obniżenie wartości współczynnika irf hipotezy poniżej wymaganego progu $\tau_2 = 0.5$, reguła R_2 nie uzyska w kolejnym kroku kwalifikacji do agendy)

Krok4

F_6: grf(1.0): true

→ Egz_SI = {+} $\odot$ with irf(0.48)

F_7: grf(0.9): true

→ Ocena_posr = {sciaga} $\odot$ with irf(0.68)

F_8: grf(0.9): true

→ Ocena_posr = {watpl} $\odot$ with irf(0.52)

(nowy fakt F_8 – wynik zastosowania odpowiednich funkcji propagacji i uaktywnienia reguły R_4)

Krok5

F_6: grf(1.0): true

→ Egz_SI = {+} $\odot$ with irf(0.54)

F_7: grf(0.9): true

→ Ocena_posr = {sciaga} $\odot$ with irf(0.68)

F_8: grf(0.9): true

→ Ocena_posr = {watpl} $\odot$ with irf(0.52)

(nowa postać faktu F_6 – wynik zastosowania odpowiednich funkcji propagacji, uaktywnienia reguły R_5 oraz obliczenia współczynników hipotezy zbiorczej Egz_SI = {+} $\odot$; ze względu na podwyższenie wartości współczynnika irf hipotezy powyżej wymaganego progu $\tau_2 = 0.5$, reguła R_2 uzyska w kolejnym kroku kwalifikację do agendy)

Krok6

F_6: grf(1.0): true

→ Egz_SI = {+} $\odot$ with irf(0.54)

F_7: grf(0.9): true

→ Ocena_posr = {sciaga} $\odot$ with irf(0.55)

F_8: grf(0.9): true

→ Ocena_posr = {watpl} $\odot$ with irf(0.52)

(nowa postać faktu F_7 – wynik zastosowania odpowiednich funkcji propagacji, uaktywnienia reguły R_2 oraz obliczenia współczynników hipotezy zbiorczej Ocena_posr = {sciaga} $\odot$)

Powyższy przykład ma jedynie charakter ilustracyjny. Warto jednak odnotować specyficzny charakter użytej w nim bazy wiedzy oraz specyficzną wartość stosunku wag współczynników $t = 0.5$ zastosowanego we wnioskowaniu. Specyfika bazy wiedzy polega na występowaniu cykli: formuła atomowa Egz_SI $= \{+\} \odot$ pełni w regułach zarówno rolę przesłanki (R_2 i R_4), jak i konkluzji (R_1, R_3 i R_5). Taka baza wiedzy może być wykorzystana do programowania działań układów ze sprzężeniem zwrotnym. Z kolei, stosunek $t < 1$ oznacza, że waga generowanych hipotez wielokrotnych będzie rosła wraz z postępem procesu wnioskowania. Przy t dążącym do 0, $t \to 0$, bieżąco wygenerowana hipoteza wielokrotna stanie się hipotezą zbiorczą. Taka wartość stosunku t może być przydatna w systemach RBS(2U) przeznaczonych do sterowania procesami dynamicznymi.