

Politechnika Opolska
Wydział Elektrotechniki, Automatyki i Informatyki
Instytut Automatyki i Informatyki



Rozprawa doktorska pt.:

**KONCEPCJA I IMPLEMENTACJA SYSTEMU
WNIOSKUJĄCEGO Z PROBABILISTYCZNO-ROZMYTA
BAZĄ WIEDZY**

Autor: mgr inż. Katarzyna Rudnik

Promotor: dr hab. inż. A. Walaszek-Babiszewska, prof. PO

Opole, 2011

Składam serdecne podziękowanie

Pani dr hab. Annie Włodarczyk-Babiszczek, prof. PŁ

za nietypowe ukrumakowanie pracy,

co do poświęcenia przy jej realizacji, dwie zielone

SPIS TREŚCI

I. OPIS PROBLEMU BADAWCZEGO.....	5
1. Motywacja dysertacji	5
2. Umiejscowienie tematu dysertacji w dyscyplinie Automatyka i Robotyka.....	6
3. Podstawowe pojęcia używane w dysertacji	7
4. Cel i teza pracy.....	8
5. Zakres pracy	8
II. PODSTAWY TEORETYCZNE BUDOWY SYSTEMU WNIOSKUJĄCEGO Z PROBABILISTYCZNO-ROZMYTĄ BAZĄ WIEDZY.....	10
1. Identyfikacja modeli w systemach sterowania opartych o modele i w systemach opartych o rozmyte bazy wiedzy	10
2. Wybrane elementy teorii zbiorów rozmytych.....	12
2.1. Definicja i własności zbiorów rozmytych	12
2.2. Operacje mnogościowe na zbiorach rozmytych	18
3. Wybrane elementy teorii prawdopodobieństwa	23
3.1. Przestrzeń zdarzeń elementarnych, zdarzenia, prawdopodobieństwo w ujęciu klasycznym	24
3.2. Prawdopodobieństwo zdarzeń rozmytych	26
4. Podstawy teorii rozmytych systemów wnioskujących.....	27
4.1. Struktura rozmytych systemów wnioskujących	27
4.2. Regułowa reprezentacja wiedzy – rozmyte modele wiedzy.....	28
4.2.1. Rozmyte modele lingwistyczne	29
4.2.2. Modele Takagi – Sugeno – Kanga	31
4.2.3. Modele probabilistyczno-rozmyte	31
4.2.4. Modele relacyjne	33
4.2.5. Modele neuronowo-rozmyte.....	34
4.3. Właściwości regułowej, rozmytej bazy wiedzy	35
4.4. Procedury wnioskowania rozmytego na podstawie bazy wiedzy	37
4.5. Rozmywanie i wyostrzanie.....	41
5. Metody automatycznego pozyskiwania baz wiedzy w systemach rozmytych.....	43
5.1. Metody pozyskiwania baz wiedzy dla modeli lingwistycznych	43
5.1.1. Metoda Wanga-Mendela	43

5.1.2.	Metoda Nozaki-Ishibuchi-Tanaki.....	45
5.1.3.	Metoda Sugeno-Yasukawy.....	45
5.1.4.	Metoda szablonowego modelowania systemów rozmytych	46
5.2.	Metody eksploracji danych i maszynowego uczenia	48
5.2.1.	Metody grupowania	49
5.2.2.	Metody rozmytych drzew decyzyjnych.....	50
5.2.3.	Rozmyte metody hybrydowe	51
6.	Reguły asocjacji	52
6.1.	Istota, rodzaje i zastosowania reguł asocjacji	52
6.2.	Wyszukiwanie ilościowych reguł asocjacji	53
6.2.1.	Algorytm Apriori.....	54
6.2.2.	Algorytm FP-Growth.....	55
6.3.	Wyszukiwanie rozmytych reguł asocjacji	56
7.	Wnioski z analizy literaturowej.....	57
III.	KONCEPCJA I IMPLEMENTACJA SYSTEMU WNIOSKUJĄCEGO Z PROBABILISTYCZNO-ROZMYTĄ BAZĄ WIEDZY	59
1.	Ogólny zarys systemu.....	59
2.	Wybór sposobu rozmywania wartości zmiennych	61
3.	Opracowanie algorytmu generowania reguł probabilistyczno- rozmytej bazy wiedzy	66
3.1.	Algorytm generujący pełny rozkład prawdopodobieństwa zdarzeń rozmytych	67
3.2.	Algorytmy oparte na regułach asocjacji	76
3.2.1.	Budowa bazy reguł z uwzględnieniem algorytmu wyszukującego ilościowe reguły asocjacji	76
3.2.2.	Budowa bazy reguł z uwzględnieniem zmodyfikowanego algorytmu Apriori...	83
3.2.3.	Budowa bazy reguł z uwzględnieniem zmodyfikowanego algorytmu FP-Growth	88
3.3.	Porównanie algorytmów generowania bazy reguł.....	94
3.3.1.	Analiza czasu generowania i rozmiarów bazy reguł z pełnym rozkładem prawdopodobieństwa	94
3.3.2.	Analiza czasu generowania i rozmiarów bazy reguł z niepełnym rozkładem prawdopodobieństwa	96
3.4.	Wnioski.....	100

4. Metoda identyfikacji modelu wiedzy w systemie z probabilistyczno-rozmytą bazą reguł	102
5. Wnioskowanie w oparciu o probabilistyczno-rozmytą bazę wiedzy	103
5.1. Wyprowadzenie stopnia aktywacji reguł plikowych.....	104
5.2. Wyprowadzenie rozmytego wniosku reguł elementarnych.....	105
5.3. Wyprowadzenie rozmytego wniosku reguł plikowych	107
5.4. Wyprowadzenie rozmytego wniosku bazy wiedzy	108
5.5. Wyznaczenie wartości numerycznej na wyjściu systemu	109
6. Implementacja systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy w środowisku Matlab	111
6.1. Założenia oraz wymaganie funkcjonalne implementacji modułu narzędziowego PFIS	111
6.2. Budowa omawianego systemu w kodzie programu Matlab.....	112
6.2.1. Tworzenie nowego systemu jako obiektu struktury	112
6.2.2. Generowanie probabilistyczno-rozmytej bazy reguł.....	114
6.2.3. Wnioskowanie w oparciu o utworzony model wiedzy.....	114
6.3. Graficzny interfejs użytkownika - PFISEDIT	115
IV. ZASTOSOWANIA SYSTEMU WNIOSKUJĄCEGO Z PROBABILISTYCZNO-ROZMYTĄ BAZĄ WIEDZY	121
1. Modelowanie własności węgla	121
1.1. Wielkości charakteryzujące węgiel jako materiał uziarniony	121
1.2. Reprezentacja probabilistyczno-rozmyta dwóch zmiennych charakteryzujących węgiel dla danych surowych.....	122
1.3. Reprezentacja probabilistyczno-rozmyta dwóch zmiennych charakteryzujących węgiel dla danych uśrednionych.....	125
1.4. Reprezentacja probabilistyczno-rozmyta dla filtracji danych dynamicznych.....	128
1.4.1. Filtracja sygnałów	128
1.4.2. Analiza właściwości systemu wnioskującego z probabilistyczno- rozmytą bazą wiedzy dla filtracji danych dynamicznych	129
1.5. System wnioskujący z probabilistyczno-rozmytą bazą wiedzy dla pełnej charakterystyki węgla	135
1.5.1. Nawiązanie do idei <i>fuzzy graph</i>	135
1.5.2. Dobór zmiennych lingwistycznych i ich wartości.....	136
1.5.3. Identyfikacja bazy wiedzy dla pełnej, statycznej charakterystyki węgla	137
2. System decydujący o wyborze algorytmu do budowy probabilistyczno-rozmytej bazy wiedzy	143

2.1.1.	Dobór zmiennych lingwistycznych i ich wartości.....	143
2.1.2.	Identyfikacja probabilistyczno-rozmytej bazy wiedzy dla systemu decyzyjnego	145
2.1.3.	Analiza probabilistyczno-rozmytej bazy wiedzy przy faktach wyrażonych za pomocą wartości lingwistycznych i umerycznych	150
V. PODSUMOWANIE I WNIOSKI KOŃCOWE		153
LITERATURA		156
SPIS RYSUNKÓW		166
SPIS TABEL.....		171
DODATEK A – WYNIKI PORÓWNANIA ALGORYTMÓW PRZY RÓŻNYCH ZAŁOŻENIACH WEJŚCIOWYCH		173
DODATEK B – OPIS FUNKCJI Z MODUŁU NARZĘDZIOWEGO PFIS		176
DODATEK C – PROBABILISTYCZNO-ROZMYTA BAZA WIEDZY DLA MODELU WIELOWYMIAROWEJ CHARAKTERYSTYKI STATYCZNEJ		180
DODATEK E – WYNIKI ANALIZY SYSTEMU DECYZYJNEGO Z WIĘKSZĄ ILOŚCIĄ ZBIORÓW ROZMYTYCH.....		185

I. Opis problemu badawczego

1. Motywacja dysertacji

Podeczas rozwiązywania problemów dotyczących identyfikacji, sterowania procesami, prognozowania, diagnozowania czy podejmowania decyzji, mamy do czynienia z wieloma zadaniami, które wiążą się z ograniczoną wiedzą i niepewnością dotyczącą przebiegu zdarzeń oraz działania modelowanych obiektów. Wynika to z tempa i zakresu zmian w pozornie lokalnym otoczeniu danych problemów, a także jest efektem zjawisk naturalnych, które wciąż zostają nieposkromione dla umysłów badaczy. Analizując źródła niepewności możemy wyróżnić niepewność informacji wynikającą z niewiedzy ludzkiej, zwanej niepewnością subiektywną, a także niepewność obiektywną, która wynika z charakterystyki analizowanych procesów [bub05]. Aby móc odkryć i usystematyzować wiedzę obarczoną wymienionymi zagadnieniami, w literaturze spotyka się różną reprezentację wiedzy niepewnej. Można wyróżnić relacyjną reprezentację wiedzy [ped84], czy opis zmiennych niepewnych (ang. *uncertain variable*) [bub05] wyrażonych poprzez zbiór wartości i funkcję stanowiącą wskaźnik pewności. Jednakże najczęściej, modelowanie niepewności dokonywane jest z użyciem opisu zmiennych rozmytych opartej na tzw. logice rozmytej [zad65], [zad68], [zad73], [zad75], [zad79]. W tym nurcie można znaleźć wiele prac poświęconych tworzeniu systemów z bazą wiedzy (ang. *knowledge-based systems*) [zad73]. Stanowią one alternatywę dla systemów opartych o modele i tradycyjne algorytmy numeryczne w sytuacjach, gdy informacja o danej dziedzinie jest niepewna, niejednoznacznie sformalizowana, bądź też zbyt trudna lub kosztowna do uzyskania [mam75], [tak85], [yag95], [wan92]. Od kilkudziesięciu lat rozmyte systemy z bazą wiedzy znajdują zastosowanie w zagadnieniach automatyki i robotyki dotyczących:

- sterowania [cza78], [kic78], [ped93], [hoy93], [yag95], [dri96], [rut97b], [wan98], [kac01], [pie03], [tan05], [kan08], [bro08],
- podejmowania decyzji [kac01],
- identyfikacji [hel97], [łęs08],
- monitorowania [osb86], [all87], [che03],
- diagnostyki [pie04],
- predykcyjnej [jon90], [tat06].

W procesach rzeczywistych modelowanych na potrzeby automatyki, w skali działania operacyjnego, taktycznego lub strategicznego, mamy również do czynienia z niepewnością wynikającą z losowości (niepewność pomiarowa, niepewność wartości parametrów zjawisk geologicznych itp.). Do modelowania takich procesów wykorzystywane są głównie metody matematyczne z uwzględnieniem metod probabilistycznych [kac81], [kus83], [bub05], [świ09]. Wówczas przy założeniu, że nieznane wielkości są zmiennymi losowymi można wprowadzić probabilistyczny opis niepewności w oparciu o rozkłady prawdopodobieństwa lub momenty procesów.

Zdaje się być naturalnym, iż łącząc obie metody analizowania zagadnień – teorię logiki rozmytej i teorię prawdopodobieństwa, możemy w sposób pełny opisać niepewność problemów rzeczywistych. Metodologię modelowania i wnioskowania w oparciu o probabilistyczno-rozmyty model wiedzy przedstawiono m.in. w pracach [wb05], [wb07], [wb08a], [wb10]. W omawianym modelu, wiedza lingwistyczna jest zawarta w regułach postaci IF-THEN z wagami, stanowiącymi brzegowe i warunkowe prawdopodobieństwo zdarzeń rozmytych znajdujących się w poprzedniku i następniku reguł. Z założenia system

rozmyty ma pozwalać na uproszczone odtworzenie złożonego problemu badawczego. Wszelkie próby wprowadzenia dużej szczegółowości do zbyt skomplikowanych problemów zmniejszą wiarygodność wyników [łęs08]. Jednakże, biorąc pod uwagę całkowity rozkład prawdopodobieństwa zdarzeń rozmytych, ilość reguł elementarnych bazy wiedzy wynosi N^m , gdzie N stanowi liczbę zmiennych modelu wiedzy, m – liczbę zbiorów rozmytych zmiennej (przy założeniu jednakowej ilości zbiorów rozmytych dla każdej zmiennej). Duża liczba reguł ma wpływ nie tylko na czas identyfikacji rozmytego modelu wiedzy, trudności wnioskowania przy użyciu utworzonej bazy wiedzy, jak również ewentualną implementację w obiekcie rzeczywistym. Stąd też pojawia się potrzeba opracowania metody identyfikacji probabilistyczno-rozmytego modelu wiedzy, która pozwalałaby na zmniejszenie liczby reguł modelu, dając jednocześnie możliwość utrzymania zadanego poziomu błędu dopasowania. Idąc dalej w kierunku praktycznego wykorzystania utworzonego modelu wiedzy, niezbędnie staje się opracowanie całościowej koncepcji systemu, który pozwoli na wnioskowanie w oparciu o utworzoną bazę wiedzy oraz ułatwi analizę, w ujęciu probabilistycznym i rozmytym, zagadnienia obarczonego niepewnością.

2. Umiejscowienie tematu dysertacji w dyscyplinie Automatyka i Robotyka

Według Grupy Roboczej Komitetu Automatyki i Robotyki Polskiej Akademii Nauk, propozycję dotyczącą Strategicznego Programu Badawczego (SPB) na temat "Rozszerzenie Internetu – Zrobotyzowane inteligentne systemy usługowe wspomagające człowieka", można potraktować jako ogólną strategię w zakresie badań nad automatyką i robotyką [kair]. Przedstawione w punkcie 4. programu, zadania badawcze i demonstratory technologii stanowią zatem propozycję zagadnień, które należy rozwiązać w zakresie dyscypliny Automatyka i Robotyka.

Tematyka pracy doktorskiej, dotyczącej utworzenia koncepcji i implementacji systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy, nawiązuje do problemów badawczych opisanych w następujących punktach programu: 4.6 "Percepcja i kognitywistyka", 4.7 "Planowanie, zachowanie elementarne i wspomaganie decyzji" i 4.9 "Współdziałanie i inteligentne środowisko". W zakresie wyżej wymienionych podpunktów, zadania badawcze mają na celu tworzenie inteligentnego środowiska współdziałającego z człowiekiem. W szczególności podpunkt 4.6 określa obszary działań, których zadaniem jest obdarzenie robotów zdolnością uczenia się, prowadzenia logicznego rozumowania oraz przypisania sygnałom pomiarowym znaczenia symbolicznego, odnoszącego się do percepcji człowieka. Praca doktorska sytuowana w pełni w tej tematyce, nawiązuje szczególnie do problemów badawczych nad rozumowaniem w obliczu niepewnych i sprzecznych danych. Badania nad algorytmami identyfikacji bazy wiedzy za pomocą metod wyszukujących reguły asocjacji, pozwalają jednocześnie na odkrywanie znaczenia i relacji między obiektami, co również stanowi przykładowy problem badawczy z zakresu "Percepcji i kognitywistyki" Strategicznego Programu Badawczego. Zastosowanie systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy do problemów podejmowania decyzji, tematycznie usytuowuje także pracę doktorską w zakresie podpunktu 4.7 Strategicznego Programu Badawczego. Rozwinięcie idei wykorzystania nauczonej bazy wiedzy do podejmowania decyzji może także w przyszłości zaowocować implementacją systemu do sterowania elementarnymi zachowaniami robota.

Chcąc jeszcze bardziej uściślić obszar badań w zakresie tematu dysertacji, należy wspomnieć, że *systemy wnioskujące z bazami wiedzy* (systemy ekspertowe) należą do metod i technik sztucznej inteligencji a także do obszaru inżynierii wiedzy.

W pracy [bub05] Z. Bubnicki stwierdza:

„Teoria sterowania jest ściśle związana z *inżynierią wiedzy*, która zajmuje się komputeryzacją rozwiązywania problemów na podstawie reprezentacji wiedzy o problemie z zastosowaniem rozumowania, a także zagadnieniami pokrewnymi, takimi jak uzyskiwanie, gromadzenie i porządkowanie wiedzy. Tak zwane *inteligentne* systemy sterowania są komputerowymi *systemami ekspertowymi*, w których generacja decyzji sterujących oparta jest na reprezentacji wiedzy o obiekcie sterowanym lub wprost wiedzy o sterowaniu i do których projektowania i realizacji stosuje się takie metody i techniki *sztucznej inteligencji*, jak komputeryzacja operacji logicznych, uczenia się, rozpoznawania, poszukiwania rozwiązań na bazie *rozmytych* opisów wiedzy oraz komputeryzacja algorytmów *neuropodobnych*.“

Mając ponadto na uwadze liczne zastosowania systemów z bazą wiedzy w zagadnieniach automatyki i robotyki, które zostały szerzej opisane w motywacji pracy doktorskiej, można zatem umieścić temat dysertacji w dyscyplinie Automatyka i Robotyka.

3. Podstawowe pojęcia używane w dysertacji

System z bazą wiedzy (ang. *knowledge-based systems*) – system operujący i korzystający z symbolicznej bazy wiedzy (ang. *knowledge base*), którą można wydzielić z systemu.

System wnioskujący (ang. *inference system*) – system, w którym można wyróżnić bazę wiedzy oraz blok wnioskowania (ang. *inference engine*), zawierający procedury rozumowania (wnioskowania) w oparciu o bazę wiedzy i nowe fakty.

Rozmyty system wnioskujący (FIS, ang. *Fuzzy Inference System*) – system, w którym można wydzielić następujące elementy: rozmyty model wiedzy, blok rozmywania, blok wnioskowania oraz blok wyostrzania.

Model wiedzy (reprezentacja wiedzy) (ang. *knowledge representation*) – reprezentacja wiedzy, zapisana w postaci symbolicznej za pomocą faktów, stwierdzeń lub reguł.

Rozmyty model wiedzy – reprezentacja wiedzy w postaci reguł, w których atrybuty są wyrażone za pomocą wartości lingwistycznych reprezentowanych przez zbiory rozmyte.

Identyfikacja rozmytego modelu wiedzy – wybór klasy modelu wiedzy, określenie struktury i parametrów rozmytego modelu wiedzy w oparciu o dane doświadczalne i kryterium jakości identyfikacji.

4. Cel i teza pracy

Rozprawa doktorska ma na celu utworzenie autorskiej koncepcji oraz budowę systemu z bazą wiedzy, który będzie zawierał reprezentację wiedzy lingwistycznej wybranego problemu oraz prawdopodobieństwa zdarzeń w kategoriach lingwistycznych, stąd przyjęto, iż głównym celem pracy jest:

Opracowanie koncepcji i implementacja narzędzia wnioskującego z probabilistyczno-rozmytą bazą wiedzy.

Do tak określonego celu głównego pracy, sformułowano następujące cele szczegółowe:

1. Zaproponowanie struktury systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy.
2. Zaproponowanie metody identyfikacji probabilistyczno-rozmytego modelu wiedzy.
3. Opracowanie algorytmu generowania reguł, zapewniającego ograniczenie liczby reguł.
4. Implementacja zaprojektowanego systemu wnioskującego w środowisku Matlab.
5. Weryfikacja działania systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy dla wybranych zastosowań.

Prowadzone badania, w powiązaniu z wyżej wymienionymi celami szczegółowym, pozwalają na sformułowanie następującej tezy:

Istnieje możliwość opracowania koncepcji i implementacji systemu wnioskującego z bazą wiedzy, który uwzględnia niepewność informacji, jednocześnie w kategoriach probabilistycznych i rozmytych, dla zadań modelowania i podejmowania decyzji.

5. Zakres pracy

Przyjęty cel pracy oraz teza wpłynęły na układ pracy. W kolejnych rozdziałach rozprawy poruszane są zatem następujące zagadnienia:

Rozdział drugi (II) zawiera analizę podstaw teoretycznych budowy systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy. Omówiono zagadnienie identyfikacji modelu w systemach sterowania opartych o modele oraz w systemach z bazą wiedzy. Przedstawiono wybrane elementy teorii zbiorów rozmytych i teorii prawdopodobieństwa zdarzeń rozmytych. Zaprezentowano podstawy teorii rozmytych systemów wnioskujących, a w szczególności skupiono się na opisaniu możliwości reprezentacji wiedzy w ujęciu rozmytym i probabilistycznym. Dokonano analizy metod automatycznego pozyskiwania bazy wiedzy w systemach rozmytych pod kątem możliwości ich zastosowania do identyfikacji probabilistyczno-rozmytego modelu wiedzy. Na koniec rozdziału, przedstawiono szczegółowo wybraną metodę eksploracji danych – wyszukiwanie reguł asocjacji. Omówiono istotę metody, jej rodzaje oraz zastosowane algorytmy.

W **rozdziale trzecim** (III) opisano propozycję koncepcji systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy. Wybrano sposób rozmywania wartości zmiennych modelu. Zaproponowano algorytmy generowania reguł probabilistyczno-rozmytego modelu wiedzy z uwzględnieniem pełnego rozkładu zdarzeń rozmytych w regułach oraz przy zastosowaniu idei wyszukiwania rozmytych reguł asocjacji. Dokonano porównania algorytmów pod kątem wyboru algorytmu najbardziej efektywnego dla zadanych parametrów modelu wiedzy. Zaproponowano metodę identyfikacji w utworzonym systemie wnioskującym, która wykorzystuje wspomniane algorytmy generowania reguł. Omówiono w pięciu etapach procedury wnioskowania w oparciu o probabilistyczno-rozmytą bazę wiedzy.

W kolejnej części rozdziału trzeciego przedstawiono implementację, w środowisku obliczeniowym Matlab, zaproponowanego systemu wnioskującego (moduł narzędziowy PFIS). Omówiono wymagania funkcjonalne, jakie w założeniu ma spełniać utworzone narzędzie. Scharakteryzowano poszczególne funkcje modułu narzędziowego w oparciu o pliki funkcyjne środowiska Matlab oraz utworzony interfejs graficzny o nazwie PFISEDIT.

Rozdział czwarty (IV) jest poświęcony weryfikacji użyteczności opracowanego systemu dla wybranych zastosowań. Wyszczególniono następujące przykłady zastosowań:

- Zaproponowano użycie systemu wnioskującego do modelowania wybranych własności węgla, które zostały pokrótko scharakteryzowane. W szczególności:
 - na podstawie reprezentacji probabilistyczno-rozmytej dla dwóch charakterystyk jakościowych węgla, pokazano możliwość identyfikacji probabilistycznych własności zmiennych rozmytych wejścia i wyjścia,
 - w przypadku probabilistyczno-rozmytej bazy wiedzy dla filtracji danych dynamicznych omówiono charakterystykę systemu wnioskującego,
 - utworzono system wnioskujący z probabilistyczno-rozmytą bazą wiedzy dla pełnej charakterystyki statycznej parametrów procesu technologicznego wzbogacania węgla. Pokazano metodykę identyfikacji modelu wiedzy. Określono zdolność utworzonego systemu do aproksymacji zależności.
- Zaproponowano i utworzono system decyzyjny, który rozstrzyga wybór efektywnego algorytmu do budowy probabilistyczno-rozmytej bazy wiedzy, przy zadanych parametrach systemu. Określono zdolności systemu do wskazania właściwego algorytmu. Pokazano możliwości analizy modelu wiedzy systemu decyzyjnego, dla faktów określonych za pomocą wartości lingwistycznych i wartości numerycznych.

Rozdział piąty (V) zawiera podsumowanie pracy oraz wskazuje kierunki dalszych badań.

Załącznik A zawiera wyniki porównania algorytmów generowania reguł modelu wiedzy, przy różnych założeniach wejściowych.

Załącznik B zawiera opis funkcji z modułu narzędziowego PFIS.

Załącznik C zawiera całociowy opis probabilistyczno-rozmytej bazy wiedzy dla systemu pełnej charakterystyki statycznej węgla.

Załącznik D zawiera wyniki identyfikacji w systemie decyzyjnym w oparciu o większą ilość zbiorów rozmytych wybranych zmiennych.

II. Podstawy teoretyczne budowy systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy

Niniejszy rozdział przedstawia pojęcia, metody oraz modele wiedzy, które stanowią podstawę do budowy systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy. W szczególności rozdział ma na celu przedstawienie niezbędnych pojęć wykorzystywanych do późniejszego omówienia koncepcji systemu, stanowiącej temat dysertacji oraz metod służących do jego budowy. Ponadto, część teoretyczna pracy pozwoli autorowi na wykazanie się, że są mu znane istniejące dotąd rozwiązania, dotyczące systemów rozmytych i automatycznego pozyskiwania baz wiedzy dla tego typu systemów. Zostanie również nakreślona potrzeba zajmowania się powyższym problemem oraz zidentyfikowane zostaną wrażliwe punkty badań.

Analizę podstaw teoretycznych rozdzieliło na następujące zagadnienia: identyfikacja modeli w systemach opartych o modele i w systemach rozmytych z bazami wiedzy, podstawy teorii zbiorów rozmytych i teorii prawdopodobieństwa zdarzeń rozmytych, podstawy teorii rozmytych systemów wnioskujących, metody automatycznego pozyskiwania baz wiedzy w systemach rozmytych oraz zagadnienia związane z wyszukiwaniem reguł asocjacji.

1. Identyfikacja modeli w systemach sterowania opartych o modele i w systemach opartych o rozmyte bazy wiedzy

Systemy z bazami wiedzy stanowią alternatywę dla systemów opartych o modele i tradycyjne algorytmy numeryczne w sytuacjach, gdy informacja o danej dziedzinie jest niepewna, niejednoznacznie sformalizowana, bądź też zbyt trudna lub kosztowna do uzyskania. Z uwagi na różne sposoby zapisu modeli w obu przypadkach, niezbędne staje się wyjaśnienie rozumienia pojęcia identyfikacji w systemach sterowania opartych o modele oraz w systemach z bazami wiedzy.

Wg [bub74], [świ09], [jan02] jako identyfikację określa się ustalenie zależności pomiędzy określonymi wielkościami obiektu w stanie ustalonym (dla charakterystyk statycznych) lub w czasie trwania procesów (dla charakterystyk dynamicznych). W przypadku charakterystyk statycznych, utworzone zależności opisane są w postaci *modelu*:

$$\hat{y} = \Phi(u, a), \quad (1)$$

w którym, dla obiektu identyfikacji, wektorowi wielkości wejściowych $u \in U$, przy określonych parametrach modelu $a \in A$, przyporządkowany jest wektor wielkości wyjściowych $\hat{y} \in Y$.

Postać funkcji Φ oraz obszar D_a , będący obszarem przestrzeni A , do której należą parametry modelu $a \in D_a \subset A$, określają odpowiednią *klasę modeli*. Opisom modeli matematycznych poświęcono wiele uwagi m.in. w pozycjach [cze08], [kac09]. Celem identyfikacji jest wybór najlepszego modelu z określonej klasy, na podstawie *kryterium jakości identyfikacji* Q , gdzie kryterium Q ocenia odległość pomiędzy sygnałem wyjściowym obiektu Y_n i modelu \hat{Y}_n dla tego samego sygnału wejściowego U_n lub odległość między sygnałem wejściowym obiektu U_n i modelu \hat{U}_n dla tego samego sygnału wyjściowego Y_n . Identyfikację dokonuje się w oparciu o dane pomiarowe lub dane aprioryczne (znane przed wykonaniem eksperymentu pomiarowego). Wówczas zależność a_n od wyników pomiarów:

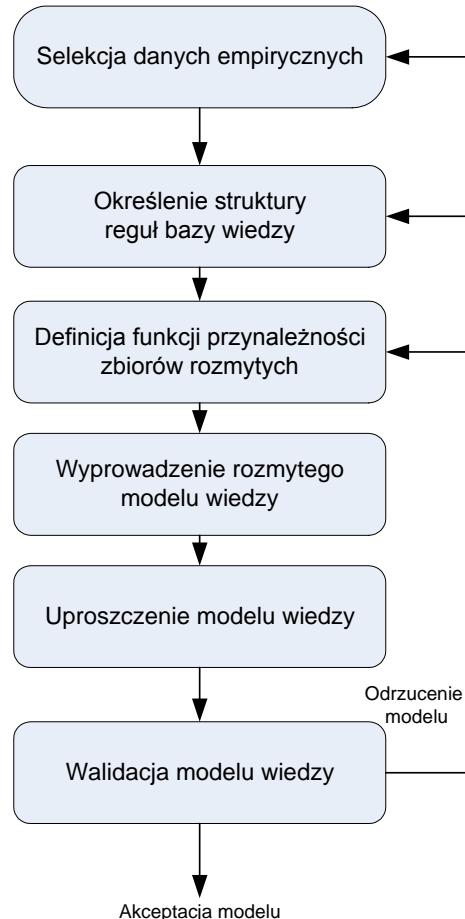
$$a_n = \Psi(U_n, Y_n) \quad (2)$$

zwana jest *algorymem identyfikacji*, natomiast jego realizacja zwana jest *identyfikatorem* (bądź po prostu *realizacją algorytmu identyfikacji*). Zatem dla systemów sterowania opartych o modele można wyróżnić następujące zadania związane z pełną identyfikacją obiektu (na podstawie [bub74], [świ09]):

- 1) określenie obiektu
- 2) określenie klasy modeli,
- 3) wyznaczenie algorytmu identyfikacji,
- 4) realizacja identyfikatora.

W razie konieczności poprawy jakości modelu, zadania te wykonywane są powtórnie.

W przypadku systemów z bazami wiedzy, a w szczególności rozmytych systemów z bazami wiedzy, identyfikacja modelu wiedzy polega na wyborze zmiennych wejściowych i wyjściowych, określeniu struktury przesłanek i konkluzji reguł bazy wiedzy, definicji funkcji przynależności zbiorów rozmytych i innych parametrów związanych z określoną strukturą modelu wiedzy. Zadania te są realizowane w oparciu o dane doświadczalne i kryterium jakości identyfikacji [oh07], [hel97]. Szczegółowe etapy identyfikacji w rozmytych systemach z bazami wiedzy przedstawiono na rysunku II-1. Zgodnie z zaleceniami modelowania i redukcji modeli zawartych w pozycji [jan02] najlepiej odwzorującym stan obiektu jest model o najmniejszej liczbie parametrów (tzw. zasada oszczędności).



Rys. II-1. Identyfikacja w rozmytych systemach z bazami wiedzy (na podstawie [hel97])

2. Wybrane elementy teorii zbiorów rozmytych

W miarę upływu czasu stało się coraz bardziej oczywiste, że reguły interesujące dla matematyków są tożsame z tymi, które wybrała Natura.

P. A. M. Dirac

Przez długi okres czasu określenie typu “niepewność”, “niejednoznaczność” miały wydźwięk typowo pejoratywny [kac86]. Traktowano je jako brak wiedzy, który w miarę badań i doświadczeń się zmniejsza. Dopiero kilkadziesiąt lat temu, zaczęto postrzegać owe pojęcia bez negatywnego nastawienia, jako odbicie rzeczywistości. Niepewność, nieprecyzyjność, niejednoznaczność wynika bowiem często nie z braku wiedzy o badanych wielkościach, ale jest efektem naturalnym, wynikającym głównie z charakterystyki analizowanych zjawisk, niepewności pomiarowych, czy subiektywnych analiz.

Taki pogląd spowodował odejście od „sztywnej” logiki dwuwartościowej, w której jakiekolwiek zdanie może być jedynie prawdziwe lub fałszywe. Powstały zatem prace z zakresu teorii mnogości logiki wielowartościowej, trójwartościowej o wartościach prawdy {0,1/2,1} [łuk20], n -wartościowej i logiki nieskończoności wartościowej [mal02]. Potrzeba matematycznego ujęcia naturalnych zjawisk nieprecyzyjnych i wieloznacznych stała się więc punktem wyjścia do wprowadzenia pojęcia i teorii zbiorów rozmytych [zad65].

Niepewność często wynika z indywidualnych odczuć ekspertów analizujących dany problem badawczy. Wprowadzenie zmiennych lingwistycznych i odpowiadających im wartości, sformułowanych w pracy [zad75], dało możliwość opisania zjawisk określeniami subiektywnych odczuć człowieka – eksperta. Opisywanie cech obiektów pojęciami nieostrymi typu: „mały”, „bardzo”, „tropię”, „nie za wiele”, „nieco więcej”, nie stanowi wówczas przeszkód. Jest to niezwykle ważne, gdyż procedura rozumowania i interpretowania pewnych wielkości przez człowieka zachodzi w sposób przybliżony, a nie ścisły. Można więc powiedzieć, że rozwój logiki rozmytej został również zainspirowany biologicznie [nał00].

Bardzo szybko teoria zbiorów rozmytych znalazła zastosowanie na gruncie teorii sterowania i teorii systemów [kac86]. W praktyce, gdzie w systemie sterowania istotnym elementem był człowiek, ścisłe metody matematyczne zazwyczaj zawodziły. L.A. Zadeh sformułował zasadę, która mówi, iż wszelkie próby odzwierciedlenia złożonego systemu sterowania, zawierającego niedokładne pojęcia, powodowały zmniejszenie wiarygodności wyników i precyzyjnej analizy ("złożoność i precyzja są ze sobą w relacji odwrotnej") [zad73]. Zastosowanie logiki rozmytej oraz naśladowanie intuicyjnego wnioskowania przybliżonego zastosowanego w umyśle ludzkiem umożliwiło budowanie rozmytych systemów sterowania. Systemy tego typu mogą stanowić alternatywę dla złożonych modeli matematycznych wszędzie tam, gdzie istnieją zależności przyczynowo-skutkowe, a model matematyczny rozważanego zjawiska jest niedostępny lub jego wyznaczenie jest zbyt kosztowne.

2.1. Definicja i własności zbiorów rozmytych

Teoria zbiorów rozmytych została zapoczątkowana przez profesora Uniwersytetu Kalifornijskiego w Berkeley Lotfiego A. Zadeha [zad65] w 1965 roku, jako uogólnienie klasycznego pojęcia zbiorów. Uogólnienie to jest dokonywane następująco. Dowolny zbiór klasyczny (ostry, ang. *crisp set*) C , można zdefiniować za pomocą określenia funkcji charakterystycznej wg poniższej definicji [kac86], [dri93].

DEFINICJA 1. Funkcja $\mu_C : X \rightarrow \{0,1\}$ jest *funkcją charakterystyczną* zbioru C wtedy i tylko wtedy, gdy dla wszystkich $x \in X$

$$\mu_C(x) = \begin{cases} 1 & \text{dla } x \in C \\ 0 & \text{dla } x \notin C \end{cases}. \quad (3)$$

Zbiór C jest całkowicie określony przez elementy $x \in C$, które go tworzą. Elementy $x \notin C$ nie należą do zbioru C . W teorii zbiorów rozmytych funkcja charakterystyczna jest uogólniona do funkcji przynależności, która pozwala zdefiniować zbiór rozmyty jako zbiór elementów, mogący w pewnym stopniu przynależeć do zbioru rozmytego. Poniższa definicja określa formalnie pojęcie zbioru rozmytego ([zad65], [kac86], [bez93], [dri93], [kli95], [zim96], [pie03], [łes08], [háj06]), który czasem określany jest jako podzbiór rozmyty [yag95].

DEFINICJA 2. Niech \mathbb{N} będzie przestrzenią (ang. *universal*), wówczas *zbior rozmyty A* (ang. *fuzzy set*) w niepustej przestrzeni \mathbb{N} jest określony przez funkcję charakterystyczną zwaną funkcją przynależności μ_A (ang. *membership function*) w formie:

$$\mu_A : \mathbb{N} \rightarrow \langle 0,1 \rangle, \quad (4)$$

gdzie: $\langle 0,1 \rangle$ oznacza przedział liczb rzeczywistych.

Zbiór rozmyty A może być opisany poprzez podanie funkcji przynależności $\mu_A(x)$ bądź też poprzez określenie zbioru uporządkowanych par $(x, \mu_A(x))$ ([kac86], [dri93], [zim96], [rut97b], [rut06], [łes08], [pie03]).

DEFINICJA 3. *Zbiór rozmyty A* określony w niepustej przestrzeni \mathbb{N} można przedstawić jako zbiór par:

$$A = \{(x, \mu_A(x)) \mid x \in \mathbb{N}, \mu_A(x) \in \langle 0,1 \rangle\}, \quad (5)$$

gdzie μ_A jest funkcją przynależności, która każdemu elementowi przestrzeni \mathbb{N} przyporządkowuje stopień przynależności (ang. *grades of membership*) do danego zbioru rozmytego. Można rozróżnić trzy przypadki:

- a) $\mu_A = 0$ ozn. brak przynależności elementu x do zbioru rozmytego A , tzn. $x \notin A$,
- b) $0 < \mu_A < 1$ ozn. częściową przynależność elementu x do zbioru rozmytego A ,
- c) $\mu_A = 1$ ozn. pełną przynależność elementu x do zbioru rozmytego A , tzn. $x \in A$.

W niniejszej dysertacji jest również używana inna notacja zapisu formalnego zbioru rozmytego [zad65], dla której zbiór rozmyty A , w dyskretniej przestrzeni \mathbb{N} , jest opisany jako:

$$A = \sum_{x \in \mathbb{N}} \mu_A(x) / x, \quad (6)$$

gdzie: symbol \sum oznacza sumę mnogościową, natomiast symbol / jest separatorem.

Stosowane w definicjach zbiorów rozmytych (def. 2-3) funkcje przynależności odzwierciedlają na obiektach z przestrzeni rozważań uporządkowanie skojarzone ze zbiorem pewnej własności (na podstawie [ryk]). W praktyce funkcje przynależności wynikają z kontekstu sytuacyjnego i są definiowane subiektywnie przez eksperta, bądź też kształt ich jest wynikiem uczenia na zbiorze danych empirycznych. Podrozdział II.1.2. zawiera spotykane interpretacje funkcji przynależności. Do najczęściej stosowanych funkcji przynależności należą funkcje odcinkowo-liniowe (symetryczne i niesymetryczne funkcje trójkątne, zwane również funkcjami klasy t [rut06] bądź λ , a także funkcje przynależności trapezowe oraz ich modyfikacje – prawa zewnętrzna funkcja klasy γ [rut06] oraz lewa zewnętrzna funkcja klasy L). Przykłady funkcji przynależności odcinkowo-liniowych wykorzystywanych w dalszej części pracy przedstawiono w tabeli II-1. Specyficzną funkcją przynależności jest rozmyty singleton (tab. II-2). Charakteryzuje go bowiem tylko jeden element, który w pełni należy do zbioru rozmytego. Wykorzystywany jest głównie do definicji zbiorów w sensie nierozmytym bądź do realizacji operacji rozmywania w systemach wnioskujących [rut06]. Opis problemów z definiowaniem i interpretacją funkcji przynależności można znaleźć w [jan07].

Tab. II-1. Przykłady odcinkowo-liniowych funkcji przynależności [rut06], [łes08], [pie03], [dri96]

Nazwa funkcji przynależności	Kształt funkcji przynależności	Wzór funkcji przynależności
F.p. trójkątna (klasy t bądź λ)		$\mu(x; a, b, c) = \begin{cases} 0 & \text{dla } x \leq a \\ \frac{x-a}{b-a} & \text{dla } a < x \leq b \\ \frac{b-a}{c-a} & \text{dla } b < x \leq c \\ 0 & \text{dla } x > c \end{cases}$ <p>oraz</p> $\mu(x; a, b, c) = \max \left[\min \left(\frac{x-a}{b-a}, \frac{c-x}{c-b} \right), 0 \right]$
F.p. trapezowa		$\mu(x; a, b, c, d) = \begin{cases} 0 & \text{dla } x \leq a \\ \frac{x-a}{b-a} & \text{dla } a < x \leq b \\ 1 & \text{dla } b < x \leq c \\ \frac{d-x}{d-c} & \text{dla } c < x \leq d \\ 0 & \text{dla } x > d \end{cases}$ <p>oraz</p> $\mu(x; a, b, c, d) = \max \left[\min \left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c} \right), 0 \right]$

Tab. II-1. c.d. Przykłady odcinkowo-liniowych funkcji przynależności [rut06], [łes08], [pie03], [dri96]

Nazwa funkcji przynależności	Kształt funkcji przynależności	Wzór funkcji przynależności
F.p. lewa zewnętrzna (klasy L)		$\mu(x; a, b) = \begin{cases} 1 & \text{dla } x \leq a \\ \frac{b-x}{b-a} & \text{dla } a < x \leq b \\ 0 & \text{dla } x > b \end{cases}$ oraz $\mu(x; a, b) = \max \left[\min \left(1, \frac{b-x}{b-a} \right), 0 \right]$
F.p. prawa zewnętrzna (klasy γ)		$\mu(x; a, b) = \begin{cases} 0 & \text{dla } x \leq a \\ \frac{x-a}{b-a} & \text{dla } a < x \leq b \\ 1 & \text{dla } x > b \end{cases}$ oraz $\mu(x; a, b) = \max \left[\min \left(\frac{x-a}{b-a}, 1 \right), 0 \right]$

Tab. II-2. Funkcja przynależności – rozmyty singleton [rut06], [pie03]

Nazwa funkcji przynależności	Kształt funkcji przynależności	Wzór funkcji przynależności
F.p. rozmyty singleton		$\mu(x; a) = \begin{cases} 1 & \text{dla } x = a \\ 0 & \text{dla } x \neq a \end{cases}$

Jako reprezentację matematyczną intuicyjnych funkcji przynależności stosuje się również bardziej złożone nieliniowe funkcje tj. symetryczne i niesymetryczne funkcje Gaussa, funkcję typu dzwonowego lub inne funkcje nieliniowe (tab. II-3) [pie03].

Tab. II-3. Przykłady nieliniowych funkcji przynależności [pie03], [łes08]

Nazwa funkcji przynależności	Kształt funkcji przynależności	Wzór funkcji przynależności
F.p. gaussowska		$\mu(x; \mu, \sigma) = \exp \left(-\frac{(x - \mu)^2}{2\sigma^2} \right),$ gdzie: μ – środek, σ^2 – wariancja

Tab. II-3. c.d. Przykłady nieliniowych funkcji przynależności [pie03], [łęs08]

Nazwa funkcji przynależności	Kształt funkcji przynależności	Wzór funkcji przynależności
F.p. dzwonowego typu		$\mu(x; \mu, \sigma, \gamma) = \frac{1}{1 + \left \frac{x - \mu}{\sigma} \right ^{\gamma}}$ <p>gdzie: μ – środek, σ – szerokość, γ – nachylenie zbioru rozmytego</p>
F.p. klasy s		$\mu(x; a, b) = \begin{cases} 0 & \text{dla } x \leq a \\ 2\left(\frac{x-a}{b-a}\right)^2 & \text{dla } a < x \leq (a+b)/2 \\ 1-2\left(\frac{x-b}{b-a}\right)^2 & \text{dla } (a+b)/2 < x \leq b \\ 1 & \text{dla } x > b \end{cases}$
F.p. klasy z		$\mu(x; a, b) = \begin{cases} 1 & \text{dla } x \leq a \\ 1-2\left(\frac{x-a}{b-a}\right)^2 & \text{dla } a < x \leq (a+b)/2 \\ 2\left(\frac{x-b}{b-a}\right)^2 & \text{dla } (a+b)/2 < x \leq b \\ 0 & \text{dla } x > b \end{cases}$

W przypadkach szczególnych, funkcja przynależności $\mu_A(x)$ jest wprost proporcjonalna do funkcji gęstości rozkładu prawdopodobieństwa $P(A|x)$ ([łęs08], [mab97], [don08]), czyli względnej częstości, z jaką element x jest uznawany w losowym eksperymencie za należący do zbioru A . Powyższa interpretacja stała się punktem wyjścia do utworzenia przez Hirotę [hir77] pojęcia zbioru probabilistycznego, w którym funkcja przynależności stanowi zmienną losową elementu x [czo85]. Jednakże, koncepcja systemu wnioskującego nie wykorzystuje idei zbiorów probabilistycznych. Więcej informacji na temat porównania probabilistycznego i rozmytego podejścia można znaleźć w pozycji [nur09].

Bardzo ważnym pojęciem dla modelowania rozmytego stała się zmienna lingwistyczna, której wartości wzięte z języka naturalnego mają swoją interpretację w przestrzeni numerycznej w postaci zbiorów rozmytych. Formalnie zmienna lingwistyczna określa się za pomocą następującej definicji [zad75b].

DEFINICJA 4. *Zmienna lingwistyczna (ang. linguistic variable) zdefiniowana jest jako piątka:*

$$\{x_{nazwa}; L(x); \aleph; G, M_x\} \quad (7)$$

gdzie:

x_{nazwa} – nazwa zmiennej lingwistycznej (np. prędkość wiatru, czas trwania zadania),

$L(x)$ – zbiór wartości (terminów, termów) lingwistycznych (ang. *linguistic value*), jakie przyjmuje x , czyli słowa lub zdania w języku naturalnym (np. wysoka, średnia, niska),

\aleph – przestrzeń rozważań (np. kwoty z zakresu 0-200 tys. zł, zakres prędkości wiatru 0-30 m/s),

G – zbiór reguł syntaktycznych, które umożliwiają utworzenie wszystkich terminów w zbiorze $L(x)$,

M_x – funkcja semantyczna, przyporządkowująca każdej wartości lingwistycznej ze zbioru $L(x)$ zbiór rozmyty zdefiniowany nad \aleph .

Na przykład, niech zmienna lingwistyczna o nazwie x ="prędkość wiatru" posiada zbiór wartości lingwistycznych:

$$L(x) = \{"bardzo silna", "silna", "dość silna", "umiarkowana", "łagodna", "słaba", "cicha"\}, \quad (8)$$

zdefiniowanych w obszarze rozważań $\aleph = \langle 0,30 \rangle$ [m/s]. Wówczas definicję wartości lingwistycznej "silna" można zapisać jako:

$$\mu_{silny}(x) = \begin{cases} \frac{1}{5}x - 4 & \text{dla } x \leq 25 \\ -\frac{1}{5}x + 6 & \text{dla } x \geq 25 \end{cases}. \quad (9)$$

Poniżej zostaną przedstawione podstawowe parametry (wskaźniki) związane ze zbiorami rozmytymi tj. nośnik, α -przekrój oraz moc zbioru rozmytego ([zad75a], [zad75b], [zad75c], [rut06], [kac86], [pie03], [łes08]). Omówione pojęcia będą używane w trakcie opisu koncepcji systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy.

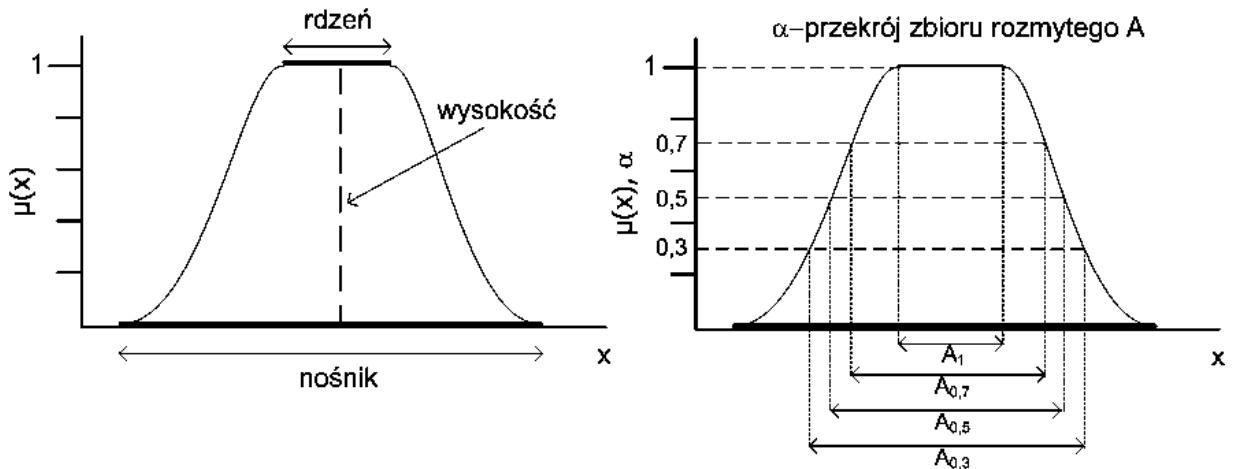
DEFINICJA 6. Nośnikiem (ang. *support*) zbioru rozmytego A nazywamy podzbiór nierożmyty zbioru A złożony z elementów, dla których funkcja przynależności jest niezerowa

$$Supp(A) = \{x \in \aleph \mid \mu_A(x) > 0\}. \quad (10)$$

DEFINICJA 7. α -przekrój zbioru (ang. *α -cut*) rozmytego A nazywamy podzbiór nierożmyty zbioru A złożony z elementów, dla których funkcja przynależności przyjmuje wartości większe bądź równe α

$$A_\alpha = \{x \in \aleph \mid \mu_A(x) \geq \alpha\}, \quad \forall \alpha \in \langle 0,1 \rangle. \quad (11)$$

Graficzne przedstawienie pojęć (def. 6-7) zawiera rysunek II-2.



Rys. II-2. Interpretacja graficzna wybranych parametrów zbioru rozmytego

DEFINICJA 8. *Moc* zbioru rozmytego (ang. *power*), zwana także *licznością nierożmytą* (ang. *sigma-count*) ([kac01], [zad73]), stanowi sumę stopni przynależności poszczególnych elementów zbioru rozmytego i jest obliczana za pomocą wzoru:

- a) dla zbioru rozmytego określonego w przestrzeni dyskretniej

$$|A| = \text{Power}(A) = \sum_{x \in \text{Supp}(A)} \mu_A(x) \quad (12)$$

- b) dla zbioru rozmytego określonego w przestrzeni ciągłej

$$|A| = \text{Power}(A) = \int_{x \in \text{Supp}(A)} \mu_A(x) dx \quad (13)$$

Moc zbioru rozmytego zwana jest często *liczbą kardynalną* i jest oznaczona jako $\text{card}(A)$ ([kac86][pie03]).

2.2. Operacje mnogościowe na zbiorach rozmytych

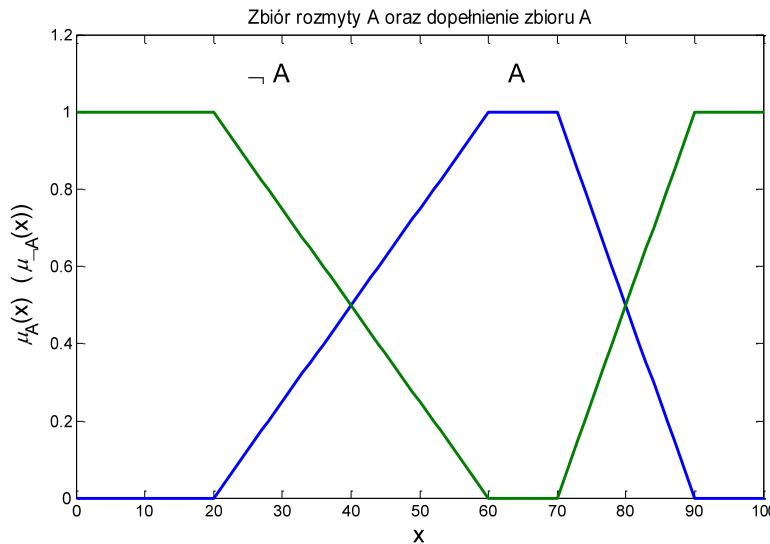
W klasycznej teorii zbiorów, podstawowe operacje, tj. suma, iloczyn, czy dopełnienie zbiorów są prostymi działaniami realizowanymi jednoznacznie. Natomiast, w przypadku zbiorów rozmytych, z uwagi na używanie stopniowanych funkcji charakterystycznych [dri03], definicja i interpretacja operacji mnogościowych nie jest już tak prosta.

Kolejne definicje będą przedstawały operacje mnogościowe na zbiorach rozmytych wprowadzone przez Zadeha [zad65]. Analizy z użyciem owych operacji można znaleźć w wielu publikacjach np. [dub84], [ian98].

DEFINICJA 9. *Dopełnieniem* (ang. *complement*) zbioru rozmytego $A \in \mathbb{N}$ nazywamy zbiór rozmyty $\neg A$ (ozn. też jako \hat{A}) i definiujemy następująco:

$$\mu_{\neg A}(x) = 1 - \mu_A(x), \quad \forall x \in \mathbb{N}. \quad (14)$$

Rysunek II-3 przedstawia przykład funkcji przynależności zbioru rozmytego A oraz funkcji przynależności dla dopełnienia zbioru rozmytego A ($\neg A$).



Rys. II-3. Funkcja przynależności zbioru rozmytego A oraz jego dopełnienia $\neg A$

DEFINICJA 10. Przecięciem (ang. *intersection*) zbiorów rozmytych A i $B \in \mathbb{N}$ (zwany też iloczynem zbiorów rozmytych) nazywamy zbiór rozmyty $A \cap B$ o następującej funkcji przynależności

$$\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \min(\mu_A(x), \mu_B(x)), \quad \forall x \in \mathbb{N}. \quad (15)$$

DEFINICJA 11. Sumą (ang. *union*) zbiorów rozmytych A i $B \in \mathbb{N}$ nazywamy zbiór rozmyty $A \cup B$ o następującej funkcji przynależności

$$\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = \max(\mu_A(x), \mu_B(x)), \quad \forall x \in \mathbb{N}. \quad (16)$$

Wzory (15-16) można uogólnić na n -argumentowe operacje na zbiorach rozmytych. Powyższe definicje są podstawowe, ale nie jedyne. Inne definicje operacji dopełnienia, przecięcia i sumy zbiorów rozmytych można znaleźć np. w polskojęzycznej literaturze w pracach [kac86], [rut97b], [pie03], [cpa09].

W niniejszej dysertacji operacje przecięcia i sumy zbiorów rozmytych, wykorzystywane przy tworzeniu bazy wiedzy oraz wnioskowania, będą realizowane z użyciem norm trójkątnych (w skrócie t-normy i s-normy). Operator t-normy jest funkcją modelującą operację przecięcia zbiorów rozmytych, natomiast operator s-normy wykorzystywany jest do modelowania operacji sumowania zbiorów rozmytych. Publikacje

[fod94], [kac01], [rut06], [pis09] wskazują również, iż normy trójkątne stały się także przydatnym narzędziem agregacji w różnych metodach podejmowania decyzji i sterowania, opartych na teorii zbiorów rozmytych.

DEFINICJA 12. Operator *t-normy* (ozn. czasem jako $\overset{T}{*}$), jest to funkcja dwóch zmiennych $T: \langle 0,1 \rangle \times \langle 0,1 \rangle \rightarrow \langle 0,1 \rangle$, jeżeli:

- a) funkcja T jest niemalejąca względem obu argumentów (ang. *monotonicity, increasing*):

$$T(\mu_A(x), \mu_C(x)) \leq T(\mu_B(x), \mu_D(x)) \quad \text{dla } \mu_A(x) \leq \mu_B(x), \mu_C(x) \leq \mu_D(x); \quad (17)$$

- b) funkcja T spełnia warunek przemienności (ang. *commutativity*):

$$T(\mu_A(x), \mu_B(x)) = T(\mu_B(x), \mu_A(x)); \quad (18)$$

- c) funkcja T spełnia warunek łączności (ang. *associativity*):

$$T[T(\mu_A(x), \mu_B(x)), \mu_C(x)] = T[\mu_A(x), T(\mu_B(x), \mu_C(x))]; \quad (19)$$

- d) funkcja T spełnia warunek brzegowy (ang. *one as a neutral element*):

$$T(\mu_A(x), 1) = \mu_A(x), \quad (20)$$

gdzie, $\mu_A(x), \mu_B(x), \mu_C(x), \mu_D(x) \in \langle 0,1 \rangle$ [łęs08].

Korzystając z warunku łączności, definicję t-normy można uogólnić na funkcję n -argumentową następująco

$$\begin{aligned} \underset{i=1}{\overset{n}{T}}(\{\mu_{A_i}(x)\}) &= T(\underset{i=1}{\overset{n-1}{T}}(\{\mu_{A_i}(x)\}), \mu_{A_n}(x)) = T(\mu_{A_1}(x), \mu_{A_2}(x), \dots, \mu_{A_n}(x)) = \\ &= \mu_{A_1}(x) \overset{T}{*} \mu_{A_2}(x) \overset{T}{*} \dots \overset{T}{*} \mu_{A_n}(x). \end{aligned} \quad (21)$$

DEFINICJA 13. Operator *s-normy* (ozn. czasem jako $\overset{s}{*}$), zwany także *t-konormą*, jest funkcją dwóch zmiennych $S: \langle 0,1 \rangle \times \langle 0,1 \rangle \rightarrow \langle 0,1 \rangle$, jeżeli funkcja S jest niemalejąca względem obu argumentów, spełnia warunek przemienności (16) i łączności (17) oraz następujący warunek brzegowy [łęs08]:

$$S(\mu_A(x), 0) = \mu_A(x), \quad \mu_A(x) \in \langle 0,1 \rangle. \quad (22)$$

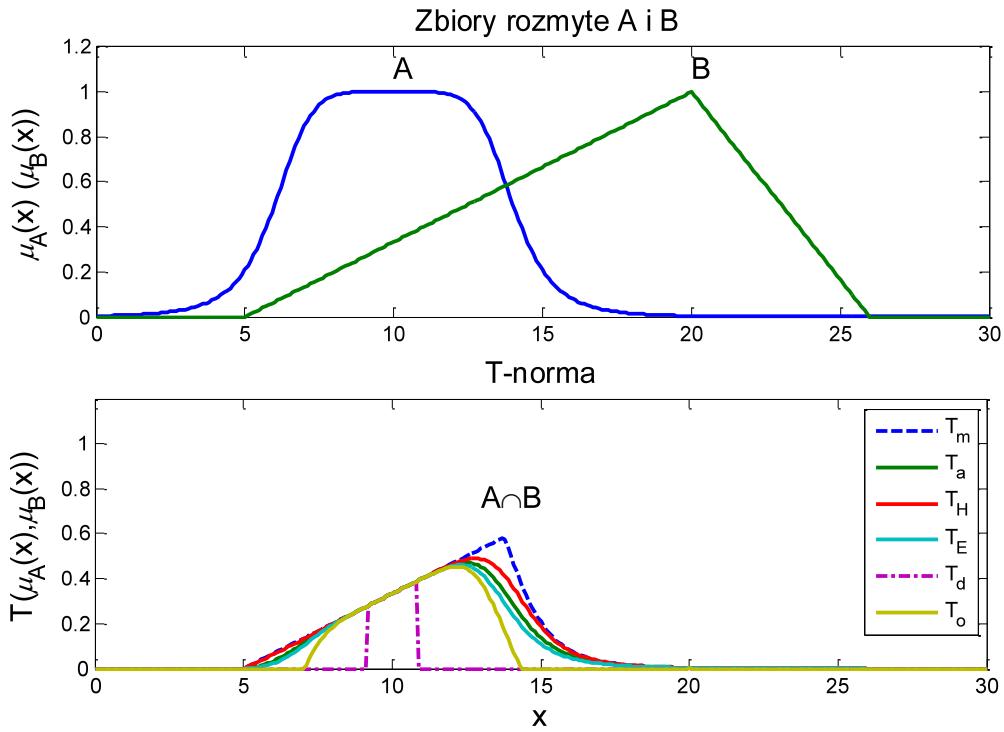
Korzystając z warunku łączności, definicję s-normy można także uogólnić na funkcję n -argumentową następująco:

$$\begin{aligned}
 S(\{\mu_{A_i}(x)\}) &= S(S(\{S(\{\mu_{A_i}(x)\}), \mu_{A_n}(x)) = S(\mu_{A_1}(x), \mu_{A_2}(x), \dots, \mu_{A_n}(x)) = \\
 &= \mu_{A_1}(x) *^S \mu_{A_2}(x) *^S \dots *^S \mu_{A_n}(x).
 \end{aligned} \tag{23}$$

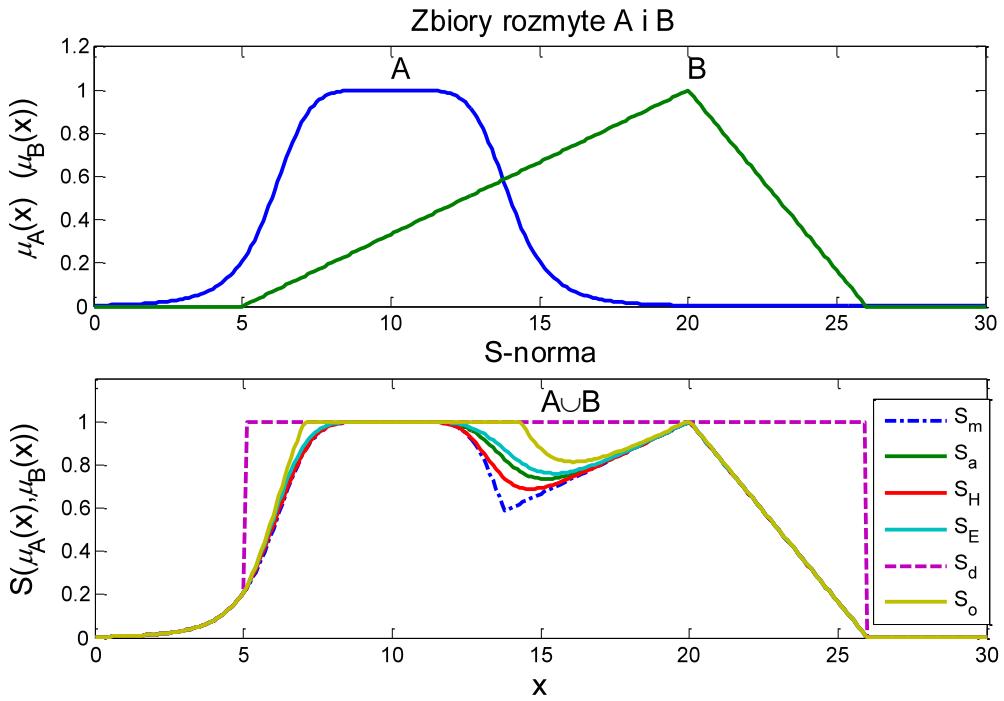
Podstawowe wzory nienastawialnych operatorów s-normy $S(\mu_A(x), \mu_B(x))$ i t-normy $T(\mu_A(x), \mu_B(x))$ przedstawione są w tabeli II-4.

Tab. II-4. Wybrane normy trójkątne [pie03], [łes08], [cpa09]

Nazwa normy trójkątnej	Operator t-normy	Operator s-normy
Normy trójkątne min/max, zwane normami trójkątnymi Zadeha	$T_m(\mu_A(x), \mu_B(x)) =$ $= MIN(\mu_A(x), \mu_B(x))$	$S_m(\mu_A(x), \mu_B(x)) =$ $= MAX(\mu_A(x), \mu_B(x))$
Normy trójkątne algebraiczne	$T_a(\mu_A(x), \mu_B(x)) =$ $= \mu_A(x) \cdot \mu_B(x)$	$S_a(\mu_A(x), \mu_B(x)) =$ $= \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x)$
Normy trójkątne Hamachera	$T_H(\mu_A(x), \mu_B(x)) =$ $= \frac{\mu_A(x) \cdot \mu_B(x)}{\mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x)}$	$S_H(\mu_A(x), \mu_B(x)) =$ $= \frac{\mu_A(x) + \mu_B(x) - 2\mu_A(x) \cdot \mu_B(x)}{1 - \mu_A(x) \cdot \mu_B(x)}$
Normy trójkątne Einsteina	$T_E(\mu_A(x), \mu_B(x)) =$ $= \frac{\mu_A(x) \cdot \mu_B(x)}{2 - (\mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x))}$	$S_E(\mu_A(x), \mu_B(x)) =$ $= \frac{\mu_A(x) + \mu_B(x)}{1 + \mu_A(x) \cdot \mu_B(x)}$
Normy trójkątne drastyczne	$T_d(\mu_A(x), \mu_B(x)) =$ $\begin{cases} MIN(\mu_A(x), \mu_B(x)) \\ \text{dla } MAX(\mu_A(x), \mu_B(x)) = 1 \\ 0 \quad \text{poza tym} \end{cases}$	$S_d(\mu_A(x), \mu_B(x)) =$ $\begin{cases} MAX(\mu_A(x), \mu_B(x)) \\ \text{dla } MIN(\mu_A(x), \mu_B(x)) = 0 \\ 1 \quad \text{poza tym} \end{cases}$
Normy trójkątne ograniczone	$T_o(\mu_A(x), \mu_B(x)) =$ $= MAX(0, \mu_A(x) + \mu_B(x) - 1)$	$S_o(\mu_A(x), \mu_B(x)) =$ $= MIN(1, \mu_A(x) + \mu_B(x))$



Rys. II-4. Funkcje przynależności przecięcia zbiorów rozmytych A i B z uwzględnieniem różnych operatorów t-normy



Rys. II-5. Funkcje przynależności sumy zbiorów rozmytych A i B z uwzględnieniem różnych operatorów s-normy

Rysunki II-4, II-5 ilustrują relacje, które spełniają normy trójkątne:

$$T_d(\mu_A, \mu_B) \leq T(\mu_A, \mu_B) \leq T_m(\mu_A, \mu_B), \quad (24)$$

$$S_m(\mu_A, \mu_B) \leq S(\mu_A, \mu_B) \leq S_d(\mu_A, \mu_B). \quad (25)$$

W zastosowaniach logiki rozmytej wykorzystywane są również parametryzowane odmiany norm trójkątnych, które charakteryzują się tym, że odpowiadające im hiperpłaszczyzny są modyfikowane w zależności od wartości parametru p . Należą do nich m.in.: normy trójkątne Webara, Hamachera, Yagera, Domi, Franka, Dubois i Pradego, Schweizera, czy Mizumoto. Szczegółowe informacje na temat norm trójkątnych zostały przedstawione m.in. w pozycjach [kle04], [rut06], [pra07], [cpa09]. Odmiany norm trójkątnych, odzwierciedlających operacje agregacji, zostały również zamieszczone w [det01].

Kolejnym ważnym pojęciem w teorii zbiorów rozmytych jest relacja rozmyta [zad65]. Pozwala ona na sformalizowanie nieprecyzyjnych stwierdzeń tj. na przykład "y jest znacznie większy od x" [duc00], [zad65]. Jest także nieodzownym elementem wykorzystywanym podczas wnioskowania rozmytego.

DEFINICJA 14. *Relacją rozmytą R (ang. fuzzy relation)* między dwoma niepustymi zbiorami (nierożmytymi) \mathbb{N} i \mathfrak{I} nazywa się zbiór rozmyty określony na iloczynie kartezjańskim $\mathbb{N} \times \mathfrak{I}$:

$$R \subseteq \mathbb{N} \times \mathfrak{I} = \{(x, y), \mu_R(x, y) \mid x \in \mathbb{N}, y \in \mathfrak{I}, \mu_R(x, y) \in \langle 0, 1 \rangle\}, \quad (26)$$

gdzie $\mu_R(x, y)$ jest funkcją przynależności relacji R , która realizuje odwzorowanie zbioru $\mathbb{N} \times \mathfrak{I}$ w przedział $\langle 0, 1 \rangle$ i może być interpretowana jako stopień powiązania elementów x i y [łęs08], [rut06]. Im elementy są bardziej ze sobą powiązane, tym stopień jest bliższy jedynce.

W procedurach wnioskowania rozmytego istotne są jeszcze dwie relacje: cylindryczne rozszerzenie zbioru (relacji) oraz projekcja relacji rozmytej [łęs08], [yag95], [rut06].

DEFINICJA 15. *Cylindryczne rozszerzenie zbioru A lub relacji R na przestrzeń $\mathbb{N} \times \mathfrak{I}$ ozn. $Ce(A)$ ($Ce(R)$) polega na uzyskaniu zbioru lub relacji rozmytej o większej wymiarowości:*

$$\forall_{x \in \mathbb{N}, y \in \mathfrak{I}} \mu_{Ce(A)}(x, y) = \mu_A(x) \quad (\text{lub} \quad \forall_{x \in \mathbb{N}, y \in \mathfrak{I}} \mu_{Ce(R)}(x, y) = \mu_R(x)). \quad (27)$$

DEFINICJA 16. *Projekcja relacji rozmytej R na przestrzeń \mathbb{N} ozn. $proj_{\mathbb{N}}(R)$ polega na uzyskaniu zbioru lub relacji rozmytej o mniejszej wymiarowości:*

$$\forall_{x \in \mathbb{N}} \mu_{proj_{\mathbb{N}}(R)}(x) = \sup_{y \in \mathfrak{I}} \mu_R(x, y), \quad (28)$$

gdzie: oznaczenie \sup (łac. *supremum*) określa kres góry zbioru.

3. Wybrane elementy teorii prawdopodobieństwa

W otaczającej nas rzeczywistości wszystkie zjawiska podlegają wpływom losowości [plu00]. Można tu wyróżnić przykłady tj. losowy charakter prądu, czas poprawnej pracy urządzeń pochodzących z masowej produkcji, ciśnienie atmosferyczne i prędkość wiatru w określonym punkcie kuli ziemskiej itp. Aby lepiej przybliżyć otaczający świat w warunkach nie laboratoryjnych, należy uwzględnić, przy opisywaniu modeli badanych zjawisk, ów czynnik losowy – niezależny od praw natury [bob86]. Pomocna jest przy tym

probabilistyka – dział matematyki zajmujący się wykrywaniem i analizowaniem prawidłowości w zakresie zdarzeń losowych. Poniżej zostaną przedstawione podstawowe definicje, związane z działem probabilistyki dotyczącym rachunku prawdopodobieństwa. Definicje te pozwolą na zrozumienie istoty rachunku prawdopodobieństwa, zarówno w ujęciu klasycznym, jak i w stosunku do zdarzeń rozmytych oraz pozwolą przybliżyć interpretację modelu wiedzy, będącego tematem niniejszej dysertacji. Dla zdefiniowania prawdopodobieństwa zdarzeń rozmytych zostaną przytoczone elementy klasycznej teorii prawdopodobieństwa.

3.1. Przestrzeń zdarzeń elementarnych, zdarzenia, prawdopodobieństwo w ujęciu klasycznym

W ujęciu teorii prawdopodobieństwa, pojęciem pierwotnym jest pojęcie *przestrzeni zdarzeń elementarnych* (Ω), stanowiącej zbiór *zdarzeń elementarnych*. W interpretacji zagadnień doświadczalnych, przestrzeń zdarzeń elementarnych można rozumieć jako zbiór wszystkich możliwych, elementarnych i niepodzielnych wyników doświadczenia lub obserwacji [plu00]. Wówczas poszczególny wynik doświadczenia stanowi zdarzenie elementarne.

DEFINICJA 17. *Zdarzenie losowe* (ang. *random event*), lub w skrócie *zdarzenie*, jest podzbiorem przestrzeni zdarzeń elementarnych Ω .

Powyższa definicja [plu00] jest podejściem uproszczonym, gdyż z teoretycznego punktu widzenia, nie każdy podzbiór przestrzeni elementarnej jest zdarzeniem. Operacje wykonywane na zdarzeniach powinny dawać w wyniku również zdarzenia. Nie dzieje się tak, w przypadku każdego podzbioru przestrzeni elementarnej. W niniejszej dysertacji będziemy rozpatrywać uproszczone przypadki, gdy przestrzeń zdarzeń elementarnych jest zbiorem przeliczalnym. Wówczas, każdy podzbiór tej przestrzeni można rozpatrywać jako zdarzenie [bob86]. Jednakże definiując ogólne, klasyczne pojęcie prawdopodobieństwa konieczne staje się zdefiniowanie odpowiedniej rodziny F podzbiorów przestrzeni elementarnej [bob86].

DEFINICJA 18. Rodzina F podzbiorów przestrzeni elementarnej Ω stanowi *ciało zdarzeń*, jeśli spełnia aksjomaty:

- 1) istnieje $A \subset \Omega$ takie, że $A \in F$ (F nie jest zbiorem pustym),
- 2) $A \in F \Rightarrow \Omega - A \in F$ (dopełnienie zdarzenia A również należy do rodziny F),
- 3) $A_1, A_2, \dots \in F \Rightarrow \bigcup_{i=1}^{\infty} A_i \in F$ (warunek przeliczalnej addytywności).

W tym ujęciu teoretycznym, zdarzenie losowe stanowi wówczas zbiór zdarzeń elementarnych należących do ciała zdarzeń. Można wyróżnić następujące specjalne podzbiory przestrzeni Ω (na podstawie [bob86], [plu00]):

DEFINICJA 19. *Zdarzenie pewne* (ozn. również jako Ω) (ang. *certain event*) jest podzbiorem przestrzeni zdarzeń elementarnych Ω zawierający wszystkie elementy tej przestrzeni.

DEFINICJA 20. *Zdarzenie niemożliwe* (ozn. \emptyset) (ang. *impossible event*) jest pustym podzbiorem przestrzeni zdarzeń elementarnych Ω .

W terminologii rachunku prawdopodobieństwa działania na zdarzeniach losowych wynikają z ogólnej teorii zbiorów. Przykładowo, pojęcia sumy i iloczynu zdarzeń oraz zdarzeń rozłącznych, uogólnione na przypadek skończonej n liczby zdarzeń losowych, definiuje się następująco (na podstawie [bob86], [plu00]):

DEFINICJA 21. *Sumą zdarzeń losowych A_1, \dots, A_n* (ang. *union of the events*) nazywamy zbiór złożony z tych i tylko tych zdarzeń elementarnych, które należą do co najmniej jednego ze zdarzeń losowych A_1, \dots, A_n . Pojęcie oznaczane jest jako:

$$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i = \{x \mid x \in A_1 \text{ lub } x \in A_2 \text{ lub } \dots \text{ lub } x \in A_n\}. \quad (29)$$

DEFINICJA 22. *Iloczynem zdarzeń losowych A_1, \dots, A_n* (ang. *intersection of the events*) nazywamy zbiór złożony z tych i tylko tych zdarzeń elementarnych, które należą do każdego ze zdarzeń losowych A_1, \dots, A_n . Pojęcie oznaczane jest jako:

$$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i = \{x \mid x \in A_1 \text{ i } x \in A_2 \text{ i } \dots \text{ i } x \in A_n\}. \quad (30)$$

DEFINICJA 23. *Zdarzeniami rozłącznymi (wykluczającymi się) A_1, \dots, A_n* (ang. *mutually exclusive events*) nazywamy zdarzenia, których iloczyn zdarzeń jest zdarzeniem niemożliwym

$$A_1 \cap A_2 \cap \dots \cap A_n = \emptyset. \quad (31)$$

W celach zrozumienia relacji zachodzących pomiędzy zdarzeniami, przedstawiane są diagramy, zwane od nazwiska autora, kołami Venna. Z diagramami Venna można się zapoznać m.in. w [bob86] [kow08].

Rozważając przestrzeń mierzalną jako (Ω, F) (patrz def. 18), prawdopodobieństwo definiowane jest następująco [wic08], [bob86]:

DEFINICJA 24. *Prawdopodobieństwem* (inaczej *miarą probabilistyczną*) (ang. *probability*) nazywamy dowolną rzeczywistą funkcję P spełniającą aksjomaty:

- 1) każdemu zdarzeniu losowemu $A \in F$ odpowiada liczba $P(A)$ taka, że $0 \leq P(A) \leq 1$,
- 2) $P(\Omega) = 1$ (prawdopodobieństwo zdarzenia pewnego),
- 3) jeśli $A_1, A_2, \dots, A_n \in F$ się parami wykluczają to $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$ (przeliczalna addytywność).

Wobec tego prawdopodobieństwo określone na ciele podzbiorów borelowskich pewnej przestrzeni liczb rzeczywistych R lub przestrzeni euklidesowej R^n dla liczby naturalnej n nazwane jest *rozkładem prawdopodobieństwa*. Natomiast jeżeli $A \in F$ jest zdarzeniem, to liczbę $P(A)$ określa się mianem *prawdopodobieństwa zdarzenia A*.

3.2. Prawdopodobieństwo zdarzeń rozmytych

Pojęcie zdarzenia rozmytego oraz definicję prawdopodobieństwa zdarzenia rozmytego zostały wprowadzone przez L.A. Zadeha w jego kolejnej pracy [zad68].

DEFINICJA 25. *Zdarzeniem rozmytym A* (ang. *fuzzy event*) nazywamy podzbiór rozmyty przestrzeni zdarzeń elementarnych \aleph ($A \in \aleph$) o funkcji przynależności $\mu_A(x)$ mierzalnej w sensie Borela.

DEFINICJA 26. *Prawdopodobieństwo* (nierożmyte) *zdarzenia rozmytego A* (ang. *probability of fuzzy event*), o funkcji przynależności $\mu_A(x)$, oznaczane jest jako $P(A)$ i definiowane następująco:

- a) dla dyskretnej przestrzeni rozważań $\aleph = \{x_1, x_2, \dots, x_n\}$:

$$P(A) = \sum_{i=1}^n p(x_i) \mu_A(x_i), \quad (32)$$

gdzie, $p(x_i) \in [0,1]$ stanowi prawdopodobieństwo (nierożmyte) zdarzenia elementarnego x_i , przy czym $\sum_{i=1}^n p(x_i) = 1$,

- b) dla ciągłej przestrzeni rozważań \aleph :

$$P(A) = \int_{\aleph} \mu_A(x) dP. \quad (33)$$

Ponadto, istnieje związek pomiędzy prawdopodobieństwem a mocą zbioru rozmytego (def. 8). W przypadku, gdy zdarzenia elementarne $x_i \in \aleph$ są jednakowo prawdopodobne to

$$P(A) = \frac{|A|}{|\aleph|} = \frac{\sum_{i=1}^n \mu_A(x_i)}{n}. \quad (34)$$

W pracach [kac86], [kac01], [pie03] wykazano, iż słowo "prawdopodobieństwo" do określenia $P(A)$ jest najwłaściwsze, z powodu spełniania właściwości klasycznego prawdopodobieństwa (def. 24). Zaproponowana przez Zadeha [zad68], definicja prawdopodobieństwa zdarzenia rozmytego jest najczęściej stosowaną w literaturze [kac01] i będzie także wykorzystywana w niniejszej dysertacji. Warto również przypomnieć pojęcie prawdopodobieństwa warunkowego zdarzenia rozmytego, które posłuży podczas tworzenia wag modelu probabilistyczno-rozmytej bazy wiedzy. Definicja ta jest analogiczna z definicją klasycznego prawdopodobieństwa warunkowego, którą można znaleźć w każdej literaturze przedmiotu, przykładowo w [kry97], [plu00].

DEFINICJA 27. *Prawdopodobieństwo warunkowe zdarzenia rozmytego A (ang. conditional probability of fuzzy set) w $\aleph = \{x_1, x_2, \dots, x_n\}$ pod warunkiem zajścia zdarzenia rozmytego B, oznaczane jest jako $P(A|B)$ i definiowane następująco [zad68], [kac86]:*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0, \quad (35)$$

przy czym, jeżeli zdarzenia rozmyte A i B są niezależne to zachodzi relacja:

$$P(A|B) = P(A). \quad (36)$$

Warunkowym rozkładem prawdopodobieństwa zdarzeń rozmytych nazywamy funkcję $P(A|B)$ argumentu A określona dla każdego $A \in \aleph$ określonego wzorem (35) (na podstawie [plu00]).

Istnieje również inne ujęcie prawdopodobieństwa - zaproponowane przez Yagera [yag79]. Definiuje on prawdopodobieństwo jako zbiór rozmyty na przedziale $\langle 0,1 \rangle$. Zainteresowanym tym podejściem proponuje się pozycje literaturowe tj. [kac01], [pie03].

4. Podstawy teorii rozmytych systemów wnioskujących

Rozmyte systemy wnioskujące stanowią systemy z bazami wiedzy (ang. knowledge-based systems), w których wykorzystane jest podejście lingwistyczne podczas modelowania i wnioskowania, zwanego również modelowaniem i wnioskowaniem rozmytym. Rozmyte systemy wnioskujące są również nazwane rozmytymi systemami ekspertowymi [kwi07], jako połączenie systemów ekspertowych i zmiennych lingwistycznych. W literaturze przedmiotu, do podstawowych prac, które zajmują się tą tematyką należą: prace Zadeha [zad73] [zad79], Mamdaniego i Assiliana [mam74] [mam75], Sugeno, Takagi oraz Kanga [tak85] [sug88], Yagera [yag80] [yag95], Dubois'a i Prade'a [dub98], Pedrycza [ped84][ped93], Kosko [kos92a,b], Janga [jan93] i wielu innych autorów.

Utworzenie rozmytego systemu wnioskującego wymaga znajomości podstawowej teorii dotyczącej rozmytych modeli wiedzy. Poniższy rozdział będzie stanowić przegląd istniejących rozwiązań w tej dziedzinie, dając także możliwość porównania tworzonego systemu ze znanyimi w literaturze.

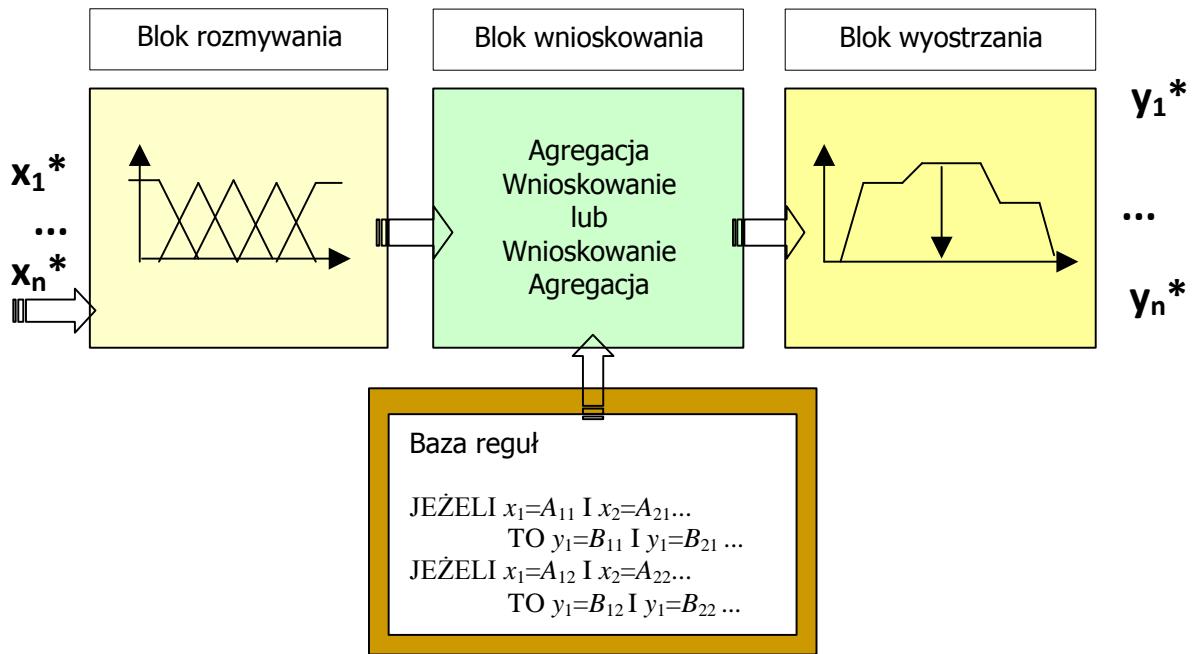
4.1. Struktura rozmytych systemów wnioskujących

W literaturze dotyczącej systemów rozmytych (m.in. [yag95], [rut97b], [kac01], [pie03], [łęs08], [now09]) wymienia się cztery podstawowe elementy struktury rozmytego systemu wnioskującego:

- blok rozmywania (ang. *fuzzification*),
- blok wnioskowania (ang. *inference*),
- baza wiedzy w postaci bazy reguł (ang. *knowledge base*),
- blok wyostrzania (ang. *defuzzification*).

Schemat systemu typu MIMO (ang. *Multiple Input, Multiple Output*), z wieloma wejściami (x_1, \dots, x_n) oraz wieloma wyjściami (y_1, \dots, y_n) o wartościach numerycznych, został przedstawiony na rysunku II-6.

Zagadnienia, dotyczące poszczególnych elementów, a w szczególności bazy wiedzy oraz metod jej pozyskania, zostaną omówione w kolejnych podrozdziałach.



Rys. II-6. Schemat rozmytego systemu wnioskującego (na podstawie [rut06], [łes08], [kac01])

4.2. Regułowa reprezentacja wiedzy – rozmyte modele wiedzy

Spośród wachlarza formalizmów reprezentacji wiedzy w dziedzinie systemów wnioskujących, najbliższą metodą zapisu wiedzy stosowaną przez człowieka jest regułowa reprezentacja wiedzy typu JEŻELI-TO. Zagadnienie to jest tematem szerszej dyskusji m.in. w [cic00], [nie00], [ban09] oraz treścią wielu badań [paj07] [nie06].

Zaczerpnięta z dziedziny logiki matematycznej, ogólna postać regułowej reprezentacji wiedzy, zwanej też *regułami wnioskowania* lub *regułami decyzji*, składa się z części warunkowej p_r , zwanej *przesłanką* bądź *poprzednikiem* reguły (ang. *antecedent, body*) oraz części decyzyjnej q_r , zwanej *konkluzją* bądź *następnikiem* reguły (ang. *consequent, head*). Zatem ogólna postać reguły ma formę:

$$\text{JEŻELI } p_r \text{ TO } q_r, \quad (37)$$

gdzie słowa JEŻELI (ang. *IF*), TO (ang. *THEN*) stanowią słowa kluczowe poprzedzające odpowiednio przesłankę i konkluzję reguły.

Szczegółowa postać bazy reguł kształtuje się w zależności od zastosowanego modelu wiedzy (zwanego też w skrócie modelem). Prostota zapisu oraz łatwość interpretacji i wnioskowania na podstawie reguł typu JEŻELI-TO wpłynęły na ogromną ich popularność zastosowania

w strukturach rozmytych, jak i neuronowo-rozmytych modeli wiedzy. Zaletą ich jest także możliwość opracowania modeli na bazie znacznie mniejszej ilości informacji o systemie [pie03] w porównaniu z modelami matematycznymi. Chęć uzyskania coraz to większej dokładności modeli dla różnorodnych systemów rzeczywistych, z różnym stopniem dostępności informacji i ich form, spowodowało intensywny rozwój struktur modeli wiedzy [łes08]. W dalszej części pracy przedstawione zostaną różne rodzaje rozmytych modeli wiedzy, opartych na regułach typu JEŻELI-TO: podstawowe tj. model Mamdaniego oraz Takagi-Sugeno-Kanga oraz ich modyfikacje, reprezentujące wiedzę w postaci reguł rozmytych z wagami.

4.2.1. Rozmyte modele lingwistyczne

W 1974 roku, pokazując jak można utworzyć model człowieka-regulatora sterującego obiektem [pie03], Mamdani wprowadził własną ideę modelowania [mam74]. Realizowany model pozwala na odwzorowanie wejścia modelu (X) na wyjście (Y): $X \rightarrow Y$ poprzez zbiór rozmytych reguł warunkowych (ang. *fuzzy conditional rules*) w postaci:

$$\text{JEŻELI } x = A_i \text{ TO } y = B_j, \quad (38)$$

gdzie x stanowi zmienną wejściową modelu, y – zmienną wyjściową, A_i, B_j – wartości zmiennych lingwistycznych, które są utożsamiane ze zbiorami rozmytymi, odpowiednio dla wejścia i dla wyjścia modelu.

W opisywanym modelu lingwistycznym stosuje się tzw. *podejście koniunkcyjne* [łes08], w którym reguły definiuje się jako relację rozmytą R_{ij} (def. 14) określona za pomocą iloczynu kartezjańskiego na przestrzeni $X \times Y$. Funkcja przynależności relacji rozmytej R_{ij} wyznaczana jest za pomocą operatora t-normy (def. 12):

$$\mu_{R_{ij}}(x, y) = \mu_{A_i \rightarrow B_j}(x, y) = T(\mu_{A_i}(x), \mu_{B_j}(y)). \quad (39)$$

W oryginalnym modelu Mamdaniego stosowanym operatorem t-normy jest minimum:

$$\mu_{R_{ij}}(x, y) = \mu_{A_i}(x) \wedge \mu_{B_j}(y) = \min(\mu_{A_i}(x), \mu_{B_j}(y)), \quad (40)$$

Reguły określone regułami Larsena stosują iloczyn arytmetyczny do definicji relacji rozmytej w przestrzeni $X \times Y$ jako:

$$\mu_{R_{ij}}(x, y) = \mu_{A_i}(x) \cdot \mu_{B_j}(y). \quad (41)$$

Przykłady innych, stosowanych operatorów t-normy przedstawiono w tabeli II-4. Należy podkreślić, iż reguły Mamdaniego nie są implikacjami w sensie logicznym [rut06], [łes08].

Jednakże, gdy relacja rozmyta R_{ij} określana jest za pomocą rozmytej implikacji $I(x, y)$ o funkcji przynależności:

$$\mu_{R_{ij}}(x, y) = \mu_{A_i \rightarrow B_j}(x, y) = \mu_I(x, y) = I(\mu_A(x), \mu_B(y)), \quad (42)$$

mówimy o podejściu do modelowania zwany *logiczną interpretacją reguł JEŻELI-TO* [łes08], [rut06]. W aksjomatycznej definicji Fodora [fod91], implikacja rozmyta jest definiowana jak poniżej.

DEFINICJA 28. *Implikacja rozmyta* (ang. *fuzzy implication*) stanowi funkcję I : $\langle 0,1 \rangle \times \langle 0,1 \rangle \rightarrow \langle 0,1 \rangle$ spełniające poniższe aksjomaty dla każdego $x, y, z \in \langle 0,1 \rangle$:

- 1) jeżeli $x \leq z$, to $I(x, y) \geq I(z, y)$ (nierosnąca funkcja dla pierwszego argumentu),
- 2) jeżeli $y \leq z$, to $I(x, y) \leq I(x, z)$ (niemalejąca funkcja dla drugiego argumentu),
- 3) $I(0, y) = 1$ (z fałszu może cokolwiek wyniknąć),
- 4) $I(x, 1) = 1$ (cokolwiek może prowadzić do prawdy),
- 5) $I(1, 0) = 0$ (z prawdy nigdy nie wynika fałsz).

W tablicy II-5 zostały przedstawione najczęściej stosowane w literaturze [kac01], [łęs08], [now09] operatory implikacji rozmytej.

Tab. II-5. Wybrane implikacje rozmyte [kac01], [łęs08], [now09]

Nazwa operatora	Wzór
Implikacja Łukasiewicza	$\mu_{A \rightarrow B}(x) = \text{MIN}(1, 1 - \mu_A(x) + \mu_B(x))$
Implikacja Fodora	$\mu_{A \rightarrow B}(x) = \begin{cases} 1 & \text{dla } \mu_A(x) \leq \mu_B(x) \\ \text{MAX}(1 - \mu_A(x), \mu_B(x)) & \text{dla } \mu_A(x) > \mu_B(x) \end{cases}$
Implikacja Goguena	$\mu_{A \rightarrow B}(x) = \text{MIN}(\mu_B(x) / \mu_A(x), 1)$
Implikacja Gödela	$\mu_{A \rightarrow B}(x) = \begin{cases} 1 & \text{dla } \mu_A(x) \leq \mu_B(x) \\ \mu_B(x) & \text{dla } \mu_A(x) > \mu_B(x) \end{cases}$
Implikacja Kleene-Dienesa	$\mu_{A \rightarrow B}(x) = \text{MAX}(1 - \mu_A(x), \mu_B(x))$
Implikacja Reichenbacha	$\mu_{A \rightarrow B}(x) = 1 - \mu_A(x) + \mu_A(x) \cdot \mu_B(x)$
Implikacja Yagera	$\mu_{A \rightarrow B}(x) = (\mu_A(x))^{\mu_B(x)}$
Implikacja Zadeha	$\mu_{A \rightarrow B}(x) = \text{MAX}(1 - \mu_A(x), \text{MIN}(\mu_A(x), \mu_B(x)))$

W [yag95], do logicznej interpretacji reguł (zwanej *podejściem destrukcyjnym*), zastosowano implikacje rozmyte spełniające warunek S-implikacji [łęs08]:

$$\mu_{R_j}(x, y) = \mu_{I_s}(x, y) = S(\mu_{\neg A_i}(x), \mu_{B_j}(y)) , \quad (43)$$

gdzie S stanowi operator s-normy (def. 13), natomiast $\neg A_i$ - negację zbioru rozmytego (def. 9).

Szczegółowe informacje na temat wnioskowania na podstawie reguł lingwistycznych zamieszczone są w podrozdziale II-3.4. Badanie nad operatorami (operatorem t-normy bądź implikacji rozmytej) oraz ich wyborem w procesie interpretacji reguł stanowi wciąż przedmiot wielu publikacji [cor97], [cor00], [bac03], [got03], [jay09].

Koniunkcyjna interpretacja reguł wymaga mniejszego nakładu obliczeniowego z uwagi na prostsze algorytmy wnioskowania. Jednakże, z punktu widzenia logiki matematycznej, bardziej uzasadnione jest użycie logicznej interpretacji reguł.

Model lingwistyczny jest jednym z częściej stosowanych modeli rozmytych jako regulatory rozmyte [dri96], [roj00], [pie03], [kan08].

4.2.2. Modele Takagi – Sugeno – Kanga

W pracach Takagi i Sugeno [tak85] oraz Sugeno i Kanga [sug88] został wprowadzony odmienny rozmyty model wiedzy zwany *modelem Takagi-Sugeno-Kanga* (w skrócie TSK). Baza wiedzy modelu TSK dla systemu typu MISO (ang. *Multiple Input, Single Output*) jest reprezentowana przez następujące reguły:

$$\text{JEŻELI } x_1 \text{ jest } A_1 \text{ I } \dots \text{ I } x_N \text{ jest } A_N \text{ TO } y = f(x_1, x_2, \dots, x_N), \quad (44)$$

gdzie: N jest liczbą wejść modelu, x_i – i -tą zmienną wejściową modelu ($i = 1, \dots, N$), y – zmienną wyjściową, A_i – wartością lingwistyczną, natomiast $f(x_1, x_2, \dots, x_N)$ – nieroźmyta funkcją wartości wejść. Dla modelu TSK pierwszego rzędu wyjście $y(x)$ ma postać:

$$y(\mathbf{x}) = f(x_1, x_2, \dots, x_N) = p_0 + p_1 x_1 + p_2 x_2 + \dots + p_N x_N = \mathbf{p}^T \mathbf{x}', \quad (45)$$

gdzie: \mathbf{p} oznacza $N+1$ wymiarowy wektor parametrów funkcji, a \mathbf{x}' wektor wyjściowy z dodatkowym elementem równym jeden: $\mathbf{x}' = \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix}$. W szczególnym przypadku, gdy $p_k=0$ ($k=1, \dots, N$) mamy do czynienia z modelem TSK zerowego rzędu.

Model TSK jest jednym z częściej stosowanych modeli wiedzy zarówno w teoretycznych pracach badawczych (np. [che98], [ahn10]), jak i praktycznych (np. [din03], [su10]). Doczekał się wielu rozwinięć samego modelu (np. [mas04], [cza09]), jak i sposobów pozyskiwania wiedzy dla parametrów modelu z zastosowaniem sieci neuronowo-rozmytych (np. [pal01]) oraz algorytmów genetycznych [rut97b]. Podstawowymi wadami modelu TSK jest niemożność zastosowania interpretacji rozmytych reguł warunkowych oraz operacji agregacji [łes08]. System będący tematem dysertacji jest pozbawiony powyższych wad.

4.2.3. Modele probabilistyczno-rozmyte

Działając w środowisku niepewnym z uwzględnieniem niepewności w sensie probabilistycznym bądź stochastycznym, możemy ująć stan rzeczywistości, za pomocą rozmytego systemu wnioskującego, który wykorzystuje empiryczny rozkład prawdopodobieństwa. W dysertacji ten model wiedzy zwany jest *modelem probabilistyczno-rozmytym*, a w literaturze spotykany jest jako *rozmyty model* bądź *rozmyta reprezentacja wiedzy z miarami prawdopodobieństwa zbiorów rozmytych* (ang. *fuzzy model using probability measures of fuzzy events, fuzzy knowledge representation using probability measures of fuzzy events*) [wb07] [wb08a]. Charakterystyczną cechą bazy wiedzy omawianego modelu dla systemu typu MISO jest reprezentacja wiedzy w postaci reguł plikowych, stanowiących zbiór J reguł elementarnych, w formie [wb07]:

$$\begin{aligned}
 R(m) : (w_m) & \quad \text{JEŻELI } x_1 = A_m^1 \text{ I } \dots \text{ I } x_N = A_m^N \\
 & \quad \text{TO } y = B_{1/m} (w_{1/m}) \\
 & \quad \dots \\
 & \quad \text{TAKŻE } y = B_{j/m} (w_{j/m}) \\
 & \quad \dots \\
 & \quad \text{TAKŻE } y = B_{J/m} (w_{J/m}),
 \end{aligned} \tag{46}$$

gdzie:

N jest liczbą zmiennych wejściowych modelu,

M – liczbą reguł plikowych,

x_1, \dots, x_N stanowi zmienne wejściowe modelu, $x_n \in X_n \subset \mathbb{R}$, $n=1, \dots, N$,

y – zmienną wyjściową modelu, $y \in Y \subset \mathbb{R}$,

A_m^n jest m -tą wartość lingwistyczna n -tej zmiennej wejściowej, $m=1, \dots, M$, $n=1, \dots, N$,

$B_{j/m}$ – wartością lingwistyczną zmiennej wyjściowej w j -tej regule elementarnej m -tej reguły plikowej, $m=1, \dots, M$, $j=1, \dots, J$.

Wagi w_m , $w_{j/m}$ stanowią miary prawdopodobieństw zbiorów rozmytych. Waga m -tej reguły plikowej w_m reprezentuje łączne prawdopodobieństwo zdarzeń rozmytych w przesłance reguły plikowej:

$$P(\mathbf{x} = A_m) = P((x_1 = A_m^1) \cap \dots \cap (x_N = A_m^N)), \tag{47}$$

gdzie $\mathbf{x} = [x_1, x_2, \dots, x_N]$.

Natomiast, waga elementarnej reguły $w_{j/m}$ reprezentuje warunkowe prawdopodobieństwo zdarzeń rozmytych (por. def. 27), które obliczane jest według wzoru:

$$P((y = B_{j/m}) / (\mathbf{x} = A_m)) = \frac{P((y = B_{j/m}) \cap (\mathbf{x} = A_m))}{P(\mathbf{x} = A_m)}. \tag{48}$$

W ujęciu klasycznym Zadeha, prawdopodobieństwo obliczane jest jako (32) dla przestrzeni dyskretnej oraz (33) dla przestrzeni ciągłe. Wg [wb05], [wb08a] wagi obliczane są dla wartości lingwistycznych A_m^n , $B_{j/m}$ określonych na niepustych przestrzeniach X_n i Y dyskretyzowanych do K rozłącznych przestrzeni wartości zmiennych, oznaczonych odpowiednio a_k^n i b_k :

$$A_m^n = \sum_{k=1, \dots, K} \mu_{A_m^n}(a_k^n) / a_k^n, \quad B_{j/m} = \sum_{k=1, \dots, K} \mu_{B_{j/m}}(b_k) / b_k, \tag{49}$$

z zachowaniem zasady podziału do jedności np. $\sum_{m=1, \dots, M} \mu_{A_m^n}(a_k^n) = 1$, dla $\forall a_k^n \in X_n$. Toteż

korzystając z definicji przecięcia zbiorów rozmytych (def. 10) opartej na operatorze t-normy (def. 12) oraz własności (47-48), waga w_m definiowana jest jako [wb07]:

$$P(\mathbf{x} = A_m) = \sum_{a_{k_1}^1, \dots, a_{k_N}^N \in X_1 \times \dots \times X_N} [p_{k_1 \dots k_N}(x_1, \dots, x_N) \cdot T(\mu_{A_m^1}(a_{k_1}^1), \dots, \mu_{A_m^N}(a_{k_N}^N))], \tag{50}$$

natomiast waga $w_{j/m}$ jako:

$$P((y = B_{j/m}) / (\mathbf{x} = A_m)) = \frac{\sum_{\substack{a_{k_1}^1, \dots, a_{k_N}^N, b_{k_{N+1}} \\ \in X_1 \times \dots \times X_N \times Y}} [p_{k_1 \dots k_N k_{N+1}}(x_1, \dots, x_N, y) \cdot T(\mu_{A_m^1}(a_{k_1}^1), \dots, \mu_{A_m^N}(a_{k_N}^N), \mu_{B_m}(b_{k_{N+1}}))]}{\sum_{\substack{a_{k_1}^1, \dots, a_{k_N}^N \\ \in X_1 \times \dots \times X_N}} [p_{k_1 \dots k_N}(x_1, \dots, x_N) \cdot T(\mu_{A_m^1}(a_{k_1}^1), \dots, \mu_{A_m^N}(a_{k_N}^N))]}, \quad (51)$$

gdzie $p_{k_1 \dots k_N k_{N+1}}(x_1, \dots, x_N, y)$ ($p_{k_1 \dots k_N}(x_1, \dots, x_N)$) stanowi prawdopodobieństwo w sensie definicji 22, jako prawdopodobieństwo zdarzenia, że $x_1 \in a_{k_1}^1$ oraz $x_2 \in a_{k_2}^2$ oraz ... oraz $y \in b_{k_{N+1}}$.

Ponieważ, tematem niniejszej dysertacji jest koncepcja systemu wnioskującego w oparciu o model probabilistyczno-rozmyty bazy wiedzy, omawiane rozważania będą kontynuowane w rozdziale III pracy.

4.2.4. Modele relacyjne

Rozpatrując strukturę rozmytych modeli wiedzy, natrafiamy na grupę modeli lingwistycznych (por. (38)) zwanych *rozmytymi modelami relacyjnymi* (ang. *fuzzy relational models*), których reguły, podobnie jak w modelu probabilistyczno-rozmytym, zawierają odpowiednie współczynniki (wagi):

$$\begin{aligned} R(m) : \text{JEŻELI } x_1 = A_m^1 \text{ I } \dots \text{ I } x_N = A_m^N \text{ TO } y &= B_{1/m} \ (w'_{1/m}) \\ y &= B_{j/m} \ (w'_{j/m}) \\ y &= B_{J/m} \ (w'_{J/m}), \end{aligned} \quad (52)$$

gdzie: oznaczenia symboli są analogiczne do wzoru (46), natomiast $w'_{j/m}$ stanowią wagi poszczególnej reguły określane jako współczynniki ufności.

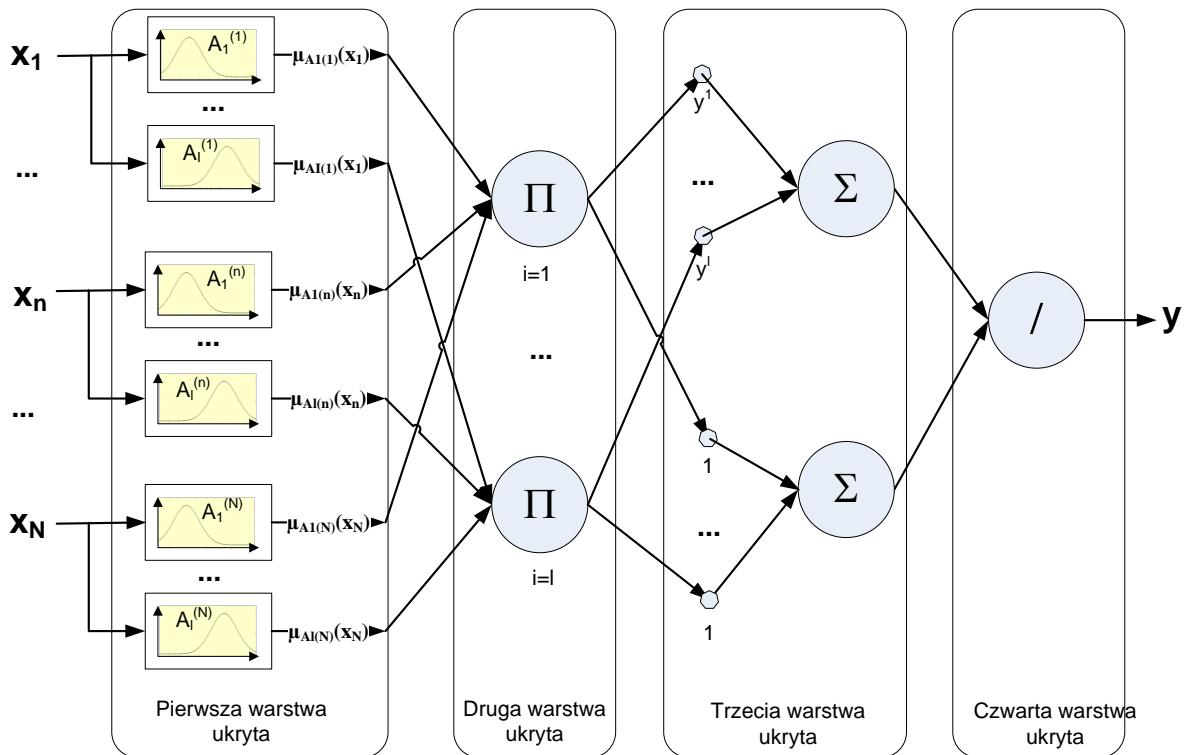
Należy wobec tego zaznaczyć, na czym polega odmienność modeli relacyjnych w porównaniu z modelem probabilistyczno-rozmytym (46).

We wprowadzonych przez Pedrycza [ped84] modelach relacyjnych, podobnie jak w tworzonym modelu wiedzy, poszczególne reguły nie są całkowicie prawdziwe, lecz jedynie częściowo. Prawdziwość reguł wynika jednakże z odmiennych uwarunkowań. W tworzonym modelu prawdziwość reguł jest związana z częstością występowania zależności typu JEŻELI-TO zdarzeń rozmytych. Reguła jest tym lepiej dopasowana, im częstotliwość występowania danej zależności jest większa. Natomiast, w modelach relacyjnych prawdziwość reguł wynika z tzw. współczynnika ufności $w'_{j/m}$, obliczanego na podstawie teorii równań relacyjnych ([ped84], [kac86]), bądź też na podstawie pomiarów doświadczalnych z użyciem rozmytych sieci neuronowych [pie03]. Toteż analizowany współczynnik nie posiada interpretacji opartej na częstości zbiorów rozmytych.

4.2.5. Modele neuronowo-rozmyte

Aby wykorzystać potencjał obliczeniowy tkwiący w sieciach neuronowych, powstały systemy zwane *systemami neuronowo-rozmytymi* (ang. *neuro-fuzzy systems*) o architekturze sieci neuronowej lub do niej podobnej, przetwarzające sygnały rozmyte. Powyższa architektura, mająca zdolność do uczenia się i uogólniania wiedzy, jest stosowana do automatycznego pozyskiwania parametrów reguł rozmytego modelu wiedzy w postaci JEŻELI-TO, bądź też do całościowego wydobywania reguł warunkowych.

Przykładowa architektura systemu neuronowo-rozmytego ([rut97b], [duc00]), w ramach pierwszego z wyżej wymienionych podejść, jest pokazana na rysunku II-7.



Rys. II-7. Struktura systemu neuronowo-rozmytego (na podstawie [duc00])

Struktura systemu neuronowo-rozmytego (rys. II-7) bazuje na regułach w postaci:

$$\text{JEŻELI } x_1 = A_i^{(1)} \text{ I } \dots \text{ I } x_N = A_i^{(N)} \text{ TO } y = B_i, i=1, \dots, l, \quad (53)$$

stosuje wnioskowanie zgodnie z uogólnioną regułą *modus ponens* (rozdz. II-3.4) oraz wnioskowanie metodą średniej środków CA (tab. II-6). Architektura neuronowa pozwala, na podstawie ciągu uczącego, dopasować parametry zbiorów rozmytych znajdujących się w przesłance reguły oraz wartości parametrów y^i (środków zbioru rozmytego B_i). Najczęściej stosowaną metodą uczenia jest algorytm wstępnej propagacji błędów ([oso96], [duc00], [rut97b]). Więcej informacji na temat dopasowania modelu neuronowo-rozmytego zamieszczono w rozdziale, dotyczącym metod automatycznego pozyskiwania baz wiedzy w systemach rozmytych.

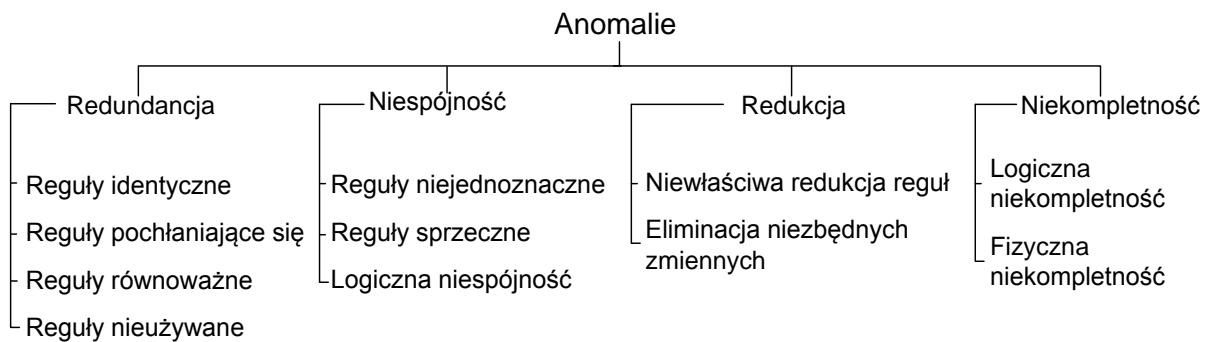
W 1993 roku Jang zaproponował system wnioskowania rozmytego, wykorzystujący sieć adaptacyjną (ANFIS, ang. *Adaptive-Network-Based Fuzzy Inference System*) [jan93].

W odróżnieniu od przedstawionego wcześniej systemu neuronowo-rozmytego, system Janga został oparty o bazę wiedzy w postaci rozmytych reguł TSK (44). ANFIS, jako jeden z pierwszych systemów neuronowo-rozmytych, został zaimplementowany w popularnym środowisku obliczeniowym Matlab. Z tego też względu, zaznacza się dużą popularnością stosowania oraz doczekał się wielu praktycznych i ciekawych zastosowań m.in. [jan97], [kor94], [lu05].

Systemy neuronowo-rozmytych w ostatnich latach były tematem badań i analiz w wielu pozycjach literaturowych, m.in. [lęs08] [jan97], [buc99], [cpa09].

4.3. Właściwości regulowej, rozmytej bazy wiedzy

Zaprojektowanie bazy wiedzy stanowi jeden z punktów krytycznych tworzenia systemów rozmytych. Różni eksperci mają prawo do przekazywania swojej wiedzy w różny, indywidualnie arbitralny sposób [ban09], natomiast projektant bazy wiedzy musi zagwarantować brak anomalii w regułach. Jednakże istnieje wiele anomalii, jakie mogą istnieć w odniesieniu do regułowej reprezentacji wiedzy. Rysunek II-8 przedstawia opis przypadków, jakie zostały uwzględnione m.in. w pozycji [lig06].



Rys. II-8. Typy anomalii regułowej reprezentacji wiedzy (na podstawie [lig08],[paj07])

Aby uniknąć błędów przy tworzeniu rozmytych baz wiedzy typu JEŻELI-TO, należy przeanalizować bazę pod kątem cech świadczących o jej prawidłowości. W literaturze dotyczącej systemów rozmytych ([pie03], [dri93], [czo85], [jin03]) najczęściej wymienia się następujące, wymagane, właściwości reguł rozmytego modelu wiedzy:

- kompletność (zupełność) modelu wiedzy,
- niesprzeczność (zgodność) bazy reguł,
- ciągłość bazy reguł,
- brak redundancji (nadmiarowości) w bazie reguł.

Rozważmy zbiór następujących reguł elementarnych, stanowiący podstawę rozmytego modelu wiedzy dla systemu typu MISO:

$$\begin{aligned}
 R(1): & \text{JEŻELI } x_1 = A_1^1 \text{ I } \dots \text{ I } x_N = A_N^1 \text{ TO } y = B^1 \\
 & \dots \\
 R(M): & \text{JEŻELI } x_1 = A_1^M \text{ I } \dots \text{ I } x_N = A_N^M \text{ TO } y = B^M. \tag{54}
 \end{aligned}$$

DEFINICJA 29. Rozmyty model wiedzy jest *kompletny* (zupełny, ang. *completeness*) jeżeli dla każdego wektora wejść $\mathbf{x}^* = [x_1^*, \dots, x_N^*]$ o wartościach ostrych można przyporządkować właściwą wartość ostrą na wyjściu modelu y^* .

Należy nadmienić, iż definicja kompletności modelu wiedzy stanowi o jego kompletności numerycznej – nie oznacza jego kompletności lingwistycznej [pie03]. Najczęszym powodem niekompletności modelu rozmytego jest niekompletność podziału rozmytego przestrzeni wejść $X_1 \times \dots \times X_N$. Formalnie oznacza to, że dla co najmniej jednej zmiennej $x_n \in X_n$ ($n = 1, \dots, N$) zachodzi zależność:

$$\sum_{m=1}^M \mu_{A_n^m}(x_n^*) = 0, \quad x_n^* \in X_n, \quad (55)$$

gdzie: M oznacza ilość zbiorów rozmytych przyjętych dla x_n . Nie zawsze jednak kompletność podziału rozmytego gwarantuje kompletność modelu wiedzy [jin03].

Należy zauważyć, że często bywają sytuacje np. w praktycznych zastosowaniach regulatorów, iż celowo regułowa baza wiedzy jest niekompletna, z uwagi na rejony dziedziny nie znajdujące się w zakresie zastosowań [dri93]. Model wiedzy jest wówczas wiarygodny jedynie w strefie zdefiniowanej przez jego reguły [pie03].

DEFINICJA 30. Rozmyty model wiedzy posiada *sprzeczną* (ang. *conflict rules*) bazę wiedzy, jeżeli istnieją w niej dwie reguły o takich samych przesłankach, ale różnych konkluzjach.

Powyzsza definicja sprzecznej bazy wiedzy jest właściwa dla reguł typu JEŻELI-TO w postaci (38). Jest natomiast niewłaściwa dla reguł w probabilistyczno-rozmytej bazie wiedzy (46), gdzie występują reguły z tą samą przesłanką i różnymi konkluzjami. Reguły te nie są regułami sprzecznymi, a jedynie regułami z inną częstością występowania.

Sprzeczność bazy wiedzy najczęściej wynika z błędów w tworzeniu reguł. Może być także spowodowana niedeterministością samego modelu.

DEFINICJA 31. Rozmyty model wiedzy posiada *ciągłą* (ang. *continuous*) bazę wiedzy, jeżeli nie ma w niej sąsiednich reguł $R(j), R(k)$ ze zbiorami rozmytymi w konkluzji B^j, B^k , których iloczyn jest zbiorem pustym ($B^j \cap B^k = \emptyset$).

Ciągłość bazy wiedzy jest pożądana [per06] z uwagi na gładkość odwzorowania modelu. Jednakże obiekty o dużych skokach wartości wyjść nie da się zamodelować za pomocą ciągłej bazy wiedzy.

Ostatnią, z wymienionych, pożądaną właściwością bazy wiedzy w modelach rozmytych jest brak nadmiarowości bazy reguł. Oznacza to unikanie sytuacji, w których w bazie reguł mamy więcej niż jedną regułę o tej samej przesłance i identycznej konkluzji. Najczęściej nie tylko unikamy nadmiarowości, ale dążymy do jak najprostszego modelu wiedzy. Mniejsza liczba reguł pozwala na łatwiejszą interpretację systemu przez człowieka.

4.4. Procedury wnioskowania rozmytego na podstawie bazy wiedzy

Wnioskowanie rozmyte (zwane też często *wnioskowaniem przybliżonym* lub *przybliżonym rozumowaniem* [kac86], ang. *fuzzy inference*) realizowane w oparciu o bazę wiedzy odnosi się do sposobu określania rozmytego zbioru wyjściowego B' zdefiniowanego na przestrzeni \mathfrak{I} na podstawie odpowiednich procedur transformacji oraz wejściowego zbioru rozmytego A' zdefiniowanego na przestrzeni \mathfrak{N} . Teorię na temat wnioskowania rozmytego zapoczątkował Zadeh w 1979 roku [zad79], odzwierciedla ona sposób rozumowania człowieka z wykorzystaniem nieprecyzyjnych stwierdzeń w języku naturalnym.

W logice klasycznej istnieją cztery tautologie rachunku zdań opisujące sposoby wnioskowania¹:

- sposób potwierdzający przez potwierdzenie (łac. *modus ponendo ponens*),

$$((p_r \rightarrow q_r) \wedge p_r) \rightarrow q_r, \quad (56)$$

- sposób zaprzeczający przez zaprzeczenie (łac. *modus tollendo tollens*),

$$((p_r \rightarrow q_r) \wedge \neg q_r) \rightarrow \neg p_r, \quad (57)$$

- sposób potwierdzający przez zaprzeczenie (łac. *modus tollendo ponens*),

$$((\neg p_r \rightarrow q_r) \wedge \neg q_r) \rightarrow p_r, \quad (58)$$

- sposób zaprzeczający przez potwierdzenie (łac. *modus ponendo tollens*),

$$((p_r \rightarrow q_r) \wedge \neg q_r) \rightarrow \neg p_r. \quad (59)$$

Zadeh wprowadził tzw. *uogólnioną regułę wnioskowania* [zad73], która pozwala na wysuwanie wniosków w oparciu o rozmyte sformułowania w przesłankach i konkluzjach. W systemach rozmytych stosuje się najczęściej dwa z wymienionych tautologii rachunku zdań ([rut06]) *modus ponendo ponens* oraz *modus tollendo tollens*, które po rozszerzeniu na zbiory rozmyte określa się jako *uogólnioną regułę wnioskowania modus ponendo ponens* (ang. *generalized modus ponendo ponens*) oraz *uogólnioną regułę wnioskowania modus tollendo tollens* (ang. *generalized modus tollendo tollens*) [zad73].

Uogólniona reguła wnioskowania modus ponendo ponens

Zgodnie z regułą (56) schemat wnioskowania z użyciem zbiorów rozmytych dla systemu MISO można przedstawić następująco [rut97b]:

Fakt	$x_1 = A'_1 \text{ I } \dots \text{ I } x_N = A'_N$
Reguła	JEŻELI $x_1 = A'_1 \text{ I } \dots \text{ I } x_N = A'_N$ TO $y = B$
Wniosek	$y = B'$

(60)

¹ Użyte oznaczenia odnoszą się do wzoru (37)

gdzie zbiór rozmyty A'_1 , określony na przestrzeni \aleph_1 , jest wartością lingwistyczną zmiennej lingwistycznej x_1 , i analogiczne dla A'_2, \dots, A'_N, B' , ponadto zbiory rozmyte A'_1, \dots, A'_N, B' są w pewnym sensie zbliżone [rut06] do zbiorów rozmytych A_1, \dots, A_N, B .

W powyższym schemacie mamy do czynienia z regułami złożeniowymi o przesłankach prostych połączonych ze sobą spójnikiem logicznym I w przesłance reguły. Wówczas, aby ocenić przesłankę tej reguły, a dokładniej określić stopień jej spełnienia [pie03] (zwany często *stopniem aktywacji* lub *zapłonem* τ), należy zastosować operator przecięcia zbiorów (def. 10) np. w postaci operatora t-normy (def. 12):

$$\mu_{A'}(x) = \mu_{A'_1 \cap \dots \cap A'_N}(x) = \sup_{x_1, \dots, x_N \in \aleph_1 \times \dots \times \aleph_N} T(\mu_{A'_1}(x_1), \dots, \mu_{A'_N}(x_N)). \quad (61)$$

W przypadku reguł o przesłankach prostych połączonych ze sobą spójnikami logicznymi LUB należy wykorzystać operatory sumowania zbiorów (def. 11,13).

Formalnie zbiór rozmyty B' określony jest, poprzez złożenie zbioru rozmytego $A' = A'_1 \times \dots \times A'_N$ występującego w przesłance (określonego na przestrzeni \aleph) oraz relacji R dla reguły $A \rightarrow B$ ($A = A_1 \times \dots \times A_N$), następująco

$$B' = A' \circ (A \rightarrow B), \quad (62)$$

gdzie \circ oznacza operację złożeniową reguły wnioskowania wprowadzonej przez Zadeha.

Złożeniową regułę wnioskowania można realizować w kolejnych etapach [łes08]:

- cylindryczne rozszerzenie (def. 15) zbioru rozmytego A' na przestrzeń $\aleph \times \mathfrak{I}$:

$$\forall_{(x,y) \in \aleph \times \mathfrak{I}} \mu_{Ce(A')}(x, y) = \mu_{A'}(x), \quad (63)$$

- operacja przecięcia zbioru rozmytego $Ce(A')$ i relacji rozmytej R z wykorzystaniem operatora t-normy:

$$\forall_{(x,y) \in \aleph \times \mathfrak{I}} \mu_{Ce(A') \cap R}(x, y) = T(\mu_{A'}(x), \mu_R(x, y)), \quad (64)$$

- operacja projekcji (def. 16) przecięcia zbioru rozmytego $Ce(A')$ i relacji rozmytej R na przestrzeń \mathfrak{I} :

$$\forall_{y \in \mathfrak{I}} \mu_{B'}(y) = \forall_{y \in \mathfrak{I}} \mu_{proj_{\mathfrak{I}}(Ce(A') \cap R)}(y) = \sup_{x \in \aleph} [T(\mu_{A'}(x), \mu_R(x, y))]. \quad (65)$$

Relacja rozmyta R może być realizowana jako minimum (40), iloczyn algebraiczny (41), t-norma (39) czy implikacja rozmyta (42). W zależności od zastosowanego podejścia stosujemy odmienne nazewnictwo reguł i ich interpretacji (patrz podrozdział II-3.2.1).

Istnieją dwie metody wnioskowania rozmytego na podstawie bazy wiedzy. Pierwsza metoda – *metoda wnioskowania opartego na złożeniu* (w skrócie AW, zwana także podejściem globalnym [duc00] [now09], ang. *FATI – First Aggregate Then Infer*) stosuje uprzednią agregację reguł, utożsamianych z relacjami rozmytymi, a później dopiero właściwe wnioskowanie. Druga metoda – metoda *wnioskowania opartego na pojedynczych regułach* (w skrócie WA, zwana także podejściem lokalnym [duc00] [now09], ang. *FITA – First Infer Then Aggregate*) stosuje najpierw wnioskowanie, a kolejno agregacje reguł [łes08] [duc00] [dria93].

Powyższe rozumowanie, dotyczące *uogólnionej reguły wnioskowania modus ponendo ponens*, stanowi wnioskowanie typu WA. Aby otrzymać wynikowy zbiór rozmyty \tilde{B}' na bazie zbioru M reguł, należy utworzyć operację agregacji:

$$\tilde{B}' = \bigoplus_{m=1}^M B'^m = \bigoplus_{m=1}^M [A' \circ (A^m \rightarrow B^m)], \quad (66)$$

gdzie \oplus stanowi operator agregacji, którym może być operator t-normy, s-normy, bądź inny dowolny operator średniej [łęs08].

W przypadku wnioskowania przybliżonego typu AW zbiór rozmyty \tilde{B}' obliczany jest jako:

$$\tilde{B}' = A' \circ \left[\bigoplus_{m=1}^M (A^m \rightarrow B^m) \right]. \quad (67)$$

Uogólniona reguła wnioskowania modus tollendo tollens

Zgodnie z *uogólnioną regułą wnioskowania modus tollendo tollens* schemat wnioskowania z użyciem zbiorów rozmytych dla systemu SISO można przedstawić następująco [łęs08]:

Fakt	$y = B'$
Reguła	JEŻELI $x = A$ TO $y = B$
Wniosek	$x = A'$

(68)

gdzie zbiór rozmyty B' , określony na przestrzeni \mathfrak{J} , jest wartością lingwistyczną zmiennej lingwistycznej y , zbiór rozmyty A' , określony na przestrzeni \mathfrak{N} , jest wartością lingwistyczną zmiennej lingwistycznej x , ponadto zbiory rozmyte A', B' są w pewnym sensie zbliżone [rut06] do zbiorów rozmytych A, B .

Zbiór A' wyznacza się stosując złożeniową regułę wnioskowania według wzoru:

$$A' = (A \rightarrow B) \circ B', \text{ czyli} \quad (69)$$

$$\forall_{x \in \mathfrak{N}} \mu_{A'}(x) = \sup_{y \in \mathfrak{J}} [T(\mu_A(x), \mu_B(y))]. \quad (70)$$

Dla zbioru rozmytego \tilde{A}' , otrzymanego w oparciu o zbiór M reguł, stosujemy operację agregacji (rozumowanie przybliżone typu WA):

$$\tilde{A}' = \bigoplus_{m=1}^M A'^m = \bigoplus_{m=1}^M [(A^m \rightarrow B^m) \circ B']. \quad (71)$$

Natomiast, w przypadku rozumowania przybliżonego typu AW, zbiór rozmyty \tilde{A}' obliczany jest jako:

$$\tilde{A}' = \bigoplus_{m=1}^M [(A^m \rightarrow B^m)] \circ B'. \quad (72)$$

Wnioskowanie w oparciu o model TSK

Z uwagi na inną strukturę reguł (44), w porównaniu z lingwistycznym modelem wiedzy, wnioskowanie w oparciu o model TSK (patrz rozdz. 3.2.2) zachodzi w odmienny sposób. Całkowite wyjście systemu y^* dla modelu TSK jest wyznaczane jako średnia ważona wartości każdego wyjścia $\mathbf{x} = [x_1, x_2, \dots, x_N]$ następująco²:

$$y^* = \frac{\sum_{j=1}^J \tau^{(j)}(\mathbf{x}) y^{(j)}(\mathbf{x})}{\sum_{j=1}^J \tau^{(j)}(\mathbf{x})} = \frac{\sum_{j=1}^J \tau^{(j)}(\mathbf{x}) \mathbf{p}^{(j)T} \mathbf{x}'}{\sum_{j=1}^J \tau^{(j)}(\mathbf{x})}, \quad (73)$$

gdzie: J stanowi liczbę reguł w bazie wiedzy, natomiast $\tau^{(j)}(\mathbf{x})$ określa stopień aktywacji j -tej reguły ($j=1, \dots, J$) wyznaczany przez odpowiedni operator t-normy wg wzoru:

$$\tau^{(j)}(\mathbf{x}) = T(\mu_{A_1^{(j)}}(x_1), \mu_{A_2^{(j)}}(x_2), \dots, \mu_{A_N^{(j)}}(x_N)). \quad (74)$$

Wnioskowanie w oparciu o model neuronowo-rozmyty

Struktura modelu neuronowo-rozmytego ma bezpośredni związek z wybraną metodą wnioskowania przybliżonego. Przykładowo wnioskowanie w systemie neuronowo-rozmytym, przedstawionym na rysunku II-7, przebiega według poniżej opisanych etapów. Każdy z etapów realizowany jest przez odpowiednią warstwę struktury neuronowej lub do niej podobnej. Pierwsza warstwa struktury odtwarza funkcje przynależności zbiorów rozmytych $\mu_{A_i^{(n)}}(x_n)$, $n = 1, \dots, N$ znajdujących się w przesłankach reguł. Druga warstwa wykonuje operację przecięcia zbiorów rozmytych z wykorzystaniem iloczynu algebraicznego jako odzwierciedlenie iloczynu logicznego I (AND). Wykonywana jest również koniunkcyjna interpretacja reguł z wykorzystaniem implikacji Larsena określonej wzorem (41) wraz z wnioskowaniem zgodnym z uogólniona regułą *modus ponens*. Kolejne warstwy (trzecia i czwarta) wykonują operacje związane z wyostrzaniem wynikowej funkcji przynależności – w tym przypadku za pomocą metody średniej środka CA [wie09] (tab. II-6). Całość architektury przedstawia w sposób blokowy wnioskowanie z uwzględnieniem następującego wzoru:

$$y^* = \frac{\sum_{i=1}^I y^i \prod_{n=1}^N \mu_{A_i^{(n)}}(x_i)}{\sum_{i=1}^I \prod_{n=1}^N \mu_{A_i^{(n)}}(x_i)}, \quad (75)$$

gdzie: y^i stanowi punkt, w którym funkcja przynależności $\mu_{B^i}(y)$ do zbioru rozmytego B^i w konkluzji reguły (53) przyjmuje wartość maksymalną, i stanowi numer reguły modelu ($i=1, \dots, I$), natomiast n – numer zmiennej wejściowej ($n=1, \dots, N$).

² Użyte oznaczenia odnoszą się do wzorów (44)-(45)

4.5. Rozmywanie i wyostrzanie

Systemy z logiką rozmytą operują na zbiorach rozmytych [rut06]. Aby zamienić wartości ostre (numeryczne) $\mathbf{x}^* = [x_1^*, \dots, x_N^*] \in \aleph$ na wejściu do systemu w zbiór rozmyty A' , definiowany na przestrzeni \aleph , dokonywana jest operacja *rozmywania* (inaczej *fuzzyfikacji*, ang. *fuzzification*) stanowiąca odwzorowanie $F : \aleph \rightarrow F_{mf}(\aleph)$.

Zagadnienie rozmywania jest jednym z kluczowych elementów tworzonego systemu. W dotychczasowych rozwiązaniach w zakresie sterowania najczęściej stosuje się operacje rozmywania typu singleton [rut06]:

$$F : \aleph \rightarrow \left\{ \mu_{A'}(\mathbf{x}) \mid \mathbf{x} \in \aleph, \mu_{A'}(\mathbf{x}) = \begin{cases} 1, & \text{dla } \mathbf{x} = \mathbf{x}^* \\ 0, & \text{dla } \mathbf{x} \neq \mathbf{x}^* \end{cases} \right\}. \quad (76)$$

Rzadziej stosuje się inne funkcje przynależności. Wówczas formalny zapis operacji rozmywania można przedstawić jako [łes08]:

$$F : \aleph \rightarrow \{\mu_{A'}(\mathbf{x}) \mid \mathbf{x} \in \aleph, \mu_{A'}(\mathbf{x}) \in \langle 0, 1 \rangle\}, \quad (77)$$

gdzie $\mu_{A'}(\mathbf{x})$ stanowi funkcję przynależności, definiowaną jako jeden z wariantów zamieszczonych w tabelach II-1, II-3.

Gdy blok wnioskowania daje w wyniku zbiór B' lub zbiory rozmyte \tilde{B}' , należy dokonać operacji odwrotnej – operacji *wyostrzania* (zwanej też *defuzzyfikacją*, ang. *defuzzification*). Celem tej operacji jest uzyskanie wartości numerycznej $y^* \in \mathfrak{I}$ na wyjściu modelu. Metod wyostrzania jest wiele, każda z nich daje w wyniku inną wartość numeryczną. Jednakże, metoda powinna spełniać odpowiednie kryteria w zależności od zastosowania systemu [lee99] oraz użytych reguł rozmytych. Metody, obecnie najczęściej spotykane w literaturze, zawiera tabela II-6.

Tab. II-6. Przykłady metod wyostrzania [lee99], [pie03], [bro06], [łes08], [now09], [duc00]

Nazwa	Oznaczenie	Wzór	Opis oznaczeń we wzorze
Metoda środkowa ciężkości zwana również jako Metoda środkowa obszaru	COG (ang. <i>Center Of Gravity</i>) lub COA (ang. <i>Center Of Area</i>)	$y^* = \frac{\int_{\mathfrak{I}} y \mu_{B'}(y) dy}{\int_{\mathfrak{I}} \mu_{B'}(y) dy}$	y^* – wartość numeryczna na wyjściu modelu $y^* \in \mathfrak{I}$, y – zmienna wyjściowa modelu, B' – zbiór rozmyty
Metoda indeksowanego środka ciężkości	ICOG _{α} (ang. <i>Indexed COG</i>)	$y^* = \frac{\int_{\mathfrak{I}_\alpha} y \mu_{B'}(y) dy}{\int_{\mathfrak{I}_\alpha} \mu_{B'}(y) dy},$ $\mathfrak{I}_\alpha = \{y \in \mathfrak{I} \mid \mu_{B'}(y) \geq \alpha\}$	y^* – wartość numeryczna na wyjściu modelu $y^* \in \mathfrak{I}$, y – zmienna wyjściowa modelu, B' – zbiór rozmyty, α – parametr o wartości z przedziału $\langle 0, 1 \rangle$

Tab. II-6. c.d. Przykłady metod wyostrzania [lee99], [pie03], [bro06], [łes08], [now09], [duc00]

Nazwa	Oznaczenie	Wzór	Opis oznaczeń we wzorze
Modyfikowana metoda indeksowanego środka ciężkości	MICOG_α (ang. <i>Modified Indexed COG</i>)	$y^* = \frac{\int_{\mathfrak{I}_\alpha} y[\mu_{B'}(y) - \alpha]dy}{\int_{\mathfrak{I}_\alpha} [\mu_{B'}(y) - \alpha]dy},$ $\mathfrak{I}_\alpha = \{y \in \mathfrak{I} \mid \mu_{B'}(y) \geq \alpha\}$	y^* – wartość numeryczna na wyjściu modelu $y^* \in \mathfrak{I}$, y – zmienna wyjściowa modelu, B' – zbiór rozmyty, α – parametr o wartości z przedziału $\langle 0,1 \rangle$
Metoda średniej środków	CA (ang. <i>Center Average</i>)	$y^* = \frac{\sum_{m=1}^M y^{*m} \cdot \mu_{B'^m}(y^{*m})}{\sum_{m=1}^M \mu_{B'^m}(y^{*m})}$	y^* – wartość numeryczna na wyjściu modelu $y^* \in \mathfrak{I}$, B'^m – zbiór rozmyty konkluzji m -tej reguły ($m=1, \dots, M$), y^{*m} – środek zbioru rozmytego B'^m , $\mu_{B'^m}(y^{*m})$ – maksymalna wartość funkcji przynależności B'^m

W tabeli II-6 przeważają metody oparte na racjonalnej metodzie COG, uważanej w pracy [lee99] jako prawdopodobnie najlepszą ze znanych metod wyostrzania. Wyostrzanie ICOG $_\alpha$, zakłada istotność jedynie tych elementów, dla których wartość stopni przynależności przekracza pewną wartość α . Natomiast wyostrzanie MICOG $_\alpha$ dodatkowo zakłada usunięcie nieinformatywnej części funkcji przynależności [łes08] – figury znajdującej się pod ograniczeniem α . Z uwagi na trudności w wyznaczaniu całek, zwykle oblicza się wartość przybliżoną operacji wyostrzania poprzez zastosowanie dyskretyzacji oraz aproksymacji [pat02].

W literaturze ([dri93], [pie03], [łes08] i innych) można spotkać również inne metody wyostrzania, takie jak:

- metodę najmniejszego maksimum (SOM, ang. *Smallest value Of Maximum*) zwaną także metodą pierwszego maksimum (FOM, ang. *First Of Maximum*),
- metodę środka maksimum (MOM, ang. *Middle value Of Maximum*) zwaną także metodą centralnego maksimum (COM, ang. *Center Of Maximum*),
- metodę największego maksimum (LOM, ang. *Largest value Of Maximum*) zwaną także metodą ostatniego maksimum (LOM, ang. *Last Of Maximum*),
- metodę średniej maksimum (MeOM, ang. *Mean value Of Maximum*),
- metodę podstawowego rozkładu wyostrzania (BADD, ang. *BAsic Defuzzification Distribution*),
- metodę środka sum (COS, ang. *Center value Of Sums*),

- metodę opartą na α -przekrojach (GLSD, ang. *Generalized Level Set Defuzzification*) [fil93],
- quasi-liniową metodę wyostrzania (SLIDE, ang. *Semi-LInear DEfuzzification*) [yag93],
- oraz inne np. przedstawione w pozycjach [yag95], [liu07], [gen97].

5. Metody automatycznego pozyskiwania baz wiedzy w systemach rozmytych

Reguły JEŽELI-TO, stanowiące podstawę bazy wiedzy systemu rozmytego, mogą być definiowane na dwa sposoby. Jako reguły fizyczne, stanowiące obiektywne modele wiedzy zdefiniowane na podstawie obserwacji i badań naturalnych zachowań analizowanego obiektu oraz zachodzących w nim prawidłowościach (tzw. obiektywny charakter reguł [ban09]). Jak również, jako reguły logiczne, stanowiące subiektywne definicje tworzone przez człowieka na podstawie jego doświadczeń i wiedzy o badanym zjawisku (tzw. subiektywny charakter reguł [ban09]). W przypadku modelowania rozmytego początkowo wykorzystywano reguły logiczne, jednakże w miarę rozwinięcia dziedziny maszynowego uczenia zaczęto stosować hybrydę omówionych reguł. Wówczas początkowe założenia dotyczące zbiorów rozmytych, a nawet reguł, są definiowane na zasadzie przekonań eksperta, natomiast pozostałe parametry zostają dopasowane do danych pomiarowych. Celem metod automatycznego pozyskiwania baz wiedzy jest uzyskanie jak najmniejszego zbioru rozmytych reguł JEŽELI-TO, który umożliwia najdokładniejsze odwzorowanie modelowanego obiektu czy zjawiska.

5.1. Metody pozyskiwania baz wiedzy dla modeli lingwistycznych

W podrozdziale zostaną przedstawione wybrane metody pozyskiwania baz wiedzy dla modeli lingwistycznych: metoda Wanga-Mendela, metoda Nozaki-Ishibuchi-Tanaki, metoda Sugeno-Yasukawy oraz metoda szablonowego modelowania systemów rozmytych. Metody Wanga-Mendela oraz szablonowego modelowania systemów rozmytych pozwalają na wydobywanie przesłanek niezależnie od wydobywania konkluzji reguł. Metoda Nozaki-Ishibuchi-Tanaki wydobywa przesłanki reguł a następnie konkluzje, natomiast metoda Sugeno-Yasukawy odwrotnie – najpierw konkluzje, dopiero później przesłanki.

5.1.1. Metoda Wanga-Mendela

Metoda Wanga-Mendela [wan92] została opracowana w 1992 r. i stanowi jedną z prostszych metod wydobywania reguł lingwistycznych (38) na podstawie danych numerycznych.

Omwiana metoda realizowana jest w następujących krokach (na podstawie [wan92], [kor02]):

- 1) Przestrzeń wejściową X_n i wyjściową Y modelu dyskretyzuje się na nieparzystą liczbę regionów. Dla każdego regionu, określa się zbiór rozmyty (w oryginalnej metodzie – o trójkątnej funkcji przynależności), odpowiednio oznaczony jako $A_1^n, A_2^n, \dots, A_{2l_n+1}^n$ dla n -tego wejścia x_n oraz $B_1, B_2, \dots, B_{2l+1}$ dla wyjścia y .
- 2) Dla każdego pomiaru wejścia-wyjście $(\mathbf{x}(i), y(i))$, w zbiorze danych uczących, ustala się rozmytą regułę:

$$R(i): \text{JEŽELI } x_1 = A_{i*}^1 \text{ I } \dots \text{ I } x_N = A_{i*}^N \text{ TO } y = B_{i*}, \quad (78)$$

gdzie: wskaźniki i^* dla $A_{i^*}^n, B_{i^*}, n \in \{1, \dots, N\}$ są określone następująco:

$$\mu_{A_{i^*}^n}(x_n(i)) = \max_{1 \leq j \leq 2t_n+1} \mu_{A_j^n}(x_n(i)), \quad (79)$$

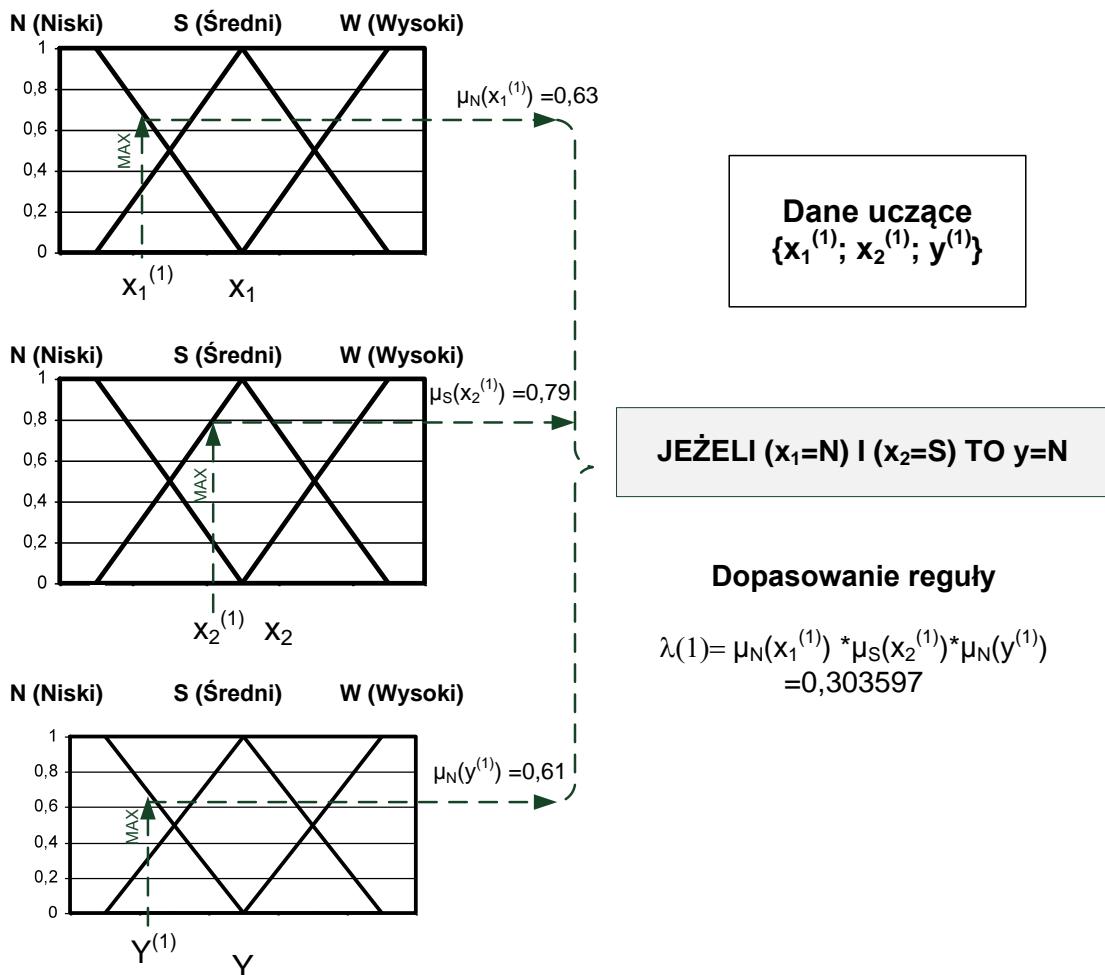
$$\mu_{B_{i^*}}(y(i)) = \max_{1 \leq j \leq 2l+1} \mu_{B_j}(y(i)). \quad (80)$$

- 3) Dla każdej i -tej reguły wyznacza się wartość dopasowania jako:

$$\lambda(i) = T(\mu_{B_{i^*}}(y(i)), \mu_{A_1^{i^*}}(x_1(i)), \dots, \mu_{A_N^{i^*}}(x_N(i))), \quad (81)$$

gdzie: operator t-normy T określony jest za pomocą normy trójkątnej algebraicznej (tab. II-4). Przykład generowania reguły wraz z wyliczeniem jej dopasowania jest zamieszczony na rysunku II-9.

- 4) Docelową bazę wiedzy otrzymuje się poprzez jej redukcję tj. eliminację reguł sprzecznych – zostawiając regułę o największym dopasowaniu λ oraz usunięcie reguł powtarzających się w danych regionie dyskretyzacji przestrzeni (patrz punkt 1) – również zostawiając regułę o największym dopasowaniu λ .



Rys. II-9. Przykład generowania reguły metodą Wanga-Mendela

5.1.2. Metoda Nozaki-Ishibuchi-Tanaki

Nozaki, Ishibuchi, Tanaki [noz97] opracowali prostą, heurystyczną metodę wydobywania reguł lingwistycznych. Metoda ta realizowana jest w następujących krokach (na podstawie [łęs08]):

- 1) Utworzenie podziału przestrzeni wejściowej X_n i wyjściowej Y na obszary rozmyte, odpowiednio $A_1^n, A_2^n, \dots, A_{2t_n+1}^n$ dla n -tego wejścia x_n oraz $B_1, B_2, \dots, B_{2l+1}$ dla wyjścia y .
- 2) Estymacja wartości położen singletonów b_{j_1, \dots, j_N} w konkluzjach reguł, na podstawie zbioru uczącego, z uwzględnieniem średniej ważonej wartości wyjściowych:

$$b_{j_1, \dots, j_N} = \frac{\sum_{i=1}^M [\tau_{j_1, \dots, j_N}(\mathbf{x}(i))]^\alpha y(i)}{\sum_{i=1}^M [\tau_{j_1, \dots, j_N}(\mathbf{x}(i))]^\alpha}, \alpha > 0, \quad (82)$$

gdzie: $j_n \in \{1, 2, \dots, 2t_n + 1\}$ dla $n \in \{1, \dots, N\}$ stanowi numer zbioru rozmytego, $(j_1, \dots, j_n, \dots, j_N)$ – wskaźnik reguły, $\mathbf{x}(i)$ – i -ty wektor wartości pomiarowych dla wejścia $i=1, \dots, M$, $\tau_{j_1, \dots, j_N}(\mathbf{x}(i))$ – stopień aktywacji dla $\mathbf{x}(i)$ określany za pomocą operatora t-normy:

$$\tau_{j_1, \dots, j_N}(\mathbf{x}(i)) = T(\mu_{A_{j_1}}(x_1(i)), \dots, \mu_{A_{j_N}}(x_N(i))) \quad (83)$$

- 3) Biorąc pod uwagę reguły typu (78) – pozyskiwanie reguł ogranicza się do wyszukania zbiorów rozmytych w konkluzji $B_1, B_2, \dots, B_{2l+1}$ o największym i drugim z kolei największym dopasowaniu do każdego singletonu:

$$C_{(j_1, \dots, j_N)^{1*}} = \max_{1 \leq j \leq J} \mu_{B_j}(b_{j_1, \dots, j_N}), \quad (84)$$

$$C_{(j_1, \dots, j_N)^{2*}} = \max_{1 \leq j \leq J; j \neq (j_1, \dots, j_N)^{1*}} \mu_{B_j}(b_{j_1, \dots, j_N}), \quad (85)$$

Wartości (84) i (85) stanowią współczynniki przekonania (ang. *certainty factor*) dla reguł typu:

$$\text{JEŻELI } x_1 = A_{j_1}^1 \text{ I } \dots \text{ I } x_N = A_{j_N}^N \text{ TO } y = B_{(j_1, \dots, j_N)^f*}(C_{(j_1, \dots, j_N)^f*}), \quad (86)$$

gdzie $f=1,2$. W pracy [noz97] podana została formuła pozwalająca na obliczenie wyjścia na podstawie bazy reguł typu (86).

5.1.3. Metoda Sugeno-Yasukawy

Metoda Sugeno-Yasukawy, w odróżnieniu do poprzednich metod, w pierwszej kolejności wydobywa konkluzje reguł. W tym celu stosuje się grupowanie rozmytą metodą k-średnich, gdzie z góry założona jest liczba grup k , stanowiąca o liczbie reguł I w modelu rozmytym. W oryginalnej metodzie, Sugeno i Yasukawa [sug93] wyznaczali liczbę reguł minimalizując wskaźnik jakości grupowania. Wartości funkcji przynależności zbiorów wyjścia y (B_1, B_2, \dots, B_I) otrzymuje się na podstawie rozmytej macierzy podziału U ,

otrzymanej podczas rozmytego grupowania. Kolejno, poprzez projekcję danych wejściowych z określonej grupy (reguły) na przestrzeń zmiennych wejściowych, otrzymuje się przesłanki reguł. Tak wyznaczone funkcje przynależności wymagają aproksymacji gładką funkcją przynależności (w oryginalnej metodzie zastosowano f. przynależności trapezowej) w celu uzyskania gładkiej funkcji przejścia modelu.

5.1.4. Metoda szablonowego modelowania systemów rozmytych

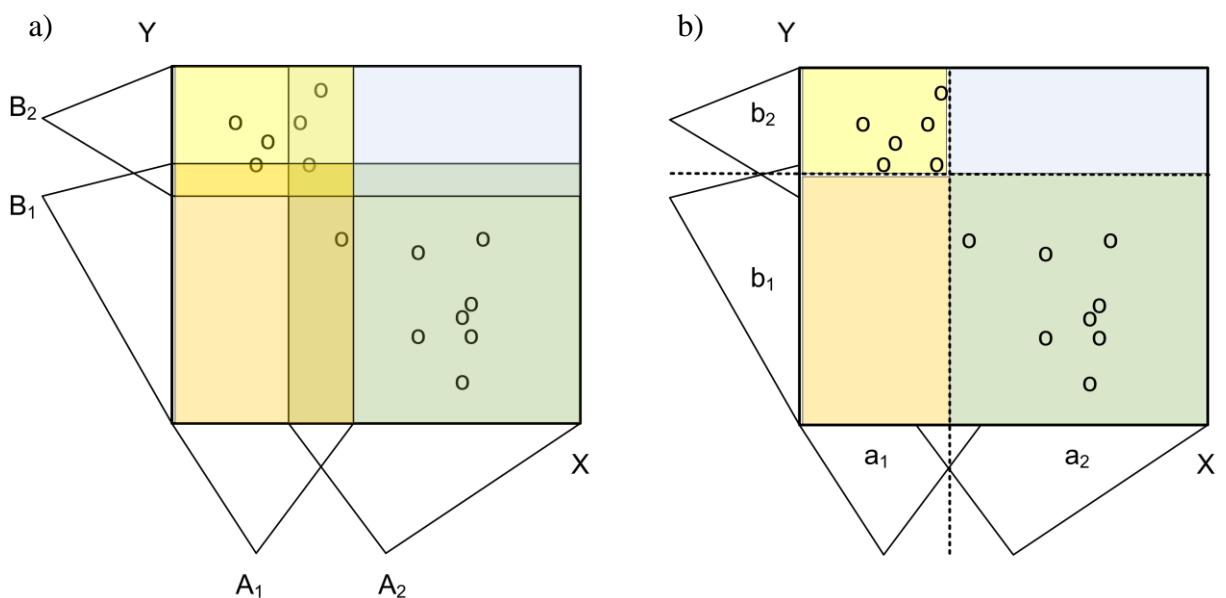
Podejście szablonowego modelowania systemów rozmytych zostało opisane w [yag95]. Metoda ta wymaga znania dwóch informacji: danych doświadczalnych (x_k, y_k) oraz określonych przez eksperta, wartości zbiorów lingwistycznych. Dla systemów SISO, podział przestrzeni wejścia X ($x \in X$) oraz wyjścia Y ($y \in Y$) na obszary rozmyte odpowiednio A_i ($i=1, \dots, I$) i B_j ($j=1, \dots, J$) tworzy tzw. szablon. Reguła elementarna powstała z szablonów A_i i B_j ($R(A_i, B_j)$) ma postać:

$$\text{JEŻELI } x = A_i \text{ TO } y = B_j \text{ z wiarygodnością } w_{ij}, \quad (87)$$

gdzie: x stanowi zmienną wejściową modelu, y – zmienną wyjściową modelu, w_{ij} stanowi wagę reguły (prawdopodobieństwo lub wiarygodność).

Istnieją dwie metody uczenia wag w_{ij} w powyższych regułach: metoda adaptacyjnego uczenia reguł rozmytej pamięci asocjacyjnej (ang. *Adaptive Fuzzy Association Memory Rule Learning*) ([yag95], [kos92a], [liu04]) oraz metoda oparta na miarach teorii Dempstera-Shafera [yag95].

W pierwszym podejściu przestrzeń X i Y zbiorów rozmytych jest przybliżona do rozłącznych przedziałów. Wówczas siatka rozmyta składająca się z obszarów rozmytych $R(A_i, B_j)$ zostaje zastąpiona nierożmytymi, rozłącznymi przedziałami $R(a_i, b_j)$ (por. rys. II-10).



Rys. II-10. Rozkład danych wejście-wyjście z uwzględnieniem szablonów: a) siatka rozmyta,
b) siatka nierożmyta (na podstawie [yag95])

Na podstawie danych empirycznych wyznaczany jest łączny rozkład prawdopodobieństwa dla siatki nierożmytej. Wówczas estymator wagi w_{ij} wynosi:

$$w_{ij} = n_{ij}/n, \quad i=1, \dots, I, \quad j=1, \dots, J, \quad (88)$$

gdzie: $n = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$ stanowi całkowitą liczbę próbek wejście-wyjście, n_{ij} – liczbę próbek w obszarze rozłącznym $a_i \times b_j$.

Estymacja grupowa w przestrzeni $X \times Y$, zwana grupowaniem w przestrzeni iloczynu (ang. *product space clustering*) ([kos92a], [kos92b]) może stanowić alternatywę dla omówionej metody tworzenia łącznego rozkładu prawdopodobieństwa.

W teorii Dempstera-Shafera podzbiorom przestrzeni zdarzeń przypisuje się podstawową miarę prawdopodobieństwa (BPA, ang. *Basic Probability Assignment*) oznaczaną często jako M (bądź m). Dlatego też, w metodzie modelowania szablonowego opartej na mierze prawdopodobieństwa wg teorii Dempstera-Shafera, zbiór reguł elementarnych, które wchodzą w skład jednej reguły plikowej, zapisuje się jako:

$$\begin{aligned} & \text{JEŻELI } x = A_i \text{ TO } y = M_{i1} \\ & \text{JEŻELI } x = A_i \text{ TO } y = M_{i2} \\ & \quad \dots \\ & \text{JEŻELI } x = A_i \text{ TO } y = M_{iJ}, \end{aligned} \quad (89)$$

przy czym M_{ij} , $j=1, \dots, J$, $i=1, \dots, I$ stanowi wiarygodną strukturę o rozmytym elemencie ogniskowym B_{ij} i wagą $P_{ij}=M_{ij}(B_{ij})$, jako prawdopodobieństwo [yag95]. Toteż, reguły (89) są innym zapisem reguły:

$$\begin{aligned} & \text{JEŻELI } x = A_i \text{ TO } y = B_{i1} \text{ z wagą } P_{i1} \\ & \text{JEŻELI } x = A_i \text{ TO } y = B_{i2} \text{ z wagą } P_{i2} \\ & \quad \dots \\ & \text{JEŻELI } x = A_i \text{ TO } y = B_{iJ} \text{ z wagą } P_{iJ}. \end{aligned} \quad (90)$$

Proces uczenia wag reguł rozpoczyna się od wyliczenia rozmytego stopnia dopasowania każdej reguły elementarnej. Dla zbioru par doświadczalnych (x_k, y_k) i reguły elementarnej w postaci JEŻELI $x = A_i$ TO $y = B_{ij}$, wynosi on:

$$\delta_{ij}(k) = \mu_{A_i}(x_k) \mu_{B_{ij}}(y_k). \quad (91)$$

Na tej podstawie, znormalizowany rozmyty stopień dopasowania każdej reguły wynosi:

$$\nu_{ij}(k) = \frac{\delta_{ij}(k)}{\sum_{i=1}^I \sum_{j=1}^J \delta_{ij}(k)}. \quad (92)$$

Skumulowany znormalizowany stopień dopasowania reguły elementarnej względem całego zbioru danych doświadczalnych wynosi:

$$v_{ij} = \sum_k v_{ij}(k). \quad (93)$$

Dla poszczególnych reguł plikowych (90)częstość warunkową otrzymujemy jako:

$$v_i = \sum_{j=1}^J v_{ij}, \quad i=1,\dots,I. \quad (94)$$

Ostatecznie, wagę P_{ij} można uzyskać za pomocą częstości warunkowej reguły elementarnej:

$$P_{ij} = \frac{v_{ij}}{v_i}. \quad (95)$$

Reguły o zerowych wagach P_{ij} są eliminowane. Rozszerzenie obu metod na model MISO jest proste i nie zmienia sensu prezentowanych podejść [yag95].

5.2. Metody eksploracji danych i maszynowego uczenia

Pojęciem nieodzwonnie związanym z automatycznym pozyskiwaniem baz wiedzy jest *maszynowe uczenie się* (ang. *machine learning*), zwane również *uczeniem się maszyn*. Pojęcie to odnosi się do całej dziedziny naukowej, będącej gałęzią sztucznej inteligencji i zajmującej się badaniami nad sztucznymi systemami uczącymi się [cic00]. Uczenie, można tutaj rozumieć jako każdą, autonomiczną zmianę w systemie, zachodzącą na podstawie zgromadzonego doświadczenia (czyli danych), która prowadzi do poprawy jego działania. Tak rozpatrywana definicja powoduje różnorodność procesu uczenia się. Stąd też istnieje szereg kryteriów, według których możemy uzyskać szczegółowe obszary nauki. Wśród nich największe znaczenie mają kryteria tj. metoda reprezentacji wiedzy, postać i źródło informacji stosowanych do uczenia czy mechanizmy wydobywania wiedzy itp. Uczenie maszynowe prowadzi do określonych celów tj. wykrywanie nieznanych prawidłowości w danych, tworzenie reguł decyzyjnych, definiowanie nowych pojęć, modyfikowanie, uogólnianie i precyzowanie danych, zdobywanie wiedzy przez interakcję z otoczeniem itp. Informacje na temat szczegółowych metod uczenia można znaleźć między innymi w pozycjach: [cic00], [mit97], [kor94], [bis06], [kon07], [kor08], [kra04], [duc00].

Innym obszarem badań, związanych z automatycznym pozyskiwaniem danych, są metody wchodzące w skład eksploracji danych (inaczej drążenia danych, ang. *data mining*). Eksploracja danych, jako główny etap procesu odkrywania wiedzy (ang. *knowledge discovery*) [fay96], zajmuje się nietrywialnymi algorytmami odkrywania „ukrytych”, dotąd nieznanych i potencjalnie potrzebnych informacji w danych [fra92] oraz zapisywania ich w postaci wzorców i modeli wiedzy.

Można zatem zauważyć, że wspólnym celem maszynowego uczenia i eksploracji danych jest odkrywanie wiedzy zapisanej w przykładach. Stąd też, mimo różnej genealogii podejść i wykorzystywanych metod [sim], pojęcia te się często przeplatają, bądź są stosowane zamiennie. Wg [kra04] algorytmy maszynowego uczenia stanowią podstawę metodyczną dla algorytmów odkrywania wiedzy z danych. Pozycja [cio98] traktuje uczenie maszynowe jako jedno z narzędzi data mining, na równi z sieciami neuronowymi, zbiorami rozmytymi, zbiorami przybliżonymi, algorytmami genetycznymi itp. Natomiast w [hol93] metody data mining stanowią wsparcie dla konstruowania i pracy systemów ekspertowych.

Ogólnie, najczęściej pojęcie maszynowego uczenia jest stosowane w stosunku do dziedzin naukowych związanych ze sztuczną inteligencją tj. robotyka, sterowanie, biocybernetyka czy

statystyka, gdzie proces uczenia się systemu ma za zadanie osiągnięcie rezultatów opartych na wiedzy fragmentarycznej, ma umożliwiać uczenie się, tworzenie nowych pojęć oraz wnioskowania indukcyjnego. Natomiast pojęcie eksploracji danych odnosi się najczęściej do dziedzin informatycznych związanych z wydobywaniem wiedzy z dużych kolekcji danych do których należą m.in. hurtownie danych, gdzie odbiorcami informacji są analitycy, menadżerowie oraz decydenci [lar06]. Frawley i inni [fra92] wyszczególniają różnice w podejściach. Wg [fra92] eksploracja danych jest trudniejsza z uwagi na dynamiczność, niekompletność i znacznie większy rozmiar danych w odróżnieniu od typowych danych uczących wykorzystywanych do maszynowego uczenia. Stały wzrost danych gromadzonych w różnych dziedzinach zastosowań powoduje, iż powyższa różnica ulega zatarciu. Dlatego też, techniki i metody obu podejść się dublują lub są wspólnie rozpatrywane. Ponadto istnieje kilka dziedzin blisko związanych z maszynowym uczeniem i eksploracją danych np. statystyka i różne formy analizy danych (oznaczone przymiotnikami: wielowymiarowe, Bayesowskie, inteligentne,...), których rozgraniczenie jest niemal niemożliwe [hül05].

Niniejsza dysertacja koncentruje się na metodach maszynowego uczenia i eksploracji danych wspomagających pozyskiwanie informacji zapisanych w postaci rozmytych reguł warunkowych. Niektóre z metod identyfikują regiony w przestrzeni zmiennych systemu, które później tworzą zdarzenia rozmyte w regułach. Może być to realizowane poprzez szukanie kластerów z użyciem algorytmów grupowania, czy też identyfikację za pomocą tzw. *covering* (inaczej *separate and conquer*) algorytmu. Inne metody bazują na stałym rozmytym podziale dla każdego atrybutu (tzw. regularnej „*fuzzy grid*”), a każdy element siatki jest rozważany jako potencjalny składnik reguły.

5.2.1. Metody grupowania

Metody grupowania (zwane *klasteringiem* bądź *analizą skupień*, ang. *clustering*) generują bezpośrednio rozmyte reguły w postaci JEŻELI-TO lub są jedynie wykorzystywane przy pozyskiwaniu funkcji przynależności [pie03]. W pierwszej metodzie, każda grupa podziału tworzy osobne funkcje przynależności dla zmiennych wejścia-wyjścia. Z matematycznego punktu widzenia metody grupowania danych polegają na podziale N – elementowego wektora danych $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ na c grup G_1, G_2, \dots, G_c zwanych *klasteryami* (*klastrami*), *skupiskami* bądź *klasami*, w taki sposób, aby dane należące do tej samej grupy były bardziej podobne do siebie, niż do danych z innych grup. Liczba grup c jest zadana arbitralnie bądź ustalona automatycznie na podstawie użytej metody [łęs08]. W literaturze istnieje wiele metod grupowania danych do których należą m. in. metody hierarchiczne (aglomeracyjne oraz dzielące), metody oparte na teorii grafów, metody oparte na dekompozycji funkcji gęstości prawdopodobieństwa, czy też metody oparte na minimalizacji funkcji kryterium lub skalarnego wskaźnika jakości. Do najpopularniejszych sposobów grupowania należy metoda z ostatniej grupy podziału – *metoda k-średnich* (ang. *k-means*, *c-means*) [oso96] [bez05] zwana czasem *metodą k-środków* [pie03] lub *ISODATA* (ang. *Iterative Self-organizing DATA Clustering*) [łęs08].

Przy założeniu, że wektor danych może należeć do więcej niż jednej grupy G_1, G_2, \dots, G_c , mamy do czynienia z rozmytymi metodami grupowania. Tą metodę zaproponował po raz pierwszy E. Ruspini w 1969 roku [rus69]. Wówczas, jeżeli oznaczymy przestrzeń wszystkich $(c \times N)$ -wymiarowych macierzy o elementach rzeczywistych jako $R^{c \times N}$, to zbiór wszystkich możliwych macierzy rozmytych podziałów ma postać:

$$\mathbf{M}_{fc} = \{\mathbf{U} \in R^{c \times N} \mid \forall_{1 \leq i \leq c} \forall_{1 \leq j \leq N} u_{ij} \in \langle 0,1 \rangle, \forall_{1 \leq j \leq N} \sum_{i=1}^c u_{ij} = 1, \forall_{1 \leq i \leq c} \sum_{j=1}^N u_{ij} \in (0, N) \}, \quad (96)$$

gdzie elementy macierzy przyjmują wartości:

$$u_{ij} \in \begin{cases} \{0\}, \mathbf{x}_j \notin G_i \\ (0,1), \mathbf{x}_j \in G_i \end{cases}. \quad (97)$$

Istnieje również wiele rozmytych sposobów grupowania danych, do najczęściej spotykanych metod należą: *rozmyta metoda c-srednich* (zwana również *rozmytą metodą ISODATA*, ang. *Fuzzy C-Means*, FCM) oparta na minimalizacji funkcji kryterium oraz *metoda Gustafsona i Kessela* dopuszczająca różnorodność kształtów geometrycznych grup, *klasteryzacja różnicowa* (ang. *subtractive clustering*) stanowiąca rozszerzenie *metody górskiej Yagera* [yag95][pri05]. Opisy wymienionych powyżej oraz innych przykładów metod grupowania rozmytego i nierożmytego można znaleźć między innymi w pozycjach: [łęs08], [rut06], [mit97], [tan06], [bar99b], [bez05], [kay00], [fuk00], [ham00], [vis09], [mij08].

5.2.2. Metody rozmytych drzew decyzyjnych

Większość metod indukcji drzew decyzyjnych nie zajmuje się problemami obarczonymi niepewnością oraz niejasnością związaną ze sposobem ludzkiego myślenia i postrzegania świata. Analiza rozmytości stwierdzeń została dopiero włączona do procesu indukcji wiedzy metodą drzew decyzyjnych w latach 80-tych.

Jak wskazuje sama nazwa, drzewa decyzyjne mają strukturę drzewiastą, w której węzły wewnętrzne zawierają testy na wartościach poszczególnych danych wejściowych. Z każdego węzła wychodzi tyle gałęzi, ile jest możliwych wyników testu. Liście drzewa zawierają odpowiedź (decyzje) dotyczącą analizowanego zagadnienia. Po utworzeniu drzewa, struktura ta może być zamieniana na zbiór rozmytych reguł warunkowych postaci (54) lub (44). Reguły uzyskujemy idąc od korzenia drzewa do każdego liścia. W ten sposób powstaje część przesłankowa reguły. Konkluzja zaś stanowi wartość (zbiór rozmyty lub funkcję, w zależności od typu rozmytych reguł decyzyjnych) znajdującą się w liściu.

Najczęstsze zastosowania drzew decyzyjnych to klasyfikacja i regresja. Jednakże, równie dobrze mogą być stosowane do generalizacji danych w szeroko pojętym obszarze podejmowania decyzji, sterowania, prognozy, diagnozy czy wykrywaniu awarii – wszędzie tam, gdzie wystarczająca jest pojedyncza wartość wyjściowa systemu.

W literaturze można znaleźć wiele algorytmów generujących rozmyte drzewa decyzyjne, wśród nich należy wymienić modyfikacje prostego, rekurencyjnego algorytmu ID3 tworzącego nierożmyte drzewa decyzyjne [qui86] [qui90] – algorytm rozmyty ID3 (ang. *fuzzy ID3*) [web92] [cio92] oraz jego warianty [uma93] [wan00], algorytm neuronowo-rozmyty ID3 (ang. *neuro-fuzzy ID3*) [ich96]. W powyższych algorytmach istnieją dwie metodologie postępowania podczas generowania drzew decyzyjnych. Główną różnicą w podejściach jest wykorzystywanie rozmytej entropii jako kryterium klasyfikacji. Niektóre algorytmy bazują na minimalizacji rozmytej entropii [web92] [uma93], inne zaś mierzą bezpośrednio jakość podziału reguł w węźle [yua95] [wan00].

Najczęściej jednak zdolność odzwierciedlenia modelu poprzez użycie drzew decyzyjnych nie jest najlepsza [qui94], [sik07]. Należy wówczas stosować metody poprawiające jakość dopasowania modelu. W celu rozwiązania problemu przeuczenia się struktury stosuje się metody obcinania drzewa decyzyjnego (ang. *pruning*), co jednak

wydłuża proces generowania reguł. Aby poprawić rezultaty, w niektórych strukturach rozmytych drzew decyzyjnych, stosuje się także procedury wygładzania (ang. *smoothing*), dzieje się to jednak kosztem czytelności reguł [sik07]. Należy również podkreślić, że w przypadku złożonych problemów, drzewa decyzyjne potrafią być również bardzo złożone [cic00]. Wiąże się to z faktem, że stosowane testy na ogólny sprawdzają wartości tylko pojedynczych atrybutów, tracąc szansę na wykorzystanie ewentualnych zależności pomiędzy wartościami różnych atrybutów, co mogłyby prowadzić do prostszej reprezentacji wiedzy. Ponadto, do znacznego rozrostu wiedzy prowadzi reprezentowanie alternatywnych warunków [cic00]. Generowanie drzew decyzyjnych ma również swoje zalety. Jedną z nich jest niewielka złożoność obliczeniowa (zakładając, że w jednym teście sprawdza się jeden atrybut), która jest ograniczona liniowo przez liczbę atrybutów. Drzewa decyzyjne są ponadto czytelne dla człowieka, a ewentualne przejście na reprezentację regułową nie stanowi problemu.

5.2.3. Rozmyte metody hybrydowe

Efektywne z punktu widzenia osiąganych wyników są metody hybrydowe wykorzystywane w systemach neuronowo-rozmytych (ang. *neuro-fuzzy systems*, połączenie systemów rozmytych z sztucznymi sieciami neuronowymi), systemach ewolucyjno-rozmytych (ang. *evolutionary-fuzzy systems*, połączenie systemów rozmytych z algorytmami ewolucyjnymi) bądź w ich hybrydach – systemach ewolucyjno-neuronowo-rozmytych. Zdolność uczenia się sieci neuronowych i algorytmów ewolucyjnych mogą być zastosowane do automatycznego pozyskiwania rozmytych reguł warunkowych. Algorytmy pozwalają na dostosowanie funkcji przynależności oraz rozmytych reguł warunkowych lub obu elementów jednocześnie. Możliwe jest również, choć rzadko spotykane w praktyce, jednoczesne dostosowanie typów funkcji przynależności poprzez zastosowanie metod ewolucyjnych [łes08]. Systemy hybrydowe w zdobywaniu wiedzy wykorzystują wiedzę eksperta oraz dane pomiarowe (np. metoda Ishibuchi i innych [ish93], metoda Horikawy i innych [hor92]) lub mogą opierać się wyłącznie na danych pomiarowych (np. system Janga – ANFIS [jan93] [jan95], system ANNBFIS [łes99], system ANBLIR [łes08]).

Algorytmy wykorzystywane w sieciach neuronowo-rozmytych zazwyczaj budują sieć w oparciu o wynik grupowania danych (patrz rozdział II-4.2.1). Liczba grup oraz liczba zbiorów rozmytych dzielących dziedzinę każdej cechy ustalane są arbitralnie lub adaptacyjnie. Wiele opisanych w literaturze metod pozyskiwania reguł opiera się na sieciach o radialnych funkcjach bazowych (np. [par91], [jan93], [hor92], [cho96]). Uczenie w powyższych systemach w większości przypadków minimalizuje błąd średniokwadratowy wyników. Zaletą architektur rozmytych sieci neuronowych jest to, że użytkownik ma możliwość interpretacji wag zawartych w połączeniach neuronowych, w przeciwieństwie do klasycznych sieci neuronowych [łes08]. Jednakże, modyfikacja funkcji przynależności wpływa na to, iż użytkownik nie może wprost interpretować reguł z użyciem zrozumiałych dla niego wartości lingwistycznych.

Zastosowanie algorytmów ewolucyjnych pozwala na uogólnianie strojenia parametrów modelu poprzez zastosowanie globalnych metod optymalizacji. Najczęściej stosowanym podejściem jest, zaproponowane przez S.F. Smitha, podejście typu Pittsburgh [smi80], w którym każdy osobnik zawiera zakodowaną całą bazę wiedzy w chromosomie o zmiennej długości. W procesie ewolucji cała baza wiedzy ulega modyfikacji, co prowadzi do wyboru optymalnego zbioru reguł. Poszukiwanie optymalnych parametrów bazy wiedzy w hiperprzestrzeni możliwych wartości jest zadaniem złożonym obliczeniowo. Toteż, w celach skrócenia czasu obliczeń stosowane są sieci neuronowe do estymacji wymaganych wartości metodą najmniejszych kwadratów [lee03],[łes08].

Systemy hybrydowe nie są bezpośrednim tematem niniejszej dysertacji, zatem zostały opisane bardzo przeglądowo. Więcej informacji na ten temat można znaleźć w pozycjach anglojęzycznych m.in. [jan93], [bac99], [gor02], [gon96], [hel97], a także w literaturze w języku polskim m.in.: [rut97b], [rut97a], [duc00], [łęs08], [cpa09].

6. Reguły asocjacji

6.1. Istota, rodzaje i zastosowania reguł asocjacji

Reguły asocjacji stanowią jedną spośród technik i algorytmów wykorzystywanych podczas eksploracji danych. Pojęcie reguł asocjacji (ang. *association rules*) zostało po raz pierwszy użyte w [agr93], jako metoda odnajdywania współwystępowania wartości atrybutów w obszernych kolekcjach danych [alv02]. W ujęciu formalnym reguły asocjacji możemy zapisać w postaci:

$$\theta \Rightarrow \varphi : X \Rightarrow Y (s, c), \quad (98)$$

gdzie X i Y są rozłącznymi zbiorami atrybutów, nazwanych często: X – zbiorem wartości warunkujących, Y – zbiorem wartości warunkowanych. A dokładniej:

$$\theta \Rightarrow \varphi : A_1 \cap \dots \cap A_n \Rightarrow A_{n+1} \cap \dots \cap A_m (s, c), \quad (99)$$

gdzie A_1 (A_n) są parami atrybut-wartość przesłanki reguły, A_{n+1} (A_m) – parami atrybut-wartość następnika reguły. Pary atrybut-wartość A_n są najczęściej zapisywane jako " $x_n = A_n$ " lub " x_n jest A_n ", gdzie x_n stanowi analizowany atrybutu, natomiast A_n jego wartość w ujęciu binarnym, jako element ze zbioru $\{0;1\}$ bądź, dla atrybutów ilościowych, jako przedział zmienności atrybutu.

Każda reguła asocjacji jest związana z dwiema miarami statystycznymi określającymi ważność i siłę reguły: *support* $s(\theta \Rightarrow \varphi)$ (w literaturze oznaczane często jako *sup*) oraz *confidence* $c(\theta \Rightarrow \varphi)$ (w literaturze oznaczane często jako *conf*). *Support* – zwane wsparciem, stanowi prawdopodobieństwo jednokrotnego występowania zbiorów θ i φ w kolekcji zbiorów. *Confidence* natomiast, zwane ufnością bądź wiarygodnością, stanowi prawdopodobieństwo warunkowe $P(\varphi | \theta)$. W literaturze [dub06] można również spotkać definicje obu miar statystycznych przedstawionych jako liczbę występowania odpowiednich wartości w kolekcjach danych, bądź też jako wartość procentowa.

Wsparcie jest istotną miarą wartościującą daną regułę asocjacji, gdyż określa liczbę transakcji w analizowanym zbiorze, które potwierdzają daną regułę. Ufność reguły określa na ile odkryta reguła asocjacji jest "pewna". Reguły o niskiej ufności są mało wiarygodne, natomiast reguły charakteryzujące się wysoką ufnością są "prawie pewne" [mor].

Na gruncie analizy danych, reguły asocjacyjne są przede wszystkim stosowane do analizy koszyka zakupów (MBA, ang. *market basket analysis*) [agr93] [han95] [sav09], organizowania promocji i sprzedaży wiązanej, do konstruowania katalogów wysyłkowych, ustalania rozmieszczenia towarów na półkach itp. [mor05]. Wówczas odnajdowane jest współwystępowanie wartości tylko jednego atrybutu – produktu np. „jeżeli klient kupi chleb i ser żółty to kupi także masło (4%,70%)”, co stanowi odnajdywanie tzw. pojedynczych reguł asocjacji. Istnieje także możliwość wyszukiwania wielowymiarowych reguł asocjacji – gdzie odnajdywane jest częste współwystępowanie wartości różnych atrybutów np. „jeżeli wiek=<20,29> i wykształcenie='średnie' to głosowanie='PO' (5%,60%)”. Wielowymiarowe

reguły asocjacji poszerzają możliwości zastosowań omawianej metody, stąd też w literaturze można spotkać zastosowania, w następujących obszarach:

- analizie rozmów telefonicznych,
- diagnostyce medycznej [he06],
- bankowości,
- analizie transakcji kartami kredytowymi,
- analizie nawigacji po stronach internetowych [kaz09],
- szeroko pojętej analizie web mining,
- badaniu wpływu czynników geograficznych na porost roślinności [shu08],
- analizie społeczno-ekonomicznej i urbanizacyjnej miasta [men05],
- analizie uwarunkowań rozwoju małych i średnich przedsiębiorstw [bla07c],
- analizie czynników pogodowych [tso01].

Rozległe możliwości zastosowań reguł asocjacji mają związek z rozwinięciem tej techniki i powstawaniem wielu jej odmian. Początkowo odkrywano jedynie wartości zamieszczone jako wartości binarne (binarne reguły asocjacji, ang. *binary* lub *Boolean association rules*), kolejno zaczęto tworzyć algorytmy służące również do odkrywania ilościowych reguł asocjacji (ang. *quantitative association rules*) [sri96], gdzie dane w regułach są danymi ciągłymi lub kategoriami. Inną modyfikacją algorytmu są wielopoziomowe (inaczej uogólnione) reguły asocjacji (ang. *multilevel* lub *generalized association rules*) [sri95], które pozwalają ukazać inny stopień abstrakcji przetwarzanych danych. Reguły wielopoziomowe, w odróżnieniu od reguł jednopoziomowych (ang. *single-level association rules*) uwzględniają taksonomię elementów wchodzących w skład reguł i umożliwiają odkrywanie reguł asocjacji, zawierających elementy z różnych poziomów taksonomii [mor06].

6.2. Wyszukiwanie ilościowych reguł asocjacji

Duży wysiłek włożono w opracowywanie efektywnych algorytmów odkrywania reguł asocjacji. Efektywne znalezienie i przeliczenie wszystkich dostępnych kombinacji atrybutów jest w większości przypadków nie możliwe ze względu na ogromną liczbę danych. Za podstawowy algorytm odkrywania reguł asocjacji uznaje się iteracyjny algorytm Apriori ([agr93], [agr94]). Algorytm Apriori doczekał się wielu modyfikacji, które zmierzają do poprawienia jego efektywności. Zaproponowano wówczas algorytm o nazwie AprioriTid [zhi05], który nie skanuje całej bazy transakcyjnej, a aktualnia przy każdym kroku strukturę danych, algorytm AprioriHybrid, który wykonuje pierwsze przeliczenia według metody Apriori, a kiedy już baza mieści się w pamięci wykonuje algorytm AprioriTid. Ponadto powstały również inne algorytmy, do których należą FP-Growth (ang. *Frequent Patterns Growth*) [han00b], [han04], [bor05], FreeSpan (ang. *Frequent Pattern-Projected Sequential Pattern Mining*) [han00a], Eclat [zak97] czy Partition [sav95]. W [mat02] przedstawiony został ewolucyjny algorytm odkrywania numerycznych reguł asocjacji.

W kolejnych rozdziałach omówione zostaną szczegółowo dwa algorytmy (Apriori i FP-Growth). Będą one wzięte pod uwagę podczas tworzenia bazy wiedzy dla systemu z probabilistyczno-rozmytą bazą wiedzy.

6.2.1. Algorytm Apriori

Algorytm Apriori dzieli problem wyznaczania reguł asocjacji na dwa zagadnienia. Po pierwsze należy wyodrębnić k -elementowe ($k=1,2,\dots$) zbiory częste tj. zbiory, dla których wsparcie jest większe od zakładanego \min_s . Wykorzystując właściwość monotoniczności miary wsparcia, k -elementowe zbiory częste powstają przez łączenie ($k-1$)-elementowych zbiorów częstych. Algorytm można przedstawić w postaci pseudo kodu (na podstawie [agr94], [han01]):

```

 $L_1 = \{ \text{zbiory częste 1-elementowe} \};$ 
for ( $k=2; L_{k-1} \neq \emptyset; k++$ ) do
begin
    //generowanie nowych zbiorów kandydujących
     $C_k = \text{apriori\_gen}(L_{k-1});$ 
    for each transakcji  $T \in D$  do begin
         $C_T = \text{podzbiory}(C_k, T)$  //podzbiory  $C_k$  zawarte w transakcji  $T$ 
        for each  $C \in C_T$  do
             $C.zlicz++;$ 
    end
    //zastosowanie warunku minimalnego wsparcia
     $L_k = \{ C \in C_k : C.zlicz \geq \min_s \};$ 
end
Zbiory częste =  $\bigcup_2^k L_k$  ;

procedure apriori_gen( $L_{k-1}$ )
for each  $l_i \in L_{k-1}$  do begin
    for each  $l_j \in L_{k-1}$  do begin
        if ( $l_i[1] = l_j[1] \wedge \dots \wedge l_i[k-2] = l_j[k-2] \wedge l_i[k-1] \neq l_j[k-1]$ ) then
             $C = \{l_i[1], \dots, l_i[k-2], l_i[k-1], l_j[k-1]\};$ 
            if  $C$  zawiera podzbiory  $\notin L_{k-1}$  then
                usuń  $C$ 
            else dodaj  $C$  do  $C_k$ ;
        end
    end
end
return  $C_k$ ;
end procedure                                     (100)

```

gdzie: C_k – rodzina zbiorów kandydujących k -elementowych, L_k – rodzina zbiorów częstych k -elementowych, \min_s – minimalne wsparcie, określone jako częstość występowania odpowiednich zbiorów.

Po drugie, ze zbiorów częstych należy wygenerować reguły asocjacji, które spełniają wymaganie minimalnej ufności ($\geq \min_c$). Ten etap algorytmu można przedstawić za pomocą następującego pseudokodu (na podstawie [tan06]):

```

for each zbiorów częstych  $k$ -elementowych  $L_k$ ,  $k \geq 2$  do
begin

```

```

H1={ i | i ∈ Lk} //1-elementowa konkluzja reguły
genrules(Lk, H1);
end

procedure genrules(Lk, Hm)
begin
    k=| Lk | //liczba elementów w zbiorze częstym
    m=| Hm | //liczba elementów w konkluzji
    if k>m+1 then
        begin
            Hm+1=apriori_gen(Hm);
            for each podzbioru hm+1 ∈ Hm+1 do
                begin
                    c= Lk.zlicz/( Lk- hm+1).zlicz;
                    if ufność c ≥ min_c then
                        return (Lk- hm+1) ⇒ hm+1
                    else
                        usuń hm+1 z Hm+1;
                end
            end
            call genrules(Lk, Hm+1)
        end
    end procedure                                     (101)

```

Kluczowe znaczenie, z punktu widzenia efektywności algorytmu odkrywania reguł asocjacji, ma pierwszy krok algorytmu – znajdowanie zbiorów częstych, to jest podzbiorów zbioru T , których wsparcie jest większe lub równe minimalnej wartości wsparcia min_s .

6.2.2. Algorytm FP-Growth

Algorytm FP-Growth (ang. *Frequent Pattern Growth*) [han00b] stanowi obecnie jeden z najszybszych i najbardziej popularnych algorytmów tego typu [bor05]. Proces odkrywania reguł asocjacji algorytmem FP-Growth ogranicza się do dwóch kroków:

- kompresji bazy danych D do drzewa tzw. FP-drzewa (ang. *FP-tree, Frequent Pattern tree*),
- eksploracji FP-drzewa w celu znalezienia zbiorów częstych.

Tworzenie FP-drzewa realizowane jest w następujących etapach (na podstawie [mor]):

- a) Znajdowanie wszystkich 1-elementowych zbiorów częstych w bazie danych D .
- b) Przekształcenie każdej transakcji $T_i \in D$ w transakcję skompresowaną T_{ki} poprzez usunięcie z niej nieczęstych 1-elementowych zbiorów.
- c) Posortowanie elementów transakcji T_{ki} według malejącej wartości ich wsparcia.
- d) Transformacja posortowanych transakcji T_{ki} do FP-drzewa, gdzie transakcje ze wspólnymi k elementami współdzielą jedną ścieżkę w drzewie, a każdy wierzchołek zawiera licznik transakcji z których składała się dana ścieżka.

Procedura rekurencyjna przeszukiwania FP-drzewa w celu znalezienia wszystkich zbiorów częstych przedstawiona została w postaci pseudokodu (na podstawie [mor]):

```

procedure FP-growth(tree, α)
    if tree zawiera pojedynczą ścieżkę P
    then
        for each kombinacji β wierzchołków ścieżki P do
        begin
            generuj zbiór  $\beta \cup \alpha$  taki, że  $s(\beta \cup \alpha)$  jest równe
            minimalnemu wsparciu elementów należących do β;
        end
    else
        for each α-i wskaźnika elementu w drzewie tree do
        begin
            generuj zbiór  $\beta = \alpha - i \cup \alpha$  taki, że  $s(\beta) = s(\alpha - i)$ ;
            utwórz warunkową bazę wzorca β;
            utwórz (tree-β) warunkowe FP-drzewo wzorca β;
            if tree-β ≠ 0 then call FP-growth(tree-β, β);
        end
    end
end procedure;

```

Algorytm zatem nie generuje kandydatów zbiorów częstych, tworzy kompresję danych w postaci drzewa (krok b)) i tylko dwukrotnie przeszukuje zbiór danych D (przy wyszukiwaniu zbiorów częstych 1-elementowych – krok a) oraz przy tworzeniu drzewa w oparciu o wymienione zbiory – krok d)) – jest zatem znacznie szybszy w porównaniu z algorymem Apriori [bor05], [han00b], [pan06]. Jego wadami są natomiast duże obciążenie pamięci (stąd też często duże FP-drzewo może się nie mieścić w pamięci komputera) oraz czasochłonność tworzenia drzewa. Jednakże raz utworzone drzewo pozwala szybko odnaleźć zbiory częste [pan06].

6.3. Wyszukiwanie rozmytych reguł asocjacji

W ostatnim czasie, zaczęto również pracować nad włączeniem teorii zbiorów rozmytych do odkrywania reguł asocjacji dla ilościowych atrybutów tworząc tzw. rozmyte reguły asocjacji (ang. *fuzzy association rules*). Reguły w takiej postaci są łatwiejsze do zrozumienia dla ludzi, gdyż operujemy w nich terminami lingwistycznymi, którymi zwykle operuje się w opisie rzeczywistości. Wówczas rozmyte, wielowymiarowe reguły asocjacji dla atrybutów $(x_1, \dots, x_n) \in D$ mają postać następującą:

$$\theta \Rightarrow \varphi : A_1 \cap \dots \cap A_n \Rightarrow A_{n+1} \cap \dots \cap A_m (s, c), \quad (101)$$

gdzie A_1 (A_n) stanowi parę atrybut-wartość lingwistyczną (zbiór rozmyty), zapisywaną częściej jako " $x_n = A_n$ " lub " x_n jest A_n ", gdzie x_n stanowi analizowany atrybut (wg definicji 4. stanowiący zmienną lingwistyczną), A_n jest wartością lingwistyczną utożsamianą ze zbiorem rozmytym definiowanym wg definicji 3. za pomocą funkcji przynależności $\mu_{A_n}(x_n)$. Zamieniając operacje na zbiorach w ujęciu klasycznym (def. 1) na operacje teoriomnogościowe, miary wsparcia i ufności reguł są obliczane następująco (na podstawie [kuo98], [dub06]):

$$s(\theta \Rightarrow \varphi) = \sum_{(x_1, \dots, x_m) \in D} T[\mu_{A_1}(x_1), \dots, \mu_{A_m}(x_m)], \quad (102)$$

$$c(\theta \Rightarrow \varphi) = \frac{\sum_{(x_1, \dots, x_m) \in D} T[\mu_{A_1}(x_1), \dots, \mu_{A_n}(x_n), \dots, \mu_{A_m}(x_m)]}{\sum_{(x_1, \dots, x_m) \in D} T[\mu_{A_1}(x_1), \dots, \mu_{A_n}(x_n)]}, \quad (103)$$

gdzie: T oznacza wybrany operator t-normy, który zastępowany jest najczęściej operatorem minimum bądź iloczynem algebraicznym.

Delgado i inni [del00] zaproponowali wyliczenie wielkość wsparcia zbiorów rozmytych na bazie ważonej sumy częstości α -przekroju zbioru rozmytego (por. def. 7). Wówczas ufność reguł rozmytych ozn. c_w definiowana jest jako:

$$c_w(\theta \Rightarrow \varphi) = \sum_{i=1}^{t-1} (\alpha_i - \alpha_{i+1}) \frac{|\theta \cap \varphi|_{\alpha_i}}{|\theta|_{\alpha_i}} \quad (104)$$

gdzie: $|...|$ oznacza moc zbioru (def. 8), $\alpha_1=1, \dots, \alpha_t=0$ jest uporządkowaną listą α -poziomów dla zbiorów $\theta \cap \varphi$ i θ . Ta metoda obliczeń faworyzuje wartości zmiennych o większym stopniu przynależności, co związane jest z faktem, że element z przynależnością na poziomie α_k pojawia się w każdym czynniku sumy $k, k+1, \dots, t$ [dub03].

W publikacjach naukowych spotyka się różne podejście do zagadnienia wyszukiwania rozmytych reguł asocjacji. Miller i Yang w [mil97] zaproponowali wyszukiwanie rozmytych reguł asocjacji w oparciu o zdefiniowaną odległość pomiędzy regułami. Częściej jednak metody wyszukiwania reguł asocjacji są oparte na zasadach algorytmu Apriori (por. podrozdział II-5.2.1). Przykładem może być algorytm opisany przez Zhang [zha99] zwany EDPFT (ang. *Equal-Depth Partition with Fuzzy Terms*), który pozwala na połączenie warunków w postaci przedziałów ilościowych i terminów rozmytych w jednej regule, bądź też algorytmy opisane przez Honga i innych w [hon01], [kuo98]. Gyenesei [gye00] zaproponował metodę wyszukiwania rozmytych reguł asocjacji w oparciu o definicję ważonego wsparcia z uwzględnieniem normalizacji i bez normalizacji (ang. *with normalization and without normalization*). Natomiast, Chan i Au [cha97] wprowadzili algorytm o nazwie F-APACS służący do odnajdywania rozmytych reguł asocjacji. W odróżnieniu od innych algorytmów wyszukuje on reguły na bazie analizy obiektywnej miary zwanej różnicą skorygowaną (ang. *adjusted difference*), zamiast warunku minimalnego wsparcia i ufności określonego przez ekspertów. Ponadto, algorytm pozwala na określenie tzw. pozytywnych i negatywnych reguł asocjacji.

Rozszerzenie metod Apriori i FP-Growth do wyszukiwania rozmytych reguł asocjacji z uwzględnieniem predefiniowanych wag reguł przedstawiono odpowiednio w [ols07] oraz [wan09]. Wielowymiarowe rozmyte reguły asocjacji są także tematem publikacji [int09]. Wyszukiwanie rozmytych reguł asocjacji za pomocą grupowania zaprezentowano w [kay02]. Połączenie eksploracji danych metodami rozmytych reguł asocjacji z wykorzystaniem algorytmów genetycznych można znaleźć m.in. w [kay05], [cha10], [alc09], [au99].

7. Wnioski z analizy literaturowej

W rozdziale II przedstawiono elementy teorii zbiorów rozmytych oraz wybrane modele wiedzy systemów rozmytych. Analiza zagadnień wskazuje, iż przez ostatnie kilkadziesiąt lat systemy rozmyte stały się obiektem szeregu badań oraz powszechnych ich

zastosowań. Wciąż możemy obserwować rozwój teorii systemów rozmytych w kierunku tzw. obliczeń miękkich (ang. *soft computing*), gdzie różne formy aparatów oraz hybrydowych metod przetwarzania wiedzy dostarczają narzędzi do modelowania złożonych, niepewnych i nieprecyzyjnie określonych procesów zachodzących w świecie rzeczywistym. To właśnie zastosowanie logiki rozmytej z regułową bazą wiedzy daje w tych systemach możliwość wyrażania informacji niepełnej i niepewnej, pod postacią symbolicznego zapisu matematycznego w języku naturalnym, w sposób charakterystyczny dla człowieka. Istnieje jednakże niewiele rozwiązań, które uwzględniają niepewność informacji w kategoriach rozmytych i probabilistycznych jednocześnie. Przedstawiona w rozdziale II-4.1.4 metoda szablonowego uczenia wskazuje na możliwość włączenia teorii prawdopodobieństwa do systemów rozmytych, ale czyni to w sposób uproszczony – wykonując obliczenia prawdopodobieństwa dla zbiorów w ujęciu klasycznym. Rozwinięcie tej metody spowodowało powstanie probabilistyczno-rozmytego modelu wiedzy, który został szczegółowo opisany w rozdziale II-3.2.3. Jednakże, w literaturze możemy spotkać jedynie rozważania teoretyczne na temat omawianego modelu bądź też krótkie przykłady zastosowań stanowiące tylko wskazówki do zastosowanej metodologii. Brak jest implementacji i badań nad identyfikacją probabilistyczno-rozmytego modelu wiedzy.

Dodanie wiarygodności reguł w postaci prawdopodobieństwa pełnej rozmytej przestrzeni wejść-wyjścia pociąga za sobą złożoność modelu, co wymusza również dopracowanie metod jego identyfikacji. W celu znalezienia odpowiedniej metody, w rozdziale II-4 przeanalizowano szereg sposobów umożliwiających automatyczne pozyskiwanie baz wiedzy w systemach rozmytych. Jednakże, specyficzna forma reguł warunkowych z wagami narzuca pewne ograniczenia dla zastosowanych metod. Stąd też zdecydowano się na nietypowe rozwiązanie – modelowanie z uwzględnieniem rozmytych reguł asocjacji. Zauważono bowiem pewną analogię istot obu podejść: probabilistyczno-rozmytego modelu wiedzy oraz rozmytych reguł asocjacji, w których zmienne możemy traktować jako zmienne losowe o rozkładzie zdarzeń rozmytych. Szczegółowe informację na temat wyszukiwania binarnych oraz rozmytych reguł asocjacji zamieszczono w rozdziale II-5. W literaturze nie znaleziono jednak bezpośrednich metod identyfikacji probabilistyczno-rozmytego modelu wiedzy. Opracowanie metody jest zatem jednym z ważniejszych punktów niniejszej dysertacji.

III. Koncepcja i implementacja systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy

1. Ogólny zarys systemu

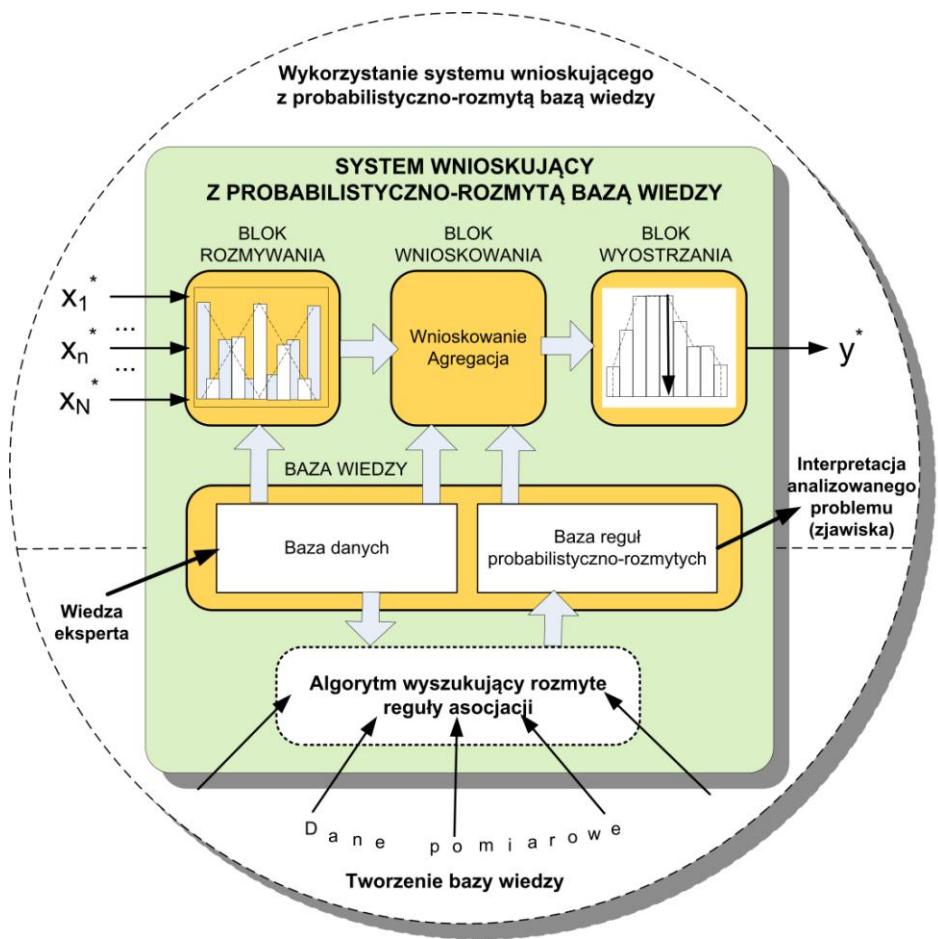
Koncepcja rozważanego w niniejszej dysertacji systemu wnioskującego opiera się na metodologii zapisu bazy wiedzy opisanej w części II-3.2.3 pracy. Analizowane reguły mogą opisywać bazę wiedzy systemu o wielu wejściach i wielu wyjściach, jednakże z uwagi na sposób budowy modelu, przestrzeń rozważań ograniczono do systemów o wielu wejściach i jednym wejściu.

Podejście traktuje proces jako czarną skrzynkę. Pozwala na to wykorzystanie teorii logiki rozmytej, której wybrane zagadnienia przedstawiono w części teoretycznej dysertacji. Użytkownik (ekspert) tworząc dany system, nie ma potrzeby znajomości szczegółowego modelu matematycznego (układu sterowania) analizowanego procesu, jednakże powinien umieć określić zmienne, które należy wziąć pod uwagę chcąc otrzymywać "dobre" wyniki modelowania i wnioskowania rozmytego. Podczas definicji zbiorów rozmytych, niezbędna staje się również umiejętność słownego wyrażenia wiedzy na temat analizowanych zmiennych wejściowych (wyjściowych) modelu. Szczegółowe probabilistyczno-rozmyte reguły modelu oraz jego parametry zostają określone na podstawie danych empirycznych, zawierających informację o rzeczywistym przebiegu danego procesu. System wykorzystywany jest zatem w sytuacji tzw. niepełnej informacji probabilistycznej [świ09], gdy zebrane dane pomiarowe pozwalają na wykorzystanie empirycznych rozkładów prawdopodobieństwa zdarzeń rozmytych.

Propozycję struktury systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy oraz powiązanie systemu z otoczeniem przedstawia rysunek III-1.

Omawiany system wnioskujący składa się z następujących części:

- bazy wiedzy, która zawiera niezbędną wiedzę zapisaną w postaci probabilistycznej i rozmytej jednocześnie, istotną dla rozważanego problemu,
- bloku rozmywania, który zamienia dane wejściowe z dziedziny ilościowej na wielkości jakościowe, reprezentowane przez zbiory rozmyte na podstawie określających je stopni przynależności zapisanych w bazie wiedzy,
- bloku wnioskowania, który korzysta z bazy wiedzy oraz zaimplementowanych metod wnioskowania i agregacji, w celu rozwiązania specjalistycznych problemów,
- bloku wyostrzania, który na podstawie wynikowych stopni przynależności oblicza ilościową (ostrą) wartość na wyjściu systemu.



Rys. III-1. Schemat struktury systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy oraz jego powiązanie z otoczeniem (na podstawie [rud10])

Podstawą bazy wiedzy probabilistyczno-rozmytego modelu są dwa komponenty: baza danych (ang. *data base*) oraz baza probabilistyczno-rozmytych reguł (ang. *probabilistic-fuzzy rule base*).

Baza danych zawiera informacje definiowane przez eksperta (inżyniera wiedzy) z danej dziedziny zastosowania, do których należą wartości lingwistyczne zmiennych rozważanych w bazie reguł oraz definicje zbiorów rozmytych utożsamianych z tymi wartościami.

Baza probabilistyczno-rozmytych reguł, jak wskazuje sama nazwa, zawiera zbiór reguł lingwistycznych z wagami (postać których omówiono szczegółowo w rozdziale II-3.2.3), które mogą być tworzone na podstawie algorytmu generującego pełen rozkład prawdopodobieństwa zdarzeń rozmytych znajdujących się w regule lub na podstawie założeń algorytmu generującego rozmyte reguły asocjacji. Obie metody pozwalają na dopasowanie modelu z wykorzystaniem danych pomiarowych. Charakterystyczna postać reguł, ukazująca empiryczny rozkład prawdopodobieństwa zdarzeń rozmytych, pozwala na łatwą interpretację zawartą w modelu wiedzy oraz umożliwia dodatkową analizę rozważanego zagadnienia. Ponadto, zapis bazy wiedzy wspomaga identyfikację statystycznych własności zmiennych rozmytych wejść i wyjścia systemu, co rozszerza pole możliwych zastosowań systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy.

Omawiany system wnioskujący wykorzystuje analogiczne podejście do definicji zbiorów rozmytych wg Zadeha (def. 3), jednakże zamiast funkcji przynależności pojedynczych

wartości zmiennych stosuje stałe stopnie przynależności, zdefiniowane dla rozłącznych przedziałów wartości zmiennych. Zasadność stosowania stopni przynależności podczas modelowania w omawianym systemie zamieszczono w rozdziale III-2.

Blok wnioskujący, umieszczony w systemie, pozwala na wnioskowanie w oparciu o uogólnioną regułę wnioskowania *modus ponendo ponens*, na podstawie bazy wiedzy oraz nowych faktów, uwzględniając niepewność informacji w kategoriach rozmytych i probabilistycznych jednocześnie [rud11].

Specyfika reprezentacji bazy wiedzy wpływa na możliwość zastosowania systemu do problemów obarczonych niepewnością oraz losowością zachodzących w nich zdarzeń. System może zatem służyć do modelowania problemów obarczonych niepewnością z dziedziny zagadnień technicznych takich jak: sterowanie, predykcja, diagnozowanie, monitorowanie, jak i nietechnicznych takich jak: podejmowanie decyzji czy ocena. Jednakże, możliwość uzyskania wiedzy o prawdopodobieństwie zajścia wyniku przy faktach zapisanych w postaci lingwistycznej bądź ilościowej, predestynuje go szczególnie dla zagadnień z zakresu podejmowania decyzji i diagnostyki.

2. Wybór sposobu rozmywania wartości zmiennych

W rozdziale II-1.1 została szczegółowo przedstawiona definicja zbiorów rozmytych, jaką zaproponował Zadeh [zad65]. Niezależnie od notacji, w jakiej opiszemy zbiór rozmyty A (def. 2-3) jest on definiowany jako zbiór par – element x przestrzeni \mathbb{N} oraz przyporządkowany mu stopień przynależności do danego zbioru rozmytego $\mu_A(x)$. Dla dyskretnej przestrzeni \mathbb{N} zbiór rozmyty może być określony jako:

$$A = \sum_{x \in \mathbb{N}} \mu_A(x) / x, \quad (105)$$

gdzie: $\mu_A(x)$ jest najczęściej ustalany poprzez wybraną funkcję przynależności, której przykłady zostały zamieszczone w tabeli III-1.

Opisany w rozdziale II-3.2.3 model probabilistyczno-rozmyty, wykorzystuje analogiczne podejście do definicji zbiorów rozmytych, jednakże stopnie przynależności określane są nie dla każdej wartości zmiennych, ale dla rozłącznych przedziałów wartości tychże zmiennych. Przedziały powstają poprzez dzielenie zakresu wartości zmiennych, na ustaloną w arbitralny sposób liczbę równych przedziałów, jednakową dla każdej zmiennej systemu. Wówczas, przy założeniu K rozłącznych przedziałów wartości zmiennej $x \in X$, oznaczonych odpowiednio $\{a_1, a_2, \dots, a_K\}$, zbiór rozmyty A definiowany jest następująco:

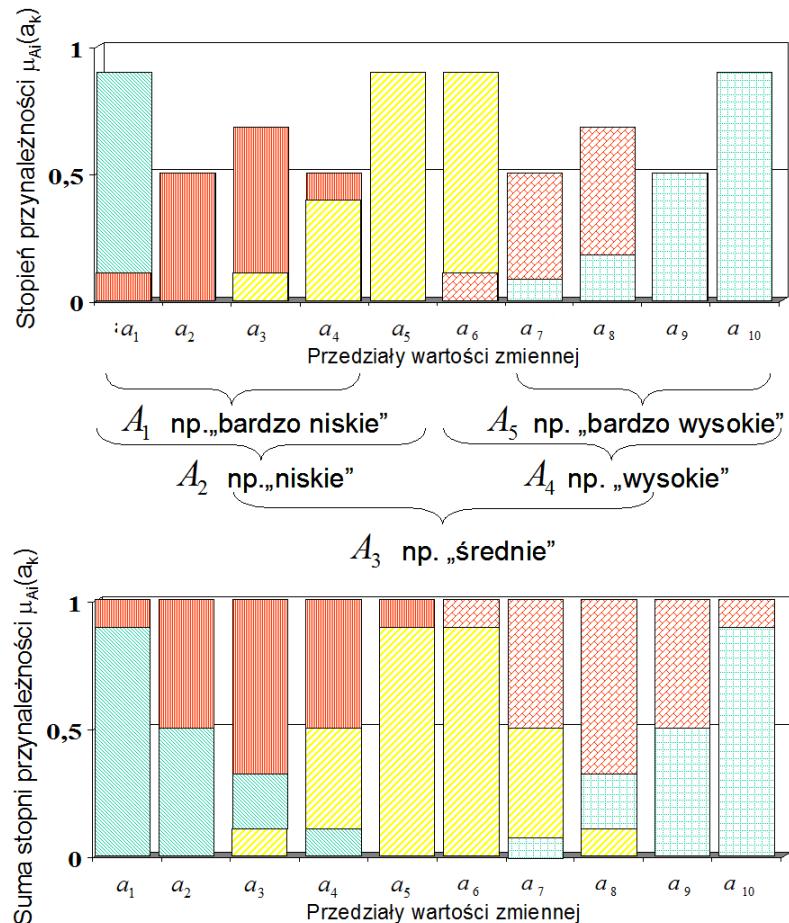
$$A = \sum_{k=1, \dots, K} \mu_A(a_k) / a_k, \quad (106)$$

gdzie: $\mu_A(a_k)$ oznacza stopień przynależności wartości zmiennej z przedziału a_k do zbioru rozmytego A .

W przypadku definicji M zbiorów rozmytych, reprezentowanych przez wartości lingwistyczne, określone na niepustej przestrzeni X , zachowana jest zasada podziału stopni przynależności do jedności wówczas gdy:

$$\sum_{m=1, \dots, M} \mu_A(a_m) = 1, \text{ dla } \forall a_m \in X. \quad (107)$$

Przykład definicji zbiorów rozmytych A_1, \dots, A_5 dla 10 rozłącznych przedziałów zmiennej x , przy zachowaniu zasady podziału do jedności, przedstawia rysunek III-2.

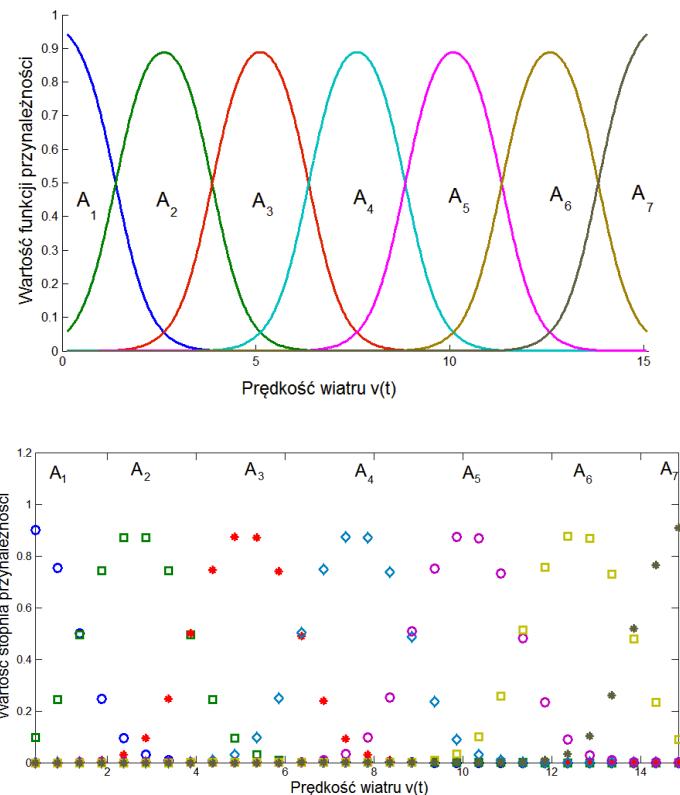


Rys. III-2. Przykład definicji zbiorów rozmytych dla 10 rozłącznych przedziałów wartości zmiennej (na podstawie [rud10])

Problem wyboru sposobu rozmywania wartości zmiennej w systemie wnioskującym z probabilistyczno-rozmytą bazą wiedzy poruszono w [bła10]. Na podstawie procesu predykcji prędkości wiatru porównywano zastosowanie metody definicji zbiorów rozmytych przy użyciu funkcji przynależności dla pojedynczych wartości zmiennych oraz stopni przynależności dla rozłącznych przedziałów wartości zmiennych. Jako algorytmu uczącego bazę wiedzy zastosowano zmodyfikowany algorytm Apriori, którego szczegółowy opis zamieszczony jest w rozdziale III-3.2.2 dysertacji. Algorytm pozwala na wyszukiwanie reguł rozmytych, które spełniają warunek minimalnego wsparcia ($\min w$ – minimalnego prawdopodobieństwa jednoczesnego zajścia zdarzeń w rozmytej regule elementarnej – patrz rozdz. II-5.1).

Predykcję prędkości wiatru $v(t)$ w czasie t dokonywano na podstawie trzech poprzednich wartości prędkości wiatru $v(t-3), v(t-2), v(t-1)$ z okresem pomiaru co 1 minutę. Do opisu wartości lingwistycznych dotyczących parametru prędkości wiatru zastosowano siedem zbiorów rozmytych. Rysunek III-3 przedstawia definicję zbiorów rozmytych z uwzględnieniem znormalizowanej funkcji gaussowskiej (tab. II-3) dla każdej wartości parametru $v(t)$ oraz stopni przynależności dla 30 rozłącznych przedziałów wartości parametru

$v(t)$, przeliczonych na podstawie tych samych funkcji przynależności. Normalizacja wartości funkcji powoduje zachowanie warunku (107). Zbiory rozmyte dla pozostałych zmiennych zostały zdefiniowane analogicznie. Do generowania reguł użyto operatora iloczynu algebraicznego (tab. II-4.), natomiast podczas procesu wnioskowania użyto następujące parametry: iloczyn algebraiczny, jako operator przecięcia zbiorów rozmytych, interpretację reguł wg Mamdaniego oraz metodę wnioskowania środka ciężkości COG (szczegóły odnośnie procesu wnioskowania w systemie z probabilistyczno-rozmytą bazą wiedzy zamieszczone w III-5.5).



Rys. III-3. Definicje zbiorów rozmytych z uwzględnieniem znormalizowanych funkcji przynależności oraz wartości stopni przynależności

Analizowano otrzymaną strukturę modelu wiedzy (liczbę reguł elementarnych), zdolność modelu do odwzorowania procesu rzeczywistego oraz czas obliczeń (zarówno czas generowania reguł, jak i czas wnioskowania przy użyciu odnalezionych reguł).

W rezultacie liczba reguł elementarnych jest niemalże jednakowa dla obu metod definicji zbiorów rozmytych (rys. III-4). Jak przedstawia rysunek III-5, również dopasowanie modeli wiedzy do wartości rzeczywistego procesu prędkości wiatru różni się nieznacznie. Można zauważać, iż celowe jest ograniczenie liczby reguł do wartości ok. 40, aby móc uprościć złożoność modelu, zachowując przy tym taką samą dokładność odwzorowania na poziomie około 0,75 m/s. Dopiero po pewnej wartości minimalnego wsparcia błąd wzrasta, gdyż model wiedzy jest zbyt uproszczony, aby odwzorować prawidłowy przebieg prędkości wiatru. Dopasowanie modelu określono za pomocą wartości błędu stanowiącego pierwiastek błędu średniokwadratowego RMSE (ang. *Root Mean Squared Error*) obliczony na podstawie obu danych uczących i testujących, wg wzoru:

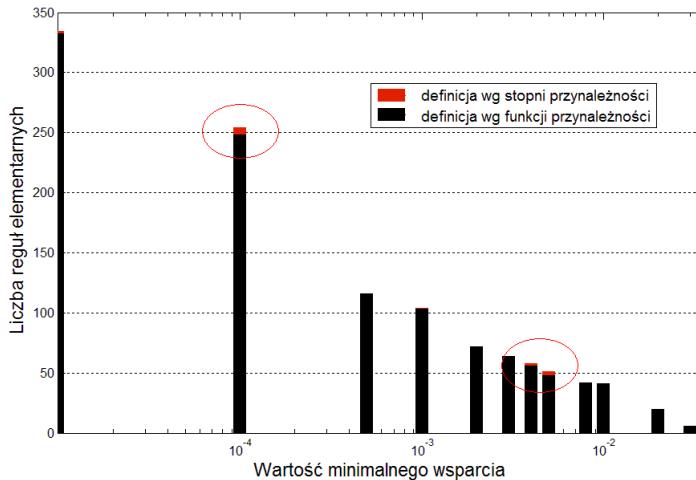
$$RMSE = \sqrt{\frac{\sum_{t=4}^N [v(t) - \hat{v}(t)]^2}{N - 3}} \quad (108)$$

gdzie:

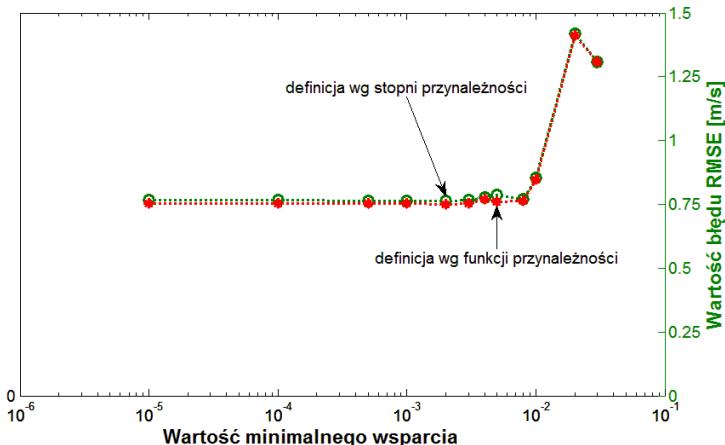
N – liczba elementów zbioru,

$\hat{v}(t)$ – wartość prognozowana w chwili t ,

$v(t)$ – wartość rzeczywista w chwili t .



Rys. III-4. Liczba elementarnych reguł w zależności od wartości minimalnego wsparcia ($min w$)

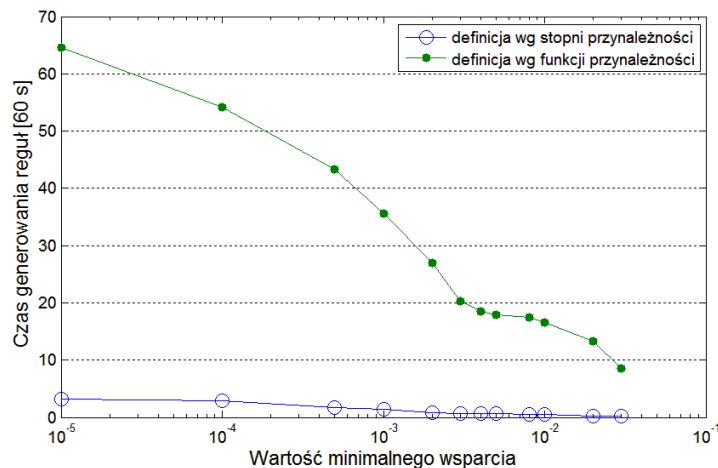


Rys. III-5. Wartości błędów RMSE dla danych uczących i testujących w zależności od parametru minimalnego wsparcia ($min w$)

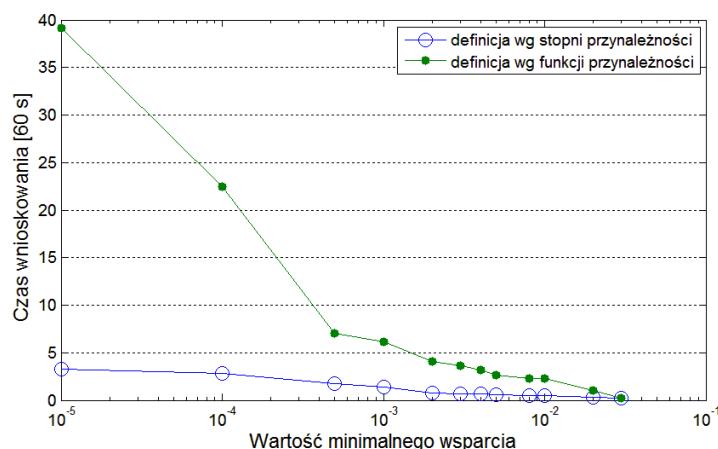
Ponieważ struktura otrzymanego modelu wiedzy oraz związana z tym zdolność do predykcji w danym systemie są zbliżone, kluczowym elementem wyboru sposobu rozmywania wartości zmiennych modelu będzie czas generowania reguł oraz wnioskowania z zastosowaniem obu metod. Na rysunkach III-6 oraz III-7 przedstawiono czas generowania reguł oraz czas wnioskowania w oparciu o wygenerowany model, dla obu metod, w zależności od wartości minimalnego wsparcia. Okazuje się, że czas generowania reguł i czas wnioskowania jest znacznie dłuższy dla metody rozmywania wartości z użyciem standardowych funkcji

przynależności. Różnica czasów jest tym większa, im mniejszy jest współczynnik minimalnego wsparcia dla reguły, czyli im więcej jest reguł w modelu o mniejszej częstotliwości występowania. Dzieje się tak dlatego, że w przypadku użycia funkcji przynależności do rozmywania wartości zmiennych, mamy do czynienia z algorytmem iteracyjnym. Natomiast, w przypadku użycia stałych stopni przynależności dla rozłącznych przedziałów istnieje możliwość wykorzystania wektoryzacji obliczeń (por. rozdz. III-3.2.2), co ma duże znaczenie w skracaniu czasu trwania obliczeń przy implementacji w środowisku Matlab. Zatem analiza udowadnia zasadność stosowania stopni przynależności dla rozłącznych przedziałów wartości zmiennych podczas modelowania rozmytego z użyciem zmodyfikowanego algorytmu Apriori. W rozdziale III-3 zostały przedstawione jeszcze inne metody generowania probabilistyczno-rozmytej bazy wiedzy, ale ponieważ rozważany algorytm stanowił punkt zaczepienia w tworzeniu modelu wiedzy zdecydowano się na definicję zbiorów rozmytych z użyciem rozłącznych przedziałów dyskretyzacji wartości zmiennych.

Wadą założonego podejścia jest obowiązek umiejętności dobrania liczby rozłącznych przedziałów wartości zmiennych. Zbyt mała liczba przedziałów w stosunku do liczby zbiorów rozmytych może uniemożliwić zapisanie wiedzy eksperta w postaci zmiennych lingwistycznych.



Rys. III-6. Czas generowania bazy reguł w zależności od wartości minimalnego wsparcia ($\text{min } w$)



Rys. III-7. Czas wnioskowania na podstawie danych uczących w zależności od wartości minimalnego wsparcia ($\text{min } w$)

3. Opracowanie algorytmu generowania reguł probabilistyczno-rozmytej bazy wiedzy

Elementem strategicznym w strukturze systemu wnioskującego jest baza wiedzy, która stanowi główną część "inteligencji" obliczeniowej. Umiejętność prawidłowego zaprojektowania bazy wiedzy jest bardzo ważnym etapem w budowie takiego systemu. Stąd też, cały rozdział III-3 będzie poświęcony analizie nad możliwymi sposobami tworzenia modelu rozmytego systemu typu MISO o N -wejściach $\{x_1, \dots, x_N\}$ i jednym wyjściu y , opartego na O regułach plikowych w postaci (na podstawie wzoru (46)):

$$\begin{aligned}
 & \text{Jeżeli } x_1 \text{ jest } A_o^{(1)} \text{ I } \dots \text{ I } x_N \text{ jest } A_o^{(N)}, \text{ z wagą } w_o \\
 & \quad \text{To } y \text{ jest } B_{1/o} \text{ z wagą } w_{1/o} \\
 & \quad \dots \\
 & \quad \text{Także } y \text{ jest } B_{l/o} \text{ z wagą } w_{l/o} \\
 & \quad \dots \\
 & \quad \text{Także } y \text{ jest } B_{L/o} \text{ z wagą } w_{L/o}, \tag{109}
 \end{aligned}$$

gdzie:

o – numer reguły plikowej,

N – liczba zmiennych wejściowych modelu,

$A_o^{(1)}, \dots, A_o^{(N)}, B_{l/o}$ – zbiory rozmyte reprezentujące wartości lingwistyczne zmiennych wejściowych x_1, \dots, x_N i zmiennej wyjściowej y w l -tej regule elementarnej o -tej reguły plikowej,

w_o – waga reguły plikowej stanowiąca prawdopodobieństwo jednoczesnego zajścia zdarzeń $(x_1 \text{ jest } A_o^{(1)}) \cap \dots \cap (x_N \text{ jest } A_o^{(N)})$ w przesłance reguły,

$w_{l/o}$ – waga reguły elementarnej, stanowiąca prawdopodobieństwo warunkowe zajścia zdarzenia $(y \text{ jest } B_{l/o})$ w konkluzji reguły, przy wcześniejszym zajęciu zdarzeń w przesłance reguły $(x_1 \text{ jest } A_o^{(1)}) \cap \dots \cap (x_N \text{ jest } A_o^{(N)})$.

Analiza dotycząca algorytmów ma na celu ustalenie metody (metod) generowania reguł, która w konsekwencji będzie zaimplementowana w module narzędziowym opracowanym w środowisku Matlab. Pod rozwagę bierze się zarówno sposób przeszukiwania przestrzeni wartości zmiennych, strukturę otrzymanego modelu, jak i czas tworzenia modelu oraz czas wnioskowania w oparciu o dany model.

Rozpatrując bazę wiedzy w postaci (109) zmienne wejściowe i zmienną wyjściową możemy traktować jako lingwistyczne zmienne losowe o określonym zbiorze wartości lingwistycznych. Wówczas, analizując zbiór zmiennych losowych zależnych od czasu $t \in (t_a, t_b)$ [mań71] wraz ze strukturą ich prawdopodobieństwa (wag modelu), system pozwala na badanie *procesów stochastycznych*. Jednakże, aby model prawidłowo odzwierciedlał proces stochastyczny i miał dobre własności uogólniające niezbędne staje się założenie *stacjonarności procesu stochastycznego*. Oznacza to, że wielowymiarowa gęstość prawdopodobieństwa wartości zmiennych $x_1(t_1 + t_0), \dots, x_1(t_n + t_0), \dots, x_N(t_1 + t_0), \dots, x_N(t_n + t_0)$ oraz $y(t_1 + t_0), \dots, y(t_n + t_0)$ powinna być zależna tylko od t_1, \dots, t_n i nie powinna zależeć od t_0 [mań83]. Zazwyczaj zakłada się stacjonarność określoną w szerszym sensie, wówczas

wartość średnia oraz wariancja procesu stacjonarnego winna się charakteryzować stałością względem czasu a funkcja kowariancji własnej winna być zależna jedynie od różnicy czasów $\tau = t_2 - t_1$. W niniejszej dysertacji zakłada się stacjonarność badanych procesów.

3.1. Algorytm generujący pełny rozkład prawdopodobieństwa zdarzeń rozmytych

Jako podstawę do rozważań na temat metody tworzenia rozmytych reguł z wagami dotyczącymi prawdopodobieństwa zdarzeń rozmytych (wzór (109)), należy wziąć pod uwagę algorytm generujący pełen rozkład prawdopodobieństwa tychże zdarzeń. Z uwagi na przeszukiwanie całej przestrzeni zmiennych, zwany jest również w dysertacji w skrócie algorytmem naiwnym. Otrzymany w modelu rozkład prawdopodobieństwa jest rozkładem empirycznym, wyznaczonym w oparciu o zbiór danych doświadczalnych T_d . Proponowaną metodologię obliczeń można znaleźć w pozycjach [wb07], [wb08a] oraz rozdziale II-3.2.3 dysertacji.

Rozważmy system typu MISO o zmiennych wejściowych $x_1 \in \aleph_1, x_2 \in \aleph_2$ i zmiennej wyjściowej $y \in \mathfrak{I}$. Zastosujemy definicję zbiorów rozmytych dla wejść i wyjścia, będących zmiennymi lingwistycznymi, jako stopnie przynależności dla K rozłącznych przedziałów tychże zmiennych. Przestrzenie rozważań $\aleph_1, \aleph_2, \mathfrak{I}$ zmiennych x_1, x_2 i y podzielono zatem na przedziały:

$$\begin{aligned} a_{k_1}^{(1)} &= \langle x_{k_{1\min}}, x_{k_{1\max}} \rangle, k_1 = 1, 2, \dots, K, \\ a_{k_2}^{(2)} &= \langle x_{k_{2\min}}, x_{k_{2\max}} \rangle, k_2 = 1, 2, \dots, K, \\ b_{k_3} &= \langle y_{k_{3\min}}, y_{k_{3\max}} \rangle, k_3 = 1, 2, \dots, K. \end{aligned} \quad (110)$$

Zdarzeniem elementarnym w przestrzeni Ω_{x1} jest spełnienie warunku $x_1 \in a_{k_1}^{(1)}$, zdarzeniem elementarnym w przestrzeni Ω_{x2} jest spełnienie warunku $x_2 \in a_{k_2}^{(2)}$, natomiast w przestrzeni Ω_y – warunku $y \in b_{k_3}$. Biorąc pod uwagę przestrzeń $\Omega_{x1} \times \Omega_{x2} \times \Omega_y$ zdarzeń elementarnych, prawdopodobieństwo (wg def. 24) jednoczesnego zajścia zdarzeń $x_1 \in a_{k_1}^{(1)}, x_2 \in a_{k_2}^{(2)}, y \in b_{k_3}$ wynosi:

$$P(x_1 \in a_{k_1}^{(1)}, x_2 \in a_{k_2}^{(2)}, y \in b_{k_3}) = p_{k_1 k_2 k_3}(x_1, x_2, y) = \frac{n_{k_1 k_2 k_3}}{N_{T_d}}; k_1, k_2, k_3 = 1, 2, \dots, K, \quad (111)$$

gdzie:

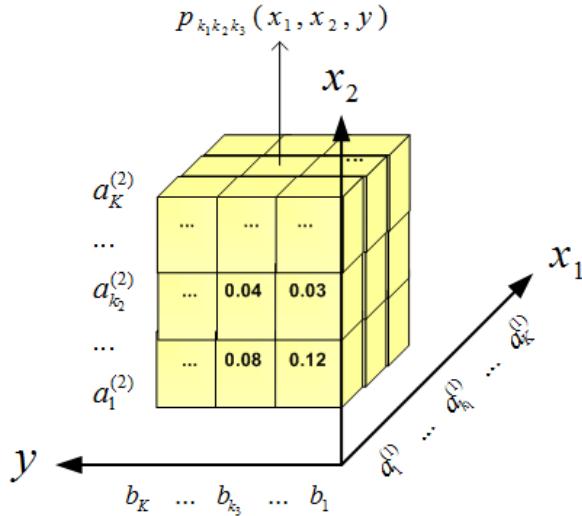
$n_{k_1 k_2 k_3}$ – jest liczbą przypadków w zbiorze danych doświadczalnych T_d , kiedy wartości zmiennych x_1, x_2 i y mieszczą się w zakresie odpowiednich przedziałów $a_{k_1}^{(1)}, a_{k_2}^{(2)}, b_{k_3}$,
 N_{T_d} – liczbą wszystkich pomiarów w zbiorze danych doświadczalnych T_d .

Definiując pełen dyskretny rozkład prawdopodobieństwa (111) zachodzi zależność:

$$\sum_{k_1=1}^K \sum_{k_2=1}^K \sum_{k_3=1}^K \frac{n_{k_1 k_2 k_3}}{N_{T_d}} = p_{k_1 k_2 k_3}(x_1, x_2, y) = 1. \quad (112)$$

Biorąc pod uwagę implementację systemu wnioskującego, dyskretny rozkład prawdopodobieństwa zdarzeń (111) jest $N+1$ -wymiarową macierzą $E^p = [p_{k_1 \dots k_{N+1}}(x_1, x_2, \dots, x_N, y)]_{K \times K \times \dots \times K}$, gdzie N stanowi liczbę wejść systemu. W rozważanym przypadku, rozkład prawdopodobieństwa zajścia jednokrotnego zdarzeń dla zmiennych wejściowych x_1, x_2 i wyjściowej y stanowi macierz trójwymiarową o elementach równych $p_{k_1 k_2 k_3}(x_1, x_2, y)$ (rys. III-8):

$$E^p = [p_{k_1 k_2 k_3}(x_1, x_2, y)]_{K \times K \times K}. \quad (113)$$



Rys. III-8. Graficzne przedstawienie rozkładu prawdopodobieństwa zajścia jednokrotnego zdarzeń dla zmiennych wejściowych x_1, x_2 i wyjściowej y (macierz E^p)

Definiując zmienną wejściową (wyjściową) jako zmienną lingwistyczną niezbędne staje się określenie dla nich zbiorów rozmytych oraz wartości lingwistycznych. Niech każda zmienna lingwistyczna wejścia x_n ($n=1,2$) i wyjścia y jest określona przez L wartości lingwistycznych reprezentowanych przez zbiory rozmyte następująco:

$$\{A_{l_n}^{(n)}, A_{2_n}^{(n)}, \dots, A_{L_n}^{(n)}\} = \{\sum_{k_n=1, \dots, K} \mu_{A_{l_n}^{(n)}}(a_{k_n}^{(n)}) / a_{k_n}^{(n)}, \sum_{k_n=1, \dots, K} \mu_{A_{2_n}^{(n)}}(a_{k_n}^{(n)}) / a_{k_n}^{(n)}, \dots, \sum_{k_n=1, \dots, K} \mu_{A_{L_n}^{(n)}}(a_{k_n}^{(n)}) / a_{k_n}^{(n)}\}, \quad (114)$$

$$\{B_1, B_2, \dots, B_L\} = \{\sum_{k_3=1, \dots, K} \mu_{B_1}(b_{k_3}) / b_{k_3}, \sum_{k_3=1, \dots, K} \mu_{B_2}(b_{k_3}) / b_{k_3}, \dots, \sum_{k_3=1, \dots, K} \mu_{B_L}(b_{k_3}) / b_{k_3}\}, \quad (115)$$

gdzie: $A_{l_n}^{(n)}$ ($l_n=1, \dots, L$) stanowi l -ty zbiór rozmyty na n -tej zmienniej wejściowej, B_l – l -ty zbiór rozmyty dla zmiennej wyjściowej, $\mu_{A_{l_n}^{(n)}}(a_{k_n}^{(n)})$ – wartość stopnia przynależności, z jakim zmienna $x_n \in a_{k_n}^{(n)}$ przynależy do zbioru rozmytego $A_{l_n}^{(n)}$, natomiast znak \sum symbolizuje sumę mnogościową.

Niech zajście warunków $x_n = A_{l_n}^{(n)}$, $y = B_l$ stanowi zdarzenie rozmyte w sensie definicji 25. Wagi w_o modelu (109) stanowią wówczas odpowiednie prawdopodobieństwa brzegowe

w łącznym rozkładzie prawdopodobieństwa zdarzenia rozmytego $A_{l_1}^{(1)} \cap A_{l_2}^{(2)} \cap B_{l_3}$ ($l_1, l_2, l_3=1, \dots, L$) (por. (47)). Natomiast wagi $w_{l/o}$ stanowią prawdopodobieństwa warunkowe (por. (48)) w tymże rozkładzie. Wg [wb07] oraz definicji prawdopodobieństwa Zadeha (def. 26), pełny rozkład prawdopodobieństwa zdarzenia rozmytego $A_{l_1}^{(1)} \cap A_{l_2}^{(2)} \cap B_{l_3}$ ($l_1, l_2, l_3=1, \dots, L$), można obliczyć jako (na podstawie (50)):

$$\begin{aligned} & \forall_{l_1, l_2, l_3=1, \dots, L} P(A_{l_1}^{(1)} \cap A_{l_2}^{(2)} \cap B_{l_3}) = \\ & \sum_{a_{k_1}^{(1)}, a_{k_2}^{(2)}, b_{k_3} \in X_1 \times X_2 \times Y} [p_{k_1 k_2 k_3}(x_1, x_2, y) \cdot T(\mu_{A_{l_1}^{(1)}}(a_{k_1}^{(1)}), \mu_{A_{l_2}^{(2)}}(a_{k_2}^{(2)}), \mu_{B_{l_3}}(b_{k_3}))], \end{aligned} \quad (116)$$

gdzie: $p_{k_1 k_2 k_3}(x_1, x_2, y)$ wynika z (111), przy warunku:

$$\sum_{l_1, l_2, l_3=1, \dots, L} P(A_{l_1}^{(1)} \cap A_{l_2}^{(2)} \cap B_{l_3}) = 1. \quad (116a)$$

W systemie wnioskującym do operacji iloczynu zbiorów rozmytych podczas generowania reguł (109) dopuszcza się wykorzystanie operatorów t-normy zamieszczonej w tabeli III-1. Jednakże w dysertacji ograniczono się do wykorzystania operatora t-normy iloczynu algebraicznego.

Tab. III-1. Operatory t-normy, które mogą być użyte do generowania reguł modelu probabilistyczno-rozmytego [pie03], [łęs08], [cpa09]

Nazwa operatora T-normy	Operator t-normy
T-norma Zadeha (min)	$T_m(\mu_A(a), \mu_B(a)) = MIN(\mu_A(a), \mu_B(a))$
T-norma Łukasiewicza	$T_L(\mu_A(a), \mu_B(a)) = MAX(\mu_A(a) + \mu_B(a) - 1, 0)$
T-norma Fodora	$T_F(\mu_A(a), \mu_B(a)) = \begin{cases} MIN(\mu_A(a), \mu_B(a)), dla \mu_A(a) + \mu_B(a) > 1 \\ 0, dla \mu_A(a) + \mu_B(a) \leq 1 \end{cases}$
T-norma iloczynu algebraicznego	$T_a(\mu_A(a), \mu_B(a)) = \mu_A(a) \cdot \mu_B(a)$
T-norma Hamachera	$T_H(\mu_A(a), \mu_B(a)) = \frac{\mu_A(a) \cdot \mu_B(a)}{\mu_A(a) + \mu_B(a) - \mu_A(a) \cdot \mu_B(a)}$
T-norma Einsteina	$T_E(\mu_A(a), \mu_B(a)) = \frac{\mu_A(a) \cdot \mu_B(a)}{2 - (\mu_A(a) + \mu_B(a) - \mu_A(a) \cdot \mu_B(a))}$
T-norma drastyczna	$T_d(\mu_A(a), \mu_B(a)) = \begin{cases} MIN(\mu_A(a), \mu_B(a)) dla MAX(\mu_A(a), \mu_B(a)) = 1 \\ 0 dla MAX(\mu_A(a), \mu_B(a)) \neq 1 \end{cases}$

Obliczanie pełnego rozkładu prawdopodobieństwa (116) jest operacją na której ciąży złożoność obliczeniowa całego procesu generowania warunkowych reguł modelu probabilistyczno-rozmytego. Aby zmniejszyć czas obliczania rozkładu, zaproponowano w implementacji algorytmu zastosowanie operacji tablicowych wykonywanych na macierzach. Wówczas, każdy wektor $[\mu_{A_{l_n}^{(n)}}(a_1^{(n)}), \dots, \mu_{A_{l_n}^{(n)}}(a_{k_n}^{(n)}), \dots, \mu_{A_{l_n}^{(n)}}(a_K^{(n)})]$, zawierający stopnie przynależności rozłącznych przedziałów $a_{k_n}^{(n)}$ ($k_n = 1, \dots, K$) wartości zmiennej x_n do

danego zbioru rozmytego $A_{l_n}^{(n)}$, jest zamieniany na macierz trójwymiarową ($N+1$ -wymiarową dla N zmiennych wejściowych) $E_{A_{l_n}^{(n)}}^{(n)} = [e_{k_1 k_2 k_3}]_{K \times K \times K}$ z elementami macierzy $e_{k_1 k_2 k_3} (k_1, k_2, k_3 = 1, \dots, K)$ o wartościach:

$$\forall_{k_1, k_2, k_3=1, \dots, K} e_{k_1 k_2 k_3} = \mu_{A_{l_n}^{(n)}}(a_{k_n}^{(n)}), \quad (117)$$

gdzie: n oznacza numer zmiennej wejściowej, dla jakiej wyznaczana jest macierz.

Analogicznie, dla zmiennej wyjściowej y , każdy wektor $[\mu_{B_{l_3}}(b_1), \dots, \mu_{B_{l_3}}(b_k), \dots, \mu_{B_{l_3}}(b_K)]$, zawierający stopnie przynależności rozłącznych przedziałów $b_{k_3} (k_3 = 1, \dots, K)$ wartości zmiennej do danego zbioru rozmytego B_{l_3} , jest zamieniany na macierz trójwymiarową $E_{B_{l_3}}^{(n+1)} = [e_{k_1 k_2 k_3}]_{K \times K \times K}$ z elementami macierzy $e_{k_1 k_2 k_3}$ o wartościach:

$$\forall_{k_1, k_2, k_3=1, \dots, K} e_{k_1 k_2 k_3} = \mu_{B_{l_3}}(b_{k_3}). \quad (118)$$

Kontynuując rozważania, pełny rozkład prawdopodobieństwa (116) można uzyskać stosując właściwy operator t-normy do operacji tablicowych, czyli operacji wykonywanych na poszczególnych elementach utworzonych macierzy, następująco:

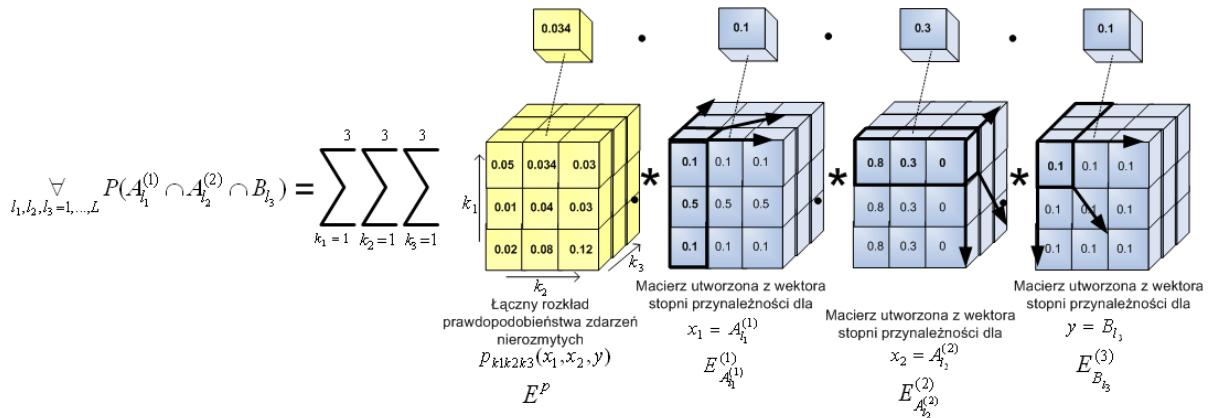
$$\forall_{l_1, l_2, l_3=1, \dots, L} P(A_{l_1}^{(1)} \cap A_{l_2}^{(2)} \cap B_{l_3}) = \sum_{a_{k_1}^{(1)}, a_{k_2}^{(2)}, b_{k_3} \in X_1 \times X_2 \times Y} E^p \cdot * T^*(E_{A_{l_1}^{(1)}}^{(1)}, E_{A_{l_2}^{(2)}}^{(2)}, E_{B_{l_3}}^{(3)}), \quad (119)$$

gdzie: E^p stanowi macierz (113), $E_{A_{l_n}^{(n)}}^{(n)}$ – macierze o elementach (117), $E_{B_{l_3}}^{(3)}$ – macierz o elementach (118), „ $\cdot *$ ” oznacza operator mnożenia tablicowego, natomiast T^* – operator t-normy wykonywany z wykorzystaniem operacji tablicowych.

W szczególności, dla operatora t-normy, jako iloczynu algebraicznego, pełny rozkład prawdopodobieństwa zdarzeń znajdujących się w przesłance i konkluzji reguł warunkowych obliczany jest ze wzoru:

$$\forall_{l_1, l_2, l_3=1, \dots, L} P(A_{l_1}^{(1)} \cap A_{l_2}^{(2)} \cap B_{l_3}) = \sum_{a_{k_1}^{(1)}, a_{k_2}^{(2)}, b_{k_3} \in X_1 \times X_2 \times Y} E^p \cdot * E_{A_{l_1}^{(1)}}^{(1)} \cdot * E_{A_{l_2}^{(2)}}^{(2)} \cdot * E_{B_{l_3}}^{(3)}. \quad (120)$$

Schemat obliczeń zamieszczono na rysunku III-9.



Rys. III-9. Schematyczne przedstawienie obliczeń (120)

Waga w_o dla reguły plikowej (109) jest wyznaczana na podstawie prawdopodobieństwa zajścia zdarzeń rozmytych w przesłance reguły, które jest prawdopodobieństwem brzegowym w łącznym rozkładzie prawdopodobieństwa zmiennych wejścia-wyjścia:

$$w_o = P(A_o^{(1)} \cap A_o^{(2)}) = \sum_{l=1, \dots, L} P(A_o^{(1)} \cap A_o^{(2)} \cap B_{l/o}), \quad (121)$$

gdzie: indeksy o ($o=1, \dots, O$) i l/o oznaczają odpowiednio zbiór rozmyty w o -tej regule plikowej oraz l -tej regule elementarnej o -tej reguły plikowej.

Natomiast waga $w_{l/o}$, jako prawdopodobieństwo zajścia zdarzenia rozmytego $B_{l/o}$ w konkluzji pod warunkiem zajścia zdarzenia rozmytego ($A_o^{(1)} \cap A_o^{(2)}$) w przesłance reguły, jest wyznaczana z wzoru (na podstawie (35) oraz (51)):

$$w_{l/o} = P(B_{l/o} | A_o^{(1)} \cap A_o^{(2)}) = \frac{P(A_o^{(1)} \cap A_o^{(2)} \cap B_{l/o})}{P(A_o^{(1)} \cap A_o^{(2)})}. \quad (122)$$

Analogicznie można wyznaczyć model dla systemu o większej ilości zmiennych wejściowych oraz różnej liczbie wartości lingwistycznych określonych dla każdej zmiennej.

PRZYKŁAD

Generowanie probabilistyczno-rozmytego modelu wiedzy według opisanego algorytmu zostanie przetestowane dla różnych parametrów modelu (takich jak liczba zmiennych, liczba zbiorów rozmytych dla poszczególnych zmiennych, liczba rozłącznych przedziałów dyskretyzacji wartości zmiennych) na przykładzie predykcji parametru węgla, jakim jest masowy udział gęstościowej frakcji lekkiej węgla w badanej próbce (w skrócie zwany udziałem frakcji lekkiej węgla). Specyfikację parametrów węgla, reprezentującą jego właściwości dla procesów przeróbczych przedstawiono w rozdziale IV-1. W przykładzie przedstawione zostaną wyniki mające na celu zbadanie poprawności obliczeń i struktury otrzymanego modelu wiedzy oraz porównanie czasu generowania bazy reguł dla różnych parametrów.

Analizie podlega 491 pomiarów udziału frakcji lekkiej węgla, które są traktowane jako dane uczące. Zakłada się, że pomiary parametru węgla $x(t)$ stanowią dyskretny proces stochastyczny autoregresji II rzędu (AR(2)) i III rzędu (AR(3)), z krokiem $\tau=1$. Przestrzeń wartości zmiennej $x(t) \in [0.24, 0.57]$ została początkowo podzielona na 40 rozłącznych

przedziałów o szerokości $0.00825 - a_i$, gdzie $i=1,\dots,40$. Dla każdej zmiennej zdefiniowano J równomiernie rozłożonych zbiorów rozmytych A_j o stopniach przynależności $\mu_{A_j}(a_i)$ przeliczanych z funkcji trójkątnych, gdzie $J=3,\dots,9$, $j=1,\dots,J$. Przykład stopni przynależności zmiennej $x(t)$ dla $J=3,5,7$ zbiorów rozmytych przedstawia tabela III-2. Z uwagi na właściwości modelu dynamicznego, zbiory rozmyte dla pozostałych zmiennych modelu są definiowane analogicznie. Podczas generowania reguł, w celu utworzenia iloczynu zbiorów rozmytych w przesłance rozmytej reguły użyto iloczynu algebraicznego, jako operatora t-normy.

Tab. III-2. Wartości stopni przynależności $\mu_{A_j}(a_i)$ ($i=1,\dots,40$) zbiorów rozmytych dla udziału frakcji lekkiej węgla gdzie: a) $j=1,\dots,3$, b) $j=1,\dots,5$, c) $j=1,\dots,7$

a)

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}
$\mu_{A1}(a_i)$	0,98	0,93	0,88	0,83	0,79	0,74	0,69	0,64	0,60	0,55	0,50	0,45	0,40	0,36	0,31	0,26	0,21	0,17	0,12	0,07
$\mu_{A2}(a_i)$	0,02	0,07	0,12	0,17	0,21	0,26	0,31	0,36	0,40	0,45	0,50	0,55	0,60	0,64	0,69	0,74	0,79	0,83	0,88	0,93
$\mu_{A3}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{28}	a_{29}	a_{30}	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	a_{37}	a_{38}	a_{39}	a_{40}
$\mu_{A1}(a_i)$	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A2}(a_i)$	0,98	0,97	0,93	0,88	0,83	0,78	0,74	0,69	0,64	0,59	0,54	0,50	0,45	0,40	0,35	0,31	0,26	0,21	0,16	0,11
$\mu_{A3}(a_i)$	0,00	0,03	0,07	0,12	0,17	0,22	0,26	0,31	0,36	0,41	0,46	0,50	0,55	0,60	0,65	0,69	0,74	0,79	0,84	0,89

b)

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}
$\mu_{A1}(a_i)$	0,96	0,86	0,76	0,67	0,57	0,48	0,38	0,29	0,19	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A2}(a_i)$	0,04	0,14	0,24	0,33	0,43	0,52	0,62	0,71	0,81	0,90	1,00	0,90	0,81	0,71	0,62	0,52	0,43	0,33	0,24	0,14
$\mu_{A3}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,19	0,29	0,38	0,48	0,57	0,67	0,76	0,86	
$\mu_{A4}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A5}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{28}	a_{29}	a_{30}	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	a_{37}	a_{38}	a_{39}	a_{40}
$\mu_{A1}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A2}(a_i)$	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A3}(a_i)$	0,96	0,95	0,85	0,76	0,66	0,57	0,47	0,37	0,28	0,18	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A4}(a_i)$	0,00	0,05	0,15	0,24	0,34	0,43	0,53	0,63	0,72	0,82	0,91	0,99	0,90	0,80	0,71	0,61	0,51	0,42	0,32	0,23
$\mu_{A5}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,10	0,20	0,29	0,39	0,49	0,58	0,68	0,77	

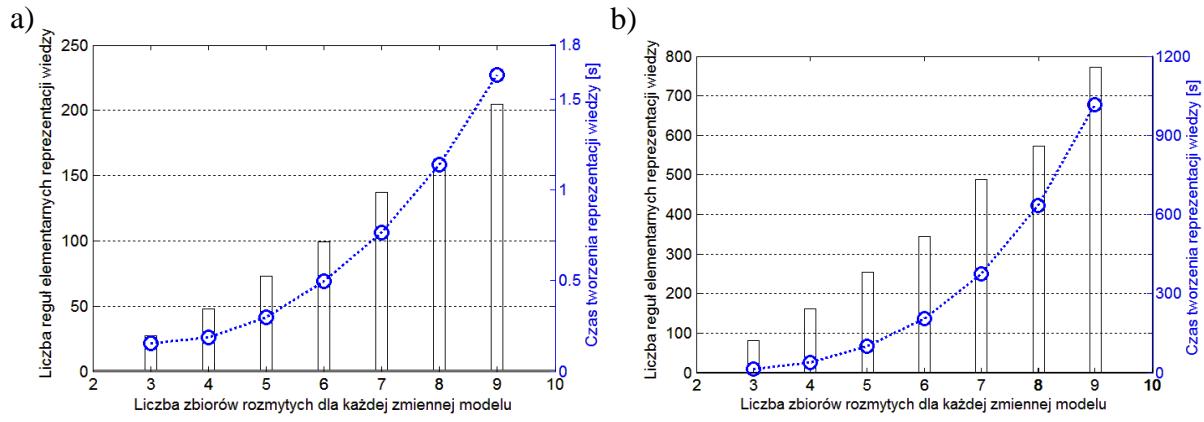
Tab. III-2. c.d. Wartości stopni przynależności $\mu_{A_j}(a_i)$ ($i=1,\dots,40$) zbiorów rozmytych dla udziału frakcji lekkiej węgla gdzie: a) $j=1,\dots,3$, b) $j=1,\dots,5$, c) $j=1,\dots,7$

c)

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}
$\mu_{A1}(a_i)$	0,93	0,79	0,65	0,50	0,36	0,22	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A2}(a_i)$	0,07	0,21	0,35	0,50	0,64	0,78	0,93	0,93	0,79	0,64	0,50	0,36	0,21	0,07	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A3}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,21	0,36	0,50	0,64	0,79	0,93	0,93	0,78	0,64	0,50	0,35	0,21
$\mu_{A4}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,22	0,36	0,50	0,65	0,79	
$\mu_{A5}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A6}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A7}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{28}	a_{29}	a_{30}	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	a_{37}	a_{38}	a_{39}	a_{40}
$\mu_{A1}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A2}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A3}(a_i)$	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A4}(a_i)$	0,93	0,92	0,78	0,64	0,49	0,35	0,21	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mu_{A5}(a_i)$	0,00	0,08	0,22	0,36	0,51	0,65	0,79	0,94	0,92	0,78	0,63	0,49	0,35	0,20	0,06	0,00	0,00	0,00	0,00	0,00
$\mu_{A6}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,22	0,37	0,51	0,65	0,80	0,94	0,92	0,77	0,63	0,49	0,34
$\mu_{A7}(a_i)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,23	0,37	0,51	0,66	

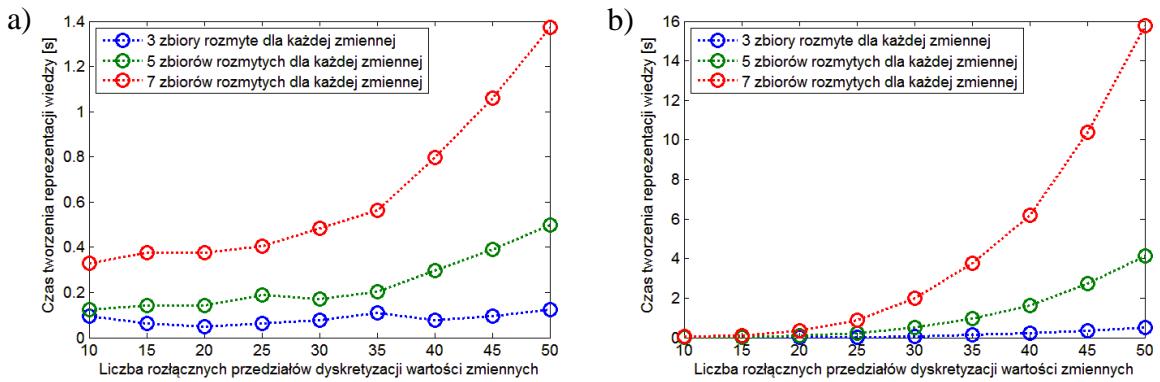
Zamieszczony na rysunku III-10 wykres przedstawia czas tworzenia bazy reguł oraz otrzymaną liczbę reguł elementarnych, w zależności od liczby zbiorów rozmytych, jakie zostały zdefiniowane dla poszczególnych zmiennych procesów autoregresji AR(2) i AR(3). Z analizy wykresów można wyciągnąć następujące wnioski:

- im więcej zostało zdefiniowanych zbiorów rozmytych, tym model wiedzy zawiera więcej reguł elementarnych o niezerowych wagach, wobec tego także czas generowania reguł tego modelu ulega wydłużeniu,
- im model zawiera więcej zmiennych wejściowych tym jego złożoność, mierzona liczbą reguł elementarnych, wzrasta a wraz ze wzrostem liczby zbiorów rozmytych dla każdej zmiennej, liczba reguł elementarnych wzrasta jeszcze szybciej.



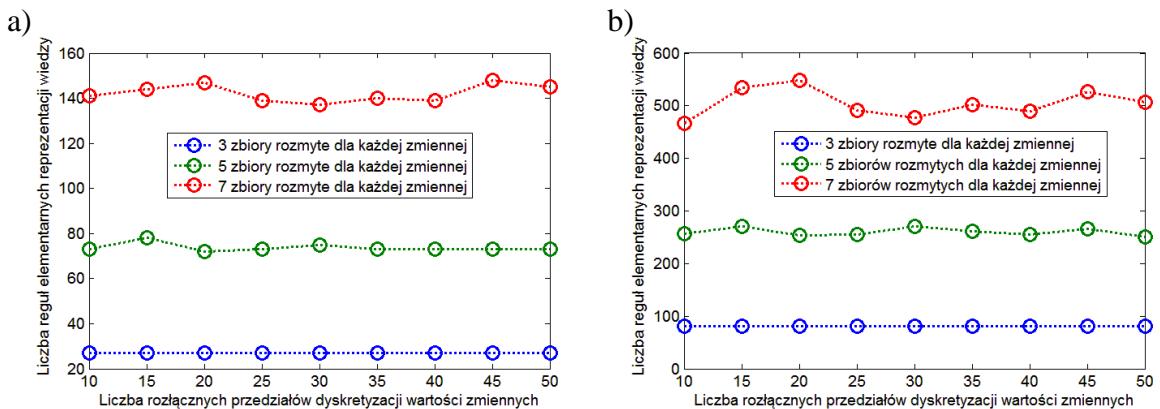
Rys. III-10. Liczba reguł elementarnych oraz czas tworzenia reprezentacji wiedzy w zależności od ilości zdefiniowanych zbiorów rozmytych dla każdej zmiennej w modelach: a) AR(2) b) AR(3)

Rysunki III-11 oraz III-12 zawierają wykresy przedstawiające czasy tworzenia reguł dla modeli AR(2) i AR(3) oraz liczbę otrzymanych reguł elementarnych z niezerowymi wagami, w zależności od liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych. Wykresy zostały przedstawione z uwzględnieniem modelu wiedzy zawierającego kolejno 3, 5 i 7 zbiorów rozmytych dla każdej zmiennej, definiowanych jak w tabeli III-2. Liczba rozłącznych przedziałów wpływa nieznacznie na strukturę modelu przy małej liczbie zbiorów rozmytych (rys. III-12). Przy siedmiu zbiorach rozmytych liczba reguł elementarnych modelu ulega większym wahaniom. Jednakże, zwiększenie liczby rozłącznych przedziałów powoduje wydłużenie czasu generowania reguł. Jest to szczególnie widoczne przy większej liczbie zbiorów rozmytych, niezależnie od ilości zmiennych modelu (rys. III-11).



Rys. III-11. Czas tworzenia reprezentacji wiedzy w zależności od liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych dla 3,5,7 zdefiniowanych zbiorów rozmytych każdej zmiennej w modelach:

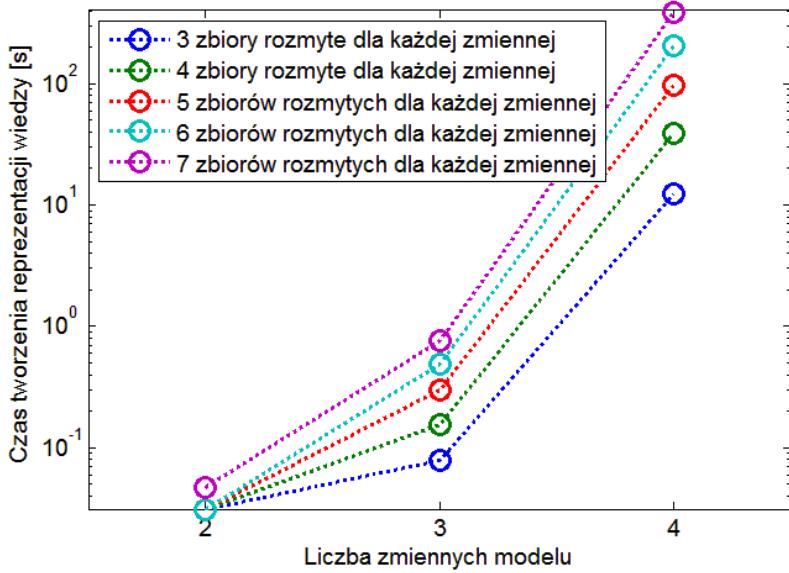
a) AR(2) b) AR(3)



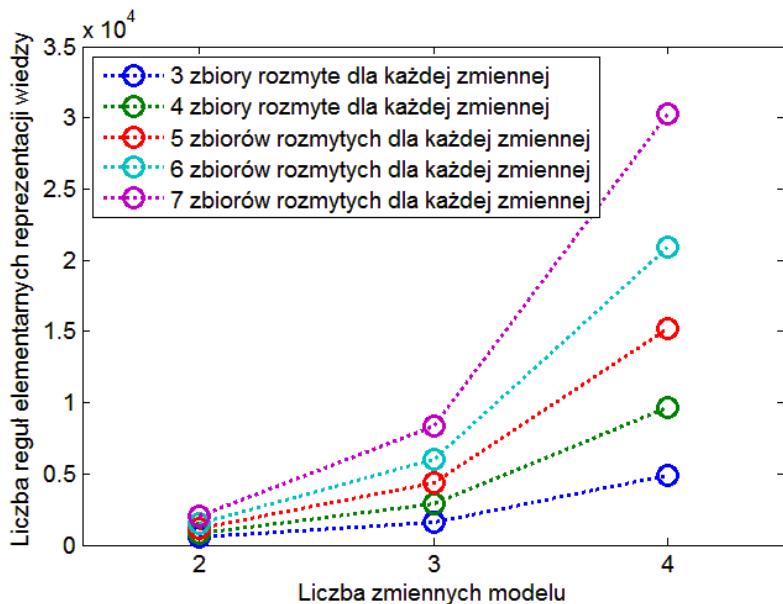
Rys. III-12. Liczba reguł elementarnych reprezentacji wiedzy w zależności od liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych dla 3,5,7 zdefiniowanych zbiorów rozmytych każdej zmiennej w modelach:

a) AR(2) b) AR(3)

Rysunki III-13 oraz III-14 zawierają wykresy pokazujące wyraźniej, iż czas generowania bazy reguł i liczba reguł elementarnych z niezerowymi wagami wzrasta wraz ze zwiększeniem się liczby zmiennych i liczby zbiorów rozmytych dla każdej zmiennej.



Rys. III-13. Zależność czasu tworzenia reprezentacji wiedzy od liczby zmiennych modelu dla różnej liczby zdefiniowanych zbiorów rozmytych każdej zmiennej



Rys. III-14. Zależność liczby reguł elementarnych reprezentacji wiedzy od liczby zmiennych modelu dla różnej liczby zdefiniowanych zbiorów rozmytych każdej zmiennej

Przedstawione wykresy (rys. III-10, III-14) potwierdzają dobrze znaną zasadę [pie03], że jeżeli oznaczymy ilość wejść modelu jako N i założymy, że każda wielkość wejściowa jest scharakteryzowana identyczną ilością J zbiorów rozmytych, to maksymalna ilość R reguł z prostymi przesłankami określona jest wzorem: $R=J^N$. W reprezentacji wiedzy (109) występują reguły z prostymi przesłankami i rozpatrujemy również wszystkie możliwe wartości lingwistyczne na wyjściu, zatem ilość reguł R zależy wykładniczo od ilości zmiennych $N+1$ oraz ilości zbiorów rozmytych J modelu.

Mimo obliczeń wektoryzowanych, przykład pokazuje, iż czas wyznaczenia pełnego rozkładu prawdopodobieństw równoczesnych zdarzeń rozmytych, zachodzących w poprzedniku i następniku reguły warunkowej, szybko wzrasta wraz z zwiększeniem się liczby zmiennych wejściowych, liczby ich wartości lingwistycznych oraz liczby rozłącznych przedziałów dyskretyzacji ich przestrzeni. Przy większych wartościach ww. parametrów utworzenie probabilistyczno-rozmytej bazy wiedzy może stać się nieosiągalne w rozsądny okresie czasu. Budowanie wielowymiarowych macierzy wpływa także na złożoność pamięciową programu, zatem niemożliwe staje się tworzenie systemu z dużą ilością wejść.

Inny przykład modelowania rozmytego dla szeregow czasowych z uwzględnieniem omówionego algorytmu został przedstawiony w [bla09].

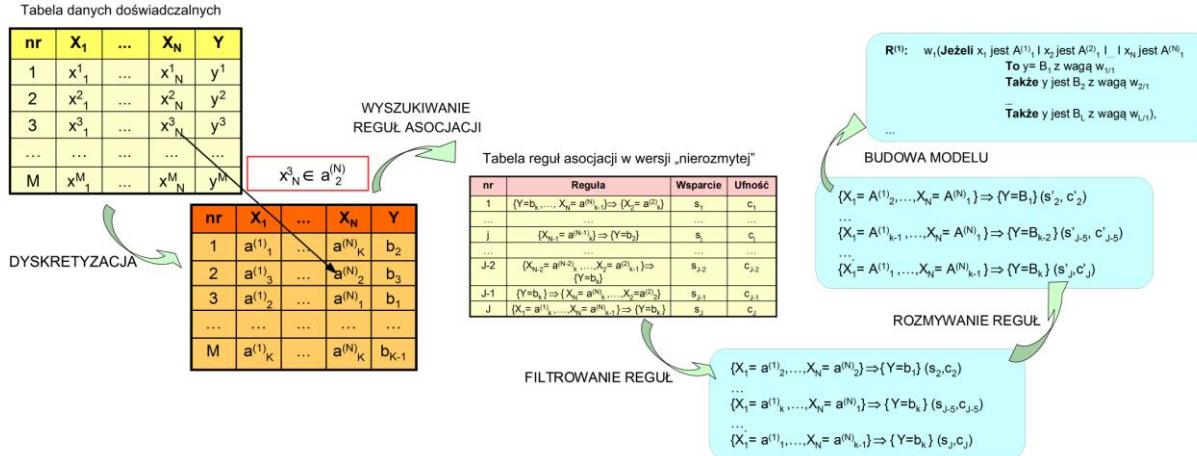
W dalszej części rozdziału III-3 dysertacji będą analizowane metody przeszukujące przestrzeń wartości zmiennych w celu generowania reguł probabilistyczno-rozmytego modelu wiedzy, które są oparte na niepełnym rozkładzie prawdopodobieństw zdarzeń rozmytych w regułach.

3.2. Algorytmy oparte na regułach asocjacji

Niezależnie od metody automatycznego pozyskiwania wiedzy, wymaga się, aby rozmyte reguły uzyskiwane były na podstawie ich optymalnego dopasowania do danych doświadczalnych. W tym sensie, generowanie reguł można rozumieć, jako wyszukiwanie reguł o dużej częstości występowania, przy czym, parametr częstości występowania charakteryzuje dopasowanie reguł. Wagi reguł w postaci IF-THEN wyrażające prawdopodobieństwo zdarzeń rozmytych w przesłance i konkluzji reguł mogą być traktowane jako miary współwystępowania rozmytych wartości zmiennych w kolekcjach danych doświadczalnych, co stanowi sens rozmytych reguł asocjacji (rozdział II-5.3). Zauważono również, iż struktura rozmytych reguł asocjacji jest zbliżona do struktury reguł probabilistyczno-rozmytego modelu wiedzy (109). Zaproponowano zatem, wykorzystanie metod odnajdywania reguł asocjacji do budowy probabilistyczno-rozmytej bazy wiedzy. Poniżej przedstawione zostaną algorytmy, które przeszukują przestrzeń wartości lingwistycznych zmiennych wejścia/wyjścia i wyszukują reguły bazy wiedzy na podstawie założeń metody odnajdywania rozmytych reguł asocjacji.

3.2.1. Budowa bazy reguł z uwzględnieniem algorytmu wyszukującego ilościowe reguły asocjacji

W pierwszym etapie poszukiwania odpowiedniej metody generowania reguł dla probabilistyczno-rozmytej bazy wiedzy, badano wykorzystanie algorytmu wyszukującego ilościowe reguły asocjacji (ang. *quantitative association rules*) [wb08b], [wb08c], [bla07a], [bla07b]. Metoda budowy modelu wiedzy polegała na wykorzystaniu algorytmu wyszukującego nieroźmyte reguły asocjacji oraz późniejszego rozmywania wygenerowanych reguł w celu otrzymania właściwego modelu stanowiącego probabilistyczno-rozmytą bazę wiedzy. Z uwagi na dostępność narzędzia Oracle Data Mining wraz z aplikacją Data Miner, do budowy modeli wykorzystano zaimplementowany w nim algorytm Apriori, który szczegółowo został omówiony w rozdziale II-5.2.1. Etapy modelowania przedstawia rysunek III-15.



Rys. III-15. Etapy modelowania z wykorzystaniem algorytmu wykrywającego ilościowe reguły asocjacji

Budowa modelu dla systemu typu MISO o N -wejściach i jednym wyjściu, składała się z następujących etapów (na podstawie [bła07b]):

- 1) Dyskretyzacja numerycznych wartości zmiennych wejściowych x_1, x_2, \dots, x_N i wyjściowej y do rozłącznych przedziałów odpowiednio oznaczonych jako $\{a_1^{(n)}, \dots, a_k^{(n)}, \dots, a_K^{(n)}\}$ dla x_n oraz $\{b_1, \dots, b_k, \dots, b_K\}$ dla y .
- 2) Wyszukiwanie reguł asocjacji z wykorzystaniem odpowiedniego algorytmu np. algorytmu Apriori, przy założonej wartości minimalnego wsparcia $\min s$ i minimalnej ufności $\min c$ dla reguł (patrz rozdział II-6.1–II-6.2).
- 3) Filtracja, wyszukanych w podpunkcie drugim, reguł asocjacji w celu otrzymania reguł typu:

$$\{x_1 = a_{k_1}^{(1)}, \dots, x_N = a_{k_N}^{(N)}\} \Rightarrow \{y = b_{k_{N+1}}\} (s_{k_1 \dots k_N k_{N+1}}, c_{k_1 \dots k_N k_{N+1}}), \quad (123)$$

które można zastosować do regułowego modelowania systemu.

- 4) Rozmywanie reguł asocjacji. Stosując podejście lingwistyczne [zad75] wprowadzamy zbiory rozmyte (wartości lingwistyczne) $A_{l_1}^{(n)}, B_{l_{N+1}}$ zmiennej wejściowej x_n i zmiennej wyjściowej y , zdefiniowanych wg definicji 3, które w rozmytych regułach asocjacji można traktować jako zdarzenia rozmyte. Korzystając z nierożmytego prawdopodobieństwa zdarzenia rozmytego wg Zadeha (na podstawie def. 24), w myśl metodyki opisanej w rozdziale III-3.2.3, dochodzimy do prawdopodobieństwa łącznego zdarzeń rozmytych $A_{l_1}^{(1)} \cap \dots \cap A_{l_N}^{(N)} \cap B_{l_{N+1}}$.

Niech zmiennej lingwistycznej x_n odpowiada zbiór wartości lingwistycznych $L(X_n)$. Wówczas prawdopodobieństwo zdarzenia rozmytego $A_{l_1}^{(1)} \cap \dots \cap A_{l_N}^{(N)} \cap B_{l_{N+1}}$ określonego w zbiorze $L(X_1) \times \dots \times L(X_n)$ zmiennej lingwistycznej będzie obliczane następująco:

$$P(A_{l_1}^{(1)} \cap \dots \cap A_{l_N}^{(N)} \cap B_{l_{N+1}}) = \sum_{k_1=1, \dots, K} \dots \sum_{k_{N+1}=1, \dots, K} p_{k_1 \dots k_{N+1}}(x_1, \dots, x_N, y) \cdot \\ \cdot T[\mu_{A_{l_1}^{(1)}}(a_{k_1}^{(1)}), \dots, \mu_{A_{l_N}^{(N)}}(a_{k_N}^{(N)}), \mu_{B_{l_{N+1}}} (b_{k_{N+1}})], \quad (124)$$

gdzie: T jest operatorem t-normy, zdefiniowanym wg def. 12.

Wzór (124) określa wparcie ($s'_{l_1 \dots l_{N+1}}$) dla rozmytych reguł asocjacji. Miara ufności ($c'_{l_1 \dots l_{N+1}}$) tych reguł wyznaczana jest natomiast na podstawie rozkładów warunkowych zdarzeń rozmytych:

$$P(B_{l_{N+1}} / A_{l_1}^{(1)} \cap \dots \cap A_{l_N}^{(N)}) = P(A_{l_1}^{(1)} \cap \dots \cap A_{l_N}^{(N)} \cap B_{l_{N+1}}) / P(A_{l_1}^{(1)} \cap \dots \cap A_{l_N}^{(N)}), \quad (125)$$

przy czym suma tych prawdopodobieństw dla wszystkich możliwych zdarzeń rozmytych konkluzji sumuje się do jedynki. W ten sposób otrzymujemy lingwistyczne reguły asocjacji w postaci:

$$\{x_1 = A_{l_1}^{(1)}, \dots, x_N = A_{l_N}^{(N)}\} \Rightarrow \{y = B_{l_{N+1}}\} \quad (s'_{l_1 \dots l_N l_{N+1}}, c'_{l_1 \dots l_N l_{N+1}}), \quad (126)$$

gdzie: $s'_{l_1 \dots l_{N+1}}$ i $c'_{l_1 \dots l_{N+1}}$ są wsparciem i ufnością wg (124-125) odpowiednio dla zbiorów rozmytych o indeksach l_1, \dots, l_N zmiennych wejściowych x_1, x_2, \dots, x_N oraz zbioru rozmytego o indeksie l_{N+1} dla zmiennej wyjściowej y .

- 5) Budowa probabilistyczno-rozmytej reprezentacji wiedzy. Reguły (126) mogą mieć zastosowanie do tworzenia modelu wiedzy dla systemu przy istnieniu niepewności o charakterze rozmytym. Wówczas pierwsza reguła plikowa R(1) modelu ma postać:

Jeżeli x_1 jest $A_{l_1}^{(1)}$ I ... I x_N jest $A_{l_N}^{(N)}$, z wagą w_1
To y jest $B_{l_{N+1}}$ z wagą $w_{1/1}$
 ...
Także y jest $B_{l_{N+1}}$ z wagą $w_{1/1}$
 ...
Także y jest $B_{l_{N+1}}$ z wagą $w_{L/1}$. (127)

Wagi reguł w_1 oraz $w_{l/1}$ są prawdopodobieństwami, wynikającymi odpowiednio z wsparcia i ufności w rozmytych regułach asocjacji, następująco:

$$w_1 = \sum_{i=1 \dots L} s'_{l_1 \dots l_N i/1}, \quad (128)$$

$$w_{l/1} = c'_{l_1 \dots l_N l/1} = \frac{s'_{l_1 \dots l_N l/1}}{\sum_{i=1 \dots L} s'_{l_1 \dots l_N i/1}}. \quad (129)$$

W przypadku założenia wartości minimalnego wsparcia $\min s$ i minimalnej ufności $\min c$ na poziomie zera dla wyszukiwania ilościowych reguł asocjacji, otrzymujemy model wiedzy identyczny jak w przypadku algorytmu generującego pełny rozkład prawdopodobieństwa (rozdz. III-3.1). Jeżeli wartość $\min s$ jest większa od zera (bądź/i wartość $\min c$ jest większa od zera), wówczas w podpunkcie 2) i 3) otrzymujemy reguły, dla których prawdopodobieństwo występowania jednocześnie zdarzeń w przesłance i konkluzji jest większe od wartości $\min s$, natomiast prawdopodobieństwo warunkowe występowania zdarzeń w konkluzji jest większe od wartości $\min c$. Toteż wagi reguły (128), (129) probabilistyczno-rozmytej reprezentacji wiedzy stanowią zanione prawdopodobieństwa występowania zdarzeń rozmytych w regule. Przy małych wartościach $\min s$ i $\min c$ wagi reguł są zbliżone do właściwego prawdopodobieństwa rozkładu empirycznego zdarzeń rozmytych.

Dla skrócenia czasu obliczeń rozmywania reguł asocjacji (podpunkt 4) można również zastosować operacje tablicowe wykonywane na macierzach:

$$s'_{l_1 \dots l_N l_{N+1}} = \sum_{a_{k1}^{(1)} \dots a_{kN}^{(N)}, b_{kN+1} \in X_1 \times \dots \times X_N \times Y} E^{p'} \cdot * T^*(E_{A_{l_1}^{(1)}}^{(1)}, \dots, E_{A_{l_N}^{(N)}}^{(N)}, E_{B_{l_{N+1}}}^{(N+1)}), \quad (130)$$

gdzie: $E^{p'}$ stanowi macierz $[s'_{l_1 \dots l_N l_{N+1}}]_{K \times K \times \dots \times K}$, $E_{A_{ln}^{(n)}}^{(n)}$ – macierze o elementach (117), $E_{B_{l_3}}^{(3)}$ – macierz o elementach (118), „ $\cdot *$ ” oznacza mnożenie tablicowe, natomiast T^* – operator t-normy wykonywany z wykorzystaniem operacji tablicowych.

W [bla07a], [bla07b] opisano szczegółowo możliwości modelowania dla szeregów czasowych.

PRZYKŁAD

Zastosowanie ilościowych reguł asocjacji do modelowania w systemie z probabilistyczno-rozmytą bazą wiedzy zostanie pokazane na przykładzie modelowania systemu dynamicznego dla wybranych cech węgla kamiennego [bla07b]. Użyto danych: udział jednej frakcji gęstościowej i zawartość popiołu w tej frakcji, oznaczone w kolejnych 496 próbkach.

Probabilistyczno-rozmyty model wiedzy udziału frakcji lekkiej węgla (x) został oparty o model autoregresji AR(1), zawierający zmienne: $x(t)$, $x(t-1)$. Zawartość frakcji lekkiej węgla w próbce jest daną ilościową, wobec tego dokonano dyskretyzacji wartości na 10 przedziałów a_1, \dots, a_{10} , stałych w przestrzeniach zmienności. Wówczas, nierozmyte reguły asocjacji dla modelu AR(1) ($\min s=0$, $\min c=0$), będące wynikiem zastosowania Oracle Data Miner, mają postać jak w tabeli III-3.

Tab. III-3. Przykład ilościowych reguł asocjacji zmiennych $x(t)$, $x(t-1)$

Nr reguły	JEŻELI	TO	c [%]	s [%]
66	$x(t-1)=(0.34,0.41>$	$x(t)=(0.34,0.41>$	30,84	6,67
84	$x(t-1)=(0.34,0.41>$	$x(t)=(0.41,0.47>$	27,10	5,86
100	$x(t-1)=(0.34,0.41>$	$x(t)=(0.47,0.54>$	19,63	4,24
48	$x(t-1)=(0.34,0.41>$	$x(t)=(0.27,0.34>$	6,54	1,41
118	$x(t-1)=(0.34,0.41>$	$x(t)=(0.54,0.61>$	5,61	1,21
32	$x(t-1)=(0.34,0.41>$	$x(t)=(0.20,0.27>$	5,61	1,21
20	$x(t-1)=(0.34,0.41>$	$x(t)=(0.14,0.20>$	1,87	0,40
8	$x(t-1)=(0.34,0.41>$	$x(t)=(0.07,0.14>$	1,87	0,40
130	$x(t-1)=(0.34,0.41>$	$x(t)=(0.61,...)$	0,93	0,20

Dokonując fuzzyfikacji wyznaczonych reguł asocjacji, wprowadzono pięć zbiorów rozmytych A_1, \dots, A_5 , utożsamianych z wartościami zmiennej lingwistycznej Z o wartościach z zakresu:

$$L(Z)=\{\text{'bardzo niska', 'niska', 'średnia', 'wysoka', 'bardzo wysoka'}\}, \quad (131)$$

w skrócie: $L(Z)=\{\text{'BN}', \text{'N}', \text{'S}', \text{'W}', \text{'BW'}\}$. Przyjęto stałe wartości stopni przynależności $\mu_{A_j}(a_i)$, z jakimi zbiory a_1, \dots, a_{10} , uzyskane z dyskretyzacji wartości atrybutów, należą do zbiorów A_1, \dots, A_5 (tab. III-4).

Tab. III-4. Wartości stopni przynależności $\mu_{A_j}(a_i)$ ($j=1, \dots, 5$; $i=1, \dots, 10$) zbiorów rozmytych dla udziału frakcji lekkiej węgla

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
$\mu_{A1}(a_i)$	0,9	0,5	0,2	0,1	0	0	0	0	0	0
$\mu_{A2}(a_i)$	0,1	0,5	0,7	0,5	0,1	0	0	0	0	0
$\mu_{A3}(a_i)$	0	0	0,1	0,4	0,9	0,9	0,4	0,1	0	0
$\mu_{A4}(a_i)$	0	0	0	0	0	0,1	0,5	0,7	0,5	0,1
$\mu_{A5}(a_i)$	0	0	0	0	0	0	0,1	0,2	0,5	0,9

Oczywiście zachodzi zależność (107), czyli:

$$\sum_{j=1}^5 \mu_{A_j}(a_i) = 1; \quad i = 1, \dots, 10. \quad (132)$$

Wyniki obliczeń dla AR(1) na podstawie wzorów (124)-(125), w których otrzymujemy rozkłady prawdopodobieństwa zdarzeń rozmytych, związanych z przynależnością wartości do zbiorów A_1, \dots, A_5 , zostały przedstawione w tabeli III-5 (rozkład łączny) i w tabeli III-6 (rozkład warunkowy).

Tab. III-5. Łączny rozkład prawdopodobieństwa dla zmiennych lingwistycznych $x(t)$, $x(t-1)$

$x(t)$	$x(t-1)$	Rozkład łączny prawdopodobieństwa $P(x(t-1)=A_i \cap x(t)=A_j), i,j=1, \dots, 5$				
		$A_1 (\text{BN})$	$A_2 (\text{N})$	$A_3 (\text{S})$	$A_4 (\text{W})$	$A_5 (\text{BW})$
$A_1 (\text{BN})$		0,0003	0,0018	0,0125	0,0080	0,0024
$A_2 (\text{N})$		0,0019	0,0079	0,0342	0,0216	0,0066
$A_3 (\text{S})$		0,0113	0,0321	0,2031	0,1367	0,0469
$A_4 (\text{W})$		0,0069	0,0203	0,1363	0,1309	0,0511
$A_5 (\text{BW})$		0,0037	0,0091	0,0442	0,0497	0,0206

Tab. III-6. Warunkowy rozkład prawdopodobieństwa dla zmiennych lingwistycznych $x(t)$, $x(t-1)$

$x(t)$ \diagdown $x(t-1)$	Rozkład warunkowy prawdopodobieństwa $P(x(t)=A_j/x(t-1)=A_i), i,j=1,\dots,5$				
	$A_1(BN)$	$A_2(N)$	$A_3(S)$	$A_4(W)$	$A_5(BW)$
$A_1(BN)$	0,0134	0,0253	0,0290	0,0231	0,0190
$A_2(N)$	0,0773	0,1108	0,0794	0,0623	0,0517
$A_3(S)$	0,4681	0,4511	0,4721	0,3941	0,3676
$A_4(W)$	0,2882	0,2852	0,3167	0,3773	0,4002
$A_5(BW)$	0,1529	0,1276	0,1028	0,1432	0,1616

Tab. III-7. Przykład rozmytych reguł asocjacji dla zmiennych lingwistycznych $x(t)$, $x(t-1)$

Nr reguły	JEŻELI	TO	c'	s'
1	$x(t-1)=S$	$x(t)=S$	0,4721	0,2031
2	$x(t-1)=W$	$x(t)=S$	0,3941	0,1367
3	$x(t-1)=S$	$x(t)=W$	0,3167	0,1363
4	$x(t-1)=W$	$x(t)=W$	0,3773	0,1309
5	$x(t-1)=BW$	$x(t)=W$	0,4002	0,0511
6	$x(t-1)=W$	$x(t)=S$	0,1432	0,0497

Przykładowe, rozmyte reguły asocjacji prezentuje tabela III-7. Wówczas, model systemu dynamicznego, z uwzględnieniem autoregresji rzędu pierwszego, zawiera 5 reguł plikowych. W każdym pliku mogą znajdować się do 5 reguł elementarnych. Otrzymano zbiór reguł, z których pierwsza jest w postaci:

R(1): Jeżeli $x(t-1)$ jest S, z wagą 0,4303

To $x(t)$ jest S z wagą 0,4721

Także $x(t)$ jest W z wagą 0,3167 (133)

Także $x(t)$ jest BW z wagą 0,1028

Także $x(t)$ jest N z wagą 0,0794

Także $x(t)$ jest BN z wagą 0,0290.

Dla większej ilości zmiennych modelu zostanie przedstawiony sposób tworzenia reguł z wykorzystaniem udziału frakcji lekkiej węgla i popiołu w tej frakcji. Podział przestrzeni zmienności dla udziału frakcji lekkiej węgla (x) i procentowego udziału popiołu (y) frakcji lekkiej dokonano, podobnie jak w poprzednim podpunkcie, na 10 równych zbiorów nieroźmytych, nazwanych odpowiednio a_1, \dots, a_{10} oraz b_1, \dots, b_{10} . Dla udziału frakcji lekkiej węgla wprowadzono pięć zbiorów rozmytych A_1, \dots, A_5 , odnoszących się do wartości lingwistycznych zmiennej (131) oraz posiadających wartości funkcji przynależności $\mu_{A_j}(a_i)$ jak w tabeli III-4. Dla popiołu wprowadzono także pięć zbiorów rozmytych B_1, \dots, B_5 z analogicznymi wartościami lingwistycznymi $L(K)$ zmiennej lingwistycznej K oraz wartościami stopni przynależności $\mu_{B_j}(b_i)$ zawartymi w tabeli III-8.

Tab. III-8. Wartości stopni przynależności $\mu_{Bj}(b_i)$ zbiorów rozmytych dla zawartości popiołu we frakcji lekkiej

	b₁	b₂	b₃	b₄	b₅	b₆	b₇	b₈	b₉	b₁₀
$\mu_{B1}(b_i)$	0,7	0,1	0	0	0	0	0	0	0	0
$\mu_{B2}(b_i)$	0,3	0,5	0,1	0	0	0	0	0	0	0
$\mu_{B3}(b_i)$	0	0,4	0,9	0,9	0,7	0,4	0	0	0	0
$\mu_{B4}(b_i)$	0	0	0	0,1	0,3	0,6	0,9	0,6	0,1	0
$\mu_{B5}(b_i)$	0	0	0	0	0	0	0,1	0,4	0,9	1

Przeprowadzono eksplorację danych empirycznych algorytmem Apriori dla dwóch atrybutów x i y zgodnie z modelem $y(t)=f(x(t-2),x(t-1),y(t-2),y(t-1))$. Wartość wyjściową modelu stanowi zawartość popiołu w frakcji lekkiej, mierzoną w czasie t . Reguły z najwyższymi miarami wsparcia dla poszukiwanej probabilistyczno-rozmytej reprezentacji wiedzy zawarte są w tabeli III-9.

Tab. III-9. Przykład ilościowych reguł asocjacji modelu $y(t)=f(x(t-2),x(t-1),y(t-2),y(t-1))$

Nr reguły	JEŻELI	TO	c [%]	s [%]
11842	$x(t-2) = <0.34, 0.41)$ AND $y(t-2) = <1.5, 2.38)$ AND $x(t-1) = <0.34, 0.41)$ AND $y(t-1) = <1.5, 2.38)$	$y(t) = <1.5, 2.38)$	81,82	1,82
11877	$x(t-2) = <0.34, 0.41)$ AND $y(t-2) = <1.5, 2.38)$ AND $x(t-1) = <0.41, 0.47)$ AND $y(t-1) = <1.5, 2.38)$	$y(t) = <1.5, 2.38)$	77,78	1,42
12267	$x(t-2) = <0.41, 0.47)$ AND $y(t-2) = <1.5, 2.38)$ AND $x(t-1) = <0.41, 0.47)$ AND $y(t-1) = (... , 1.5)$	$y(t) = <1.5, 2.38)$	60,00	1,21

Po dokonaniu obliczeń według (130), wprowadzających rozmycie do reguł asocjacji, otrzymano rozkłady prawdopodobieństwa łącznego w postaci tablicy o wymiarach $5 \times 5 \times 5 \times 5 \times 5$. Przykładowy rozkład łącznego prawdopodobieństwa, gdy $x(t-2)=A_3$, $y(t-2)=B_3$ oraz $y(t-1)=B_3$, przedstawia tabela III-10.

Tab. III-10. Łączny rozkład prawdopodobieństwa dla poszukiwanego probabilistyczno-rozmytego modelu wiedzy

y(t)	x(t-1)	Rozkład łączny prawdopodobieństwa				
		P[x(t-2)= A ₃ ∩ y(t-2)= B ₃ ∩ x(t-1)=A _i ∩ y(t-1)= B _j] , i,j=1,...,5	A ₁ (BN)	A ₂ (N)	A ₃ (S)	A ₄ (W)
B ₁ (BN)		0,0001	0,0007	0,0029	0,0017	0,0007
B ₂ (N)		0,0008	0,0032	0,0146	0,0101	0,0043
B ₃ (S)		0,0025	0,0101	0,0477	0,0379	0,0175
B ₄ (W)		0,0003	0,0015	0,0050	0,0039	0,0017
B ₅ (BW)		0,0000	0,0003	0,0006	0,0007	0,0006

Kolejno, wagi reguł elementarnych zostały podane jako warunkowe prawdopodobieństwo zdarzeń rozmytych $P[y(t)= B_{3j} / (x(t-2)= A_{1i} \cap y(t-2)= B_{1j} \cap x(t-1)=A_{2i} \cap y(t-1)= B_{2j})]$,

analogicznie do (128)-(129). Otrzymano reguły plikowe z wagami niezerowymi, z których pierwsza reguła jest w postaci:

R(1): Jeżeli $x(t-2)$ jest S **I** $y(t-2)$ jest S **I** $x(t-1)$ jest S **I** $y(t-1)$ jest S, z wagą 0,07082

To $y(t)$ jest S z wagą 0,67303

Także $y(t)$ jest N z wagą 0,20683

Także $y(t)$ jest W z wagą 0,06990

Także $y(t)$ jest BN z wagą 0,04110

Także $y(t)$ jest BW z wagą 0,00914.

Na podstawie przedstawionego przykładu można stwierdzić, iż:

- generowanie bazy wiedzy z wykorzystaniem algorytmu tworzącego ilościowe reguły asocjacji stanowi stosunkowo niewielką część w przedstawionym procesie modelowania (zwłaszcza 20-25% czasu),
- najwięcej czasu pochłania etap przygotowania danych oraz etap "rozmywania" wcześniej otrzymanych reguł,
- czas obliczeń znacznie wzrasta wraz ze zwiększeniem liczb zmiennych systemu (wejść i wyjść), liczby rozłącznych przedziałów (na jakie przestrzeń wartości rozważanych zmiennych została podzielona) oraz liczby wartości lingwistycznych tychże zmiennych,
- wstępna granulacja przestrzeni zmiennych, w celu znalezienia wiarygodnych reguł z rozkładem prawdopodobieństwa zdarzeń nieroźmytych oraz późniejsze rozmywanie zdarzeń, zmniejsza złożoność modelu wiedzy pod kątem liczby reguł, a nieznacznie tylko skracia czas obliczeń,
- ponadto, algorytmy wyszukujące reguły asocjacji (w tym algorytm Apriori) generują dużo więcej reguł, niż jest to potrzebne do budowy modelu wiedzy.

Jednym z rozwiązań jest zastosowanie metod obliczeń bazujących na algorytmach wyszukujących bezpośrednio rozmyte reguły asocjacji. Ponieważ omawiane algorytmy nie są zaimplementowane w znanych autorowi, komercyjnych narzędziach eksploracji, postanowiono utworzyć własne narzędzie, które pozwoli nie tylko na modelowanie z użyciem idei wyszukiwania rozmytych reguł asocjacji ale również późniejsze wnioskowanie w oparciu o zbudowany model wiedzy.

3.2.2. Budowa bazy reguł z uwzględnieniem zmodyfikowanego algorytmu Apriori

W ostatnich latach temat wyszukiwania rozmytych reguł asocjacji jest często poruszany w anglojęzycznych publikacjach. Spotyka się jednak różne podejście w stosunku do omawianych metod (patrz rozdział II-5). Elementem najczęściej analizowanym jest czas wyszukiwania reguł asocjacji oraz jakość utworzonych reguł. Istnieje zatem wiele algorytmów przeszukujących przestrzeń wartości zmiennych w celu znalezienia rozmytych reguł asocjacji, a same reguły są różnie definiowane (patrz rozdział II-5.3). Najczęściej spotykane definicje wielkości miar statystycznych wsparcia i ufności rozmytych reguł asocjacji są określone za pomocą wzorów (102)-(104). W dysertacji proponuje się również obliczanie miary wsparcia na podstawie zależności (116), która pozwala na obliczenie prawdopodobieństwa zajścia jednoczesnego zdarzeń rozmytych zdefiniowanych na

rozłącznych przedziałach wartości zmiennych. Założenie wartości progowej minimalnego wsparcia pozwoli na ograniczenie liczby reguł rozmytych do tych, których wsparcie jest wyższe od wartości zakładanej $\min w$. Pojedyncze zdarzenia rozmyte ($k=1$) lub iloczyny k zdarzeń rozmytych, które spełniają warunek minimalnego wsparcia są wówczas nazwane *rozmytymi zdarzeniami częstymi k-elementowymi*. Zbiory rozmytych zdarzeń częstych ($k=1$) lub iloczynów k rozmytych zdarzeń będą nazwane *zbiorami częstymi k-elementowymi*. Analizowany algorytm generowania probabilistyczno-rozmytej bazy wiedzy jest oparty na głównym założeniu algorytmu Apriori (szczegółowe omówienie algorytmu znajduje się w rozdziale II-5.2.1), które mówi o tym, iż podzbiory zdarzeń częstych stanowią także zdarzenia częste. Stąd też budowa reguł rozpocznie się od wybrania zbioru jednoelementowych rozmytych zdarzeń A_j^n , B_j , których prawdopodobieństwo występowania jest większe od wartości $\min w$. Można zatem zauważać, iż konkluzje reguł są wydobywane jednocześnie z przesłankami. Omawiane podejście zostało przedstawione w [wb09].

Zmodyfikowany algorytm Apriori

Wejścia proponowanego algorytmu:

- zbiór I pomiarów użytych do identyfikacji modelu,
- predefiniowana baza danych (wartości lingwistyczne zmiennych rozważanych w modelu oraz definicje zbiorów rozmytych utożsamianych z tymi wartościami),
- wartość progowa minimalnego wsparcia ($\min w$).

Wyjście: baza reguł probabilistyczno-rozmytej bazy wiedzy.

Notacja użyta do przedstawienia algorytmu jest następująca:

I – liczba pomiarów użytych do identyfikacji modelu wiedzy,

$N+1$ – całkowita liczba zmiennych (N zmiennych wejściowych, 1 zmienna wyjściowa),

K – liczba rozłącznych przedziałów o równej szerokości w przestrzeniach zmiennych,

x_n – zmienne wejściowe modelu, $x_n \in X_n \subset \mathbb{R}$, $n=1, \dots, N$,

y – zmienna wyjściowa modelu, $y \in Y \subset \mathbb{R}$,

$|A_j^{(n)}|$ ($|B_j|$) – liczba wartości lingwistycznej dla n -tej zmiennej wejściowej (zmiennej wyjściowej)

$A_j^{(n)}$ – j -ta wartość lingwistyczna n -tej zmiennej wejściowej, $j=1, \dots, |A_j^{(n)}|$, $n=1, \dots, N$,

B_j – j -ta wartość lingwistyczna zmiennej wyjściowej, $j=1, \dots, |B_j|$,

$\mathbf{a}^{(n)} = (a_1^{(n)}, \dots, a_k^{(n)}, \dots, a_K^{(n)})$ – rozłączne przedziały wartości n -tej zmiennej wejściowej x_n , $n=1, \dots, N$,

$\mathbf{b} = (b_1, \dots, b_k, \dots, b_K)$ – rozłączne przedziały wartości zmiennej wyjściowej y ,

w – obliczona wartość wsparcia dla kandydatów zbiorów,

$\min w$ – założona, minimalna wartość wsparcia,

C_r – r -elementowy zbiór kandydatów, składający się z rozmytych zdarzeń dla r ($1 \leq r \leq N+1$) zmiennych systemu,

F_r – zbiór częsty r -elementowy, składający się z rozmytych zdarzeń częstych r -elementowych (zdarzeń rozmytych dla r ($1 \leq r \leq N+1$) zmiennych systemu),

D – dane empiryczne dotyczące badanego systemu, w terminologii data mining często określane jako dane transakcyjne (ang. *transaction data*),

D^i – i -ty zbiór wartości empirycznych zmiennych modelu $\{x_1^i, \dots, x_N^i, y^i\}$, $i=1, \dots, I$ (i -ty pomiar).

Algorytm do generowania reguł dla systemu wnioskującego typu MISO z probabilistyczno-rozmytą bazą wiedzy został przedstawiony poniżej.

Krok 1: Dyskretyzacja wartości liczbowych każdego zbioru D^i ($i=1, \dots, I$) na K rozłączne przedziały ($\mathbf{a}^{(n)}$; $n=1, \dots, N$ oraz \mathbf{b}) o jednakowej szerokości w przestrzeni zmiennych.

Krok 2: Na podstawie próby D , obliczenie rozkładu prawdopodobieństwa dla zdarzenia nierożmytego:

$$P_{k_1 \dots k_N k_{N+1}}(x_1, \dots, x_N, y) = P(x_1 \in a_{k_1}^{(1)}, \dots, x_N \in a_{k_N}^{(N)}, y \in b_{k_{N+1}}), k_1, \dots, k_{N+1}=1, \dots, K. \quad (135)$$

Krok 3: Zdefiniowanie kandydatów zbioru częstego 1-elementowego C_1 . Obliczenie wartości wsparcia dla zmiennych wejściowych:

$$w(A_j^{(n)}) = P(A_j^{(n)}) = \sum_{k_n=1}^K P(x_n \in a_{k_n}^{(n)}) \cdot \mu_{A_j^{(n)}}(a_{k_n}^{(n)}), j=1, \dots, |A|, n=1, \dots, N, \quad (136)$$

oraz zmiennej wyjściowej:

$$w(B_j) = P(B_j) = \sum_{k_{N+1}=1}^K P(y \in b_{k_{N+1}}) \cdot \mu_{B_j}(b_{k_{N+1}}), j=1, \dots, |B|. \quad (137)$$

Krok 4: Wybranie zdarzeń rozmytych ze zbioru C_1 , które spełniają założenie minimalnego wsparcia. Wybrane zdarzenia rozmyte będą stanowić zbiór F_1 składający się z rozmytych zdarzeń częstych 1-elementowych:

$$F_1 = \{A_j^{(n)} \mid \forall j, n : w(A_j^{(n)}) \geq \min w, B_j \mid \forall j : w(B_j) \geq \min w\}. \quad (138)$$

Krok 5: Założenie $r=1$ jako liczby elementów w rozmytym zdarzeniu częstym. Jeżeli zbiór F_r nie jest pusty przejście do kolejnego punktu algorytmu. W przeciwnym wypadku przerwanie wykonywania algorytmu. Brak reguł rozmytych spełniających warunek minimalnego wsparcia ($\min w$).

Krok 6: Utworzenie C_{r+1} zbioru kandydatów zbioru częstego $r+1$ -elementowego na podstawie zbioru częstego r -elementowego F_r :

$$F_r = \{\{A_{j_1}^{(n_1)}, \dots, A_{j_{r-1}}^{(n_{r-1})}, A_{j_r}^{(n_r)}\}, \{A'_{j_1}^{(n_1)}, \dots, A'_{j_{r-1}}^{(n_{r-1})}, A'_{j_r}^{(n_r)}\}, \dots\}, \quad (139)$$

przy czym któryś z zdarzeń rozmytych $A_{j_{r-s}}^{(n_{r-s})}$, $s = 0, \dots, r-1$ ($A'_{j_{r-s}}^{(n_{r-s})}$) może stanowić zdarzenie rozmyte $B_{j_{r-s}}$ ($B'_{j_{r-s}}$).

Zbiór C_{r+1} składa się z połączenia elementów zbioru F_r spełniających warunek:

$$(A_{j_1}^{(n_1)} = A'_{j_1}^{(n_1)}) \wedge (A_{j_2}^{(n_2)} = A'_{j_2}^{(n_2)}) \wedge \dots \wedge (A_{j_{r-1}}^{(n_{r-1})} = A'_{j_{r-1}}^{(n_{r-1})}) \wedge (A_{j_r}^{(n_r)} \neq A'_{j_r}^{(n_r)}). \quad (140)$$

Wówczas, zbiór C_{r+1} jest równy:

$$C_{r+1} = \{A_{j_1}^{(n_1)}, A_{j_2}^{(n_2)}, \dots, A_{j_{r-1}}^{(n_{r-1})}, A_{j_r}^{(n_r)}, A_{j_r}^{(n_r)}\}. \quad (141)$$

Każdy element zbioru stanowi zdarzenie rozmyte dotyczące innej zmiennej wejściowej lub wyjściowej. Należy także usunąć zbiory z powtarzającymi się elementami.

Krok 7: Obliczenie miary wsparcia dla $r+1$ -elementowych kandydatów rozmytego zbioru częstego następująco:

$$\begin{aligned} w(A_{j_1}^{(n_1)}, \dots, A_{j_{r+1}}^{(n_{r+1})}) &= P(A_{j_1}^{(n_1)} \cap \dots \cap A_{j_{r+1}}^{(n_{r+1})}) = \\ &\sum_{k_1=1}^K \dots \sum_{k_{r+1}=1}^K [p_{k_1 \dots k_{r+1}}(x_{n_1}, \dots, x_{n_{r+1}}) \cdot T(\mu_{A_{j_1}^{(n_1)}}(a_{k_1}^{(n_1)}), \dots, \mu_{A_{j_{r+1}}^{(n_{r+1})}}(a_{k_{r+1}}^{(n_{r+1})}))], \end{aligned} \quad (142)$$

gdzie: symbol T oznacza operator t-normy wg tabeli III-1.

Należy zauważyć, że któryś ze zmiennych $x_{n_{r-s}}, s = -1, 0, \dots, r-1$ może stanowić zmienną y , wówczas $A_{j_{r-s}}^{(n_{r-s})} = B_{j_{r-s}}$ oraz $a_{k_{r-s}}^{(n_{r-s})} = b_{k_{r-s}}$.

Krok 8: Wybranie kandydatów rozmytych zdarzeń częstych $r+1$ -elementowych ze zbioru C_{r+1} , które spełniają założenie minimalnego wsparcia i utworzenie z nich zbioru F_{r+1} składającego się z rozmytych zdarzeń częstych $r+1$ -elementowych:

$$F_{r+1} = \{\{A_{j_1}^{(n_1)}, \dots, A_{j_{r+1}}^{(n_{r+1})}\} \mid w(A_{j_1}^{(n_1)}, \dots, A_{j_{r+1}}^{(n_{r+1})}) \geq \min w\}. \quad (143)$$

Krok 9: Jeżeli zbiór F_{r+1} nie jest pusty i $r < N$ to powtórzenie kroków 6-9, przy założeniu $r=r+1$, jako liczby zdarzeń rozmytych w zbiorze częstym. Jeżeli zbiór F_{r+1} jest pusty i $r \leq N$ to zakończ wykonywanie algorytmu – brak reguł o podanym wsparciu $\min w$. W przeciwnym wypadku przejście do kroku 10.

Krok 10: Utworzenie bazy reguł modelu probabilistyczno-rozmytego na podstawie rozmytych zdarzeń częstych zbioru F_{r+1} . Otrzymujemy zbiór reguł elementarnych w postaci:

$$(w(A_{j_1}^{(n_1)}, \dots, A_{j_r}^{(n_r)}, B_{l/j})) \text{ JEŻELI } x_1 \text{ jest } A_{j_1}^{(n_1)} \text{ I } \dots \text{ I } x_N \text{ jest } A_{j_r}^{(n_r)} \text{ TO } y \text{ jest } B_{l/j} \quad (144)$$

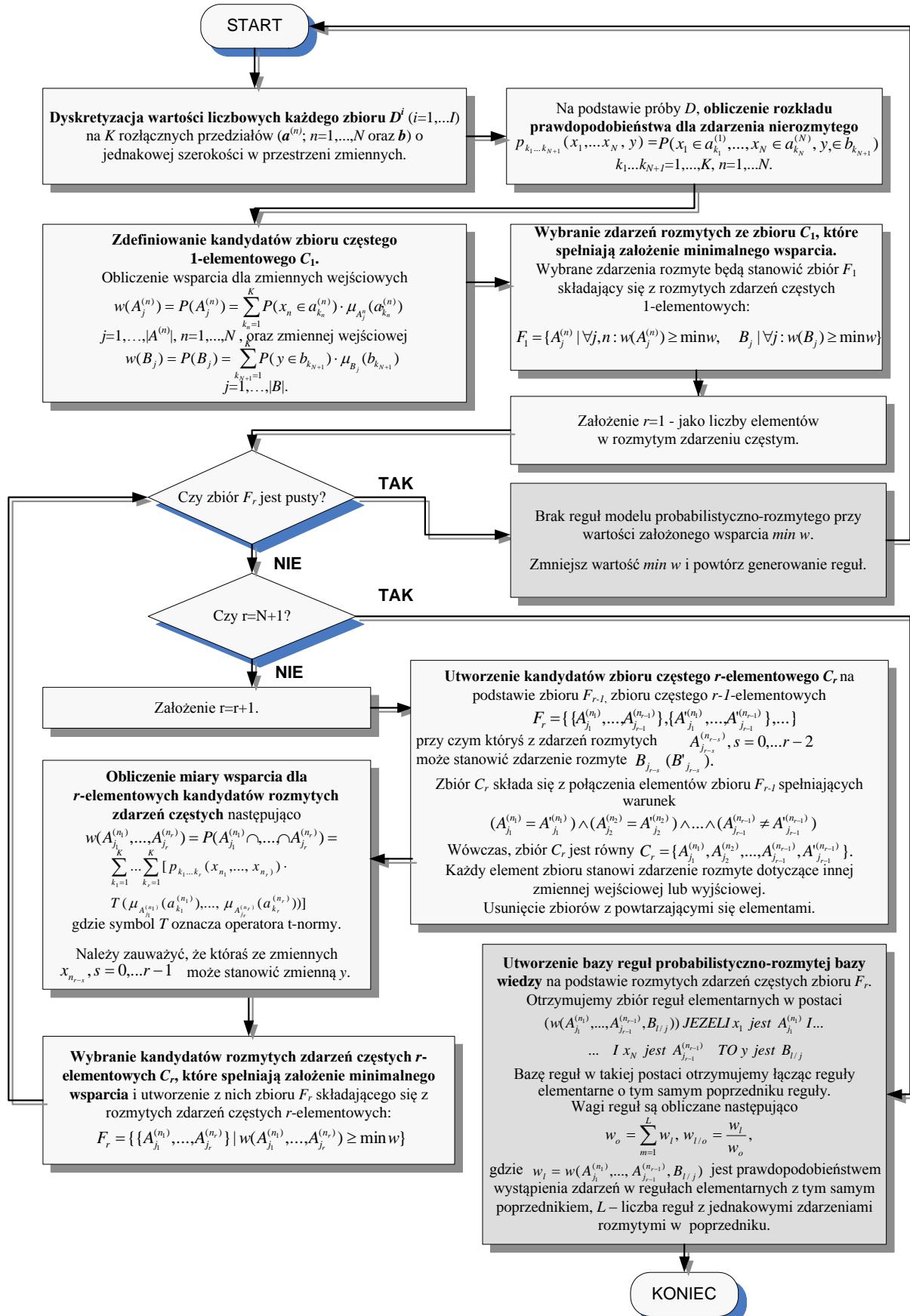
Bazę reguł w postaci (109) otrzymujemy łącząc reguły elementarne (144) o tym samym poprzedniku reguły. Wagi reguł są obliczane następująco:

$$w_o = \sum_{l=1}^L w_l, \quad w_{l/o} = \frac{w_l}{w_o} \quad (145)$$

gdzie $w_l = w(A_{j_1}^{(n_1)}, \dots, A_{j_r}^{(n_r)}, B_{l/j})$ jest prawdopodobieństwem wystąpienia zdarzeń rozmytych w regułach elementarnych z tym samym poprzednikiem (x_1 jest $A_{j_1}^{(n_1)}$ I ... I x_N jest $A_{j_r}^{(n_r)}$), L – liczbą reguł z jednakowymi zdarzeniami rozmytymi w poprzedniku.

W przypadku, gdy nie ma rozmytych zdarzeń częstych $N+1$ elementowych, wówczas należy powtórzyć kroki 4-10 algorytmu dla mniejszej wartości parametru wsparcia reguł rozmytych [wb09].

Proces wydobywania reguł dla probabilistyczno-rozmytej bazy wiedzy w postaci schematu algorytmu przedstawiony jest na rysunku III-16.



Rys. III-16. Algorytm generowania reguł modelu probabilistyczno-rozmytego bazującego na założeniach algorytmu Apriori (na podstawie [rud10])

Założenia przedstawionego algorytmu, które odróżniają go od innych algorytmów wyszukujących reguły asocjacji z uwzględnieniem właściwości algorytmu Apriori [agr93], są następujące:

- uwzględnienie wszystkich zmiennych (atrybutów) w otrzymanej rozmytej regule asocjacji,
- uwzględnienie konkretnych zmiennych (atrybutów) w przesłance i konkluzji reguł,
- nieuwzględnianie ograniczenia minimalnej ufności reguł przy tworzeniu reguły,
- przekształcenie postaci rozmytych reguł asocjacji do postaci reguł probabilistyczno-rozmytej bazy wiedzy (109).

3.2.3. Budowa bazy reguł z uwzględnieniem zmodyfikowanego algorytmu FP-Growth

Innym, znanim algorytmem wyszukującym ilościowe reguły asocjacji jest algorytm FP-Growth. Szczegóły algorytmu, jego zalety i wady zostały omówione w rozdziale II-5.2.2. Literatura udowadnia [bor05] [han00b] [pan06], że zastosowanie algorytmu FP-Growth do wyszukiwania ilościowych reguł asocjacji jest efektywniejsze, z punktu wiedzenia czasu trwania obliczeń, w stosunku do tradycyjnego algorytmu Apriori (efektywność algorytmów jest rozumiana pod kątem minimalizacji czasu trwania obliczeń). Artykuł [wan09] stosuje algorytm oparty na zasadach algorytmu FP-Growth do eksploracji danych w celu znalezienia rozmytych reguł asocjacji. Efektywność tego algorytmu jest zwiększoną z uwagi na wprowadzenie, wyznaczanych przez ekspertów, ocen wagowych dla zbiorów rozmytych zmiennych, które są rozważane w regułach. Użycie dodatkowych wag, które użytkownik (ekspert) musiałby zdefiniować podczas budowy bazy wiedzy dla analizowanego w dysertacji systemu, nie jest elementem pożądanym. Bowiem dla wielu danych doświadczalnych, określenie ważności terminów rozmytych zmiennych, byłoby kłopotliwe bądź też nawet niemożliwe z uwagi na nieznany charakter obiektu. Stąd też, w omawianym rozdziale analizuje się metodę rozmytego modelowania wykorzystującą ideę tworzenia rozmytych reguł asocjacji z uwzględnieniem zasad algorytmu FP-Growth, gdzie potencjalnym zbiorem częstym może być każdy zdefiniowany zbiór rozmyty. Wymieniona wada algorytmu – duże obciążenie pamięci z uwagi na rozmiar FP-Growth – nie stanowi przeszkody, gdyż wykorzystanie metody reguł asocjacji do budowy bazy reguł dla systemów wnioskujących nie pociąga za sobą dużego rozmiaru FP-drzewa. Ponadto budowa struktury FP-drzewa jest ograniczona jedynie do zapisu informacji, które zdarzenia rozmyte mogą być analizowane jednocześnie. Jest to możliwe dzięki założeniu występowania określonych zmiennych w przesłance i konkluzji reguł rozmytych.

Zmodyfikowany algorytm FP-Growth

Wejścia proponowanego algorytmu:

- zbiór I pomiarów użytych do identyfikacji modelu,
- predefiniowana baza danych (wartości lingwistyczne zmiennych rozważanych w modelu oraz definicje zbiorów rozmytych utożsamianych z tymi wartościami),
- wartość progowa minimalnego wsparcia ($min\ w$).

Wyjście: baza reguł probabilistyczno-rozmytej bazy wiedzy.

Notacja użyta do przedstawienia algorytmu jest analogiczna do podpunktu II-3.2.2.

Dla takich samych założeń wejściowych zmodyfikowanego algorytmu Apriori i FP-Growth otrzymuje się bazę wiedzy o identycznej strukturze i takich samych parametrach reguł. Inny natomiast jest sposób osiągnięcia wyniku.

Algorytm do generowania reguł dla systemu wnioskującego typu MISO z probabilistyczno-rozmytą bazą wiedzy jest przedstawiony poniżej.

Krok 1: Dyskretyzacja wartości liczbowych każdego zbioru D^i ($i=1,\dots,I$) na K rozłączne przedziały $(\mathbf{a}^{(n)}; n=1,\dots,N$ oraz \mathbf{b}) o jednakowej szerokości w przestrzeni zmiennych.

Krok 2: Na podstawie próby D , obliczenie rozkładu prawdopodobieństwa dla zdarzenia nierożmytego:

$$P_{k_1 \dots k_N k_{N+1}}(x_1, \dots, x_N, y) = P(x_1 \in a_{k_1}^{(1)}, \dots, x_N \in a_{k_N}^{(N)}, y \in b_{k_{N+1}}), k_1, \dots, k_{N+1}=1, \dots, K. \quad (146)$$

Krok 3: Zdefiniowanie kandydatów zbioru częstego 1-elementowego C_1 . Obliczenie wartości wsparcia dla zmiennych wejściowych:

$$w(A_j^{(n)}) = P(A_j^{(n)}) = \sum_{k_n=1}^K P(x_n \in a_{k_n}^{(n)}) \cdot \mu_{A_j^{(n)}}(a_{k_n}^{(n)}), j=1, \dots, |A^{(n)}|, n=1, \dots, N, \quad (147)$$

oraz zmiennej wyjściowej:

$$w(B_j) = P(B_j) = \sum_{k_{N+1}=1}^K P(y \in b_{k_{N+1}}) \cdot \mu_{B_j}(b_{k_{N+1}}), j=1, \dots, |B|. \quad (148)$$

Krok 4: Wybranie zdarzeń rozmytych ze zbioru C_1 , które spełniają założenie minimalnego wsparcia. W wyniku czego otrzymujemy zbiór F_1 składający się z rozmytych zdarzeń częstych 1-elementowych F_1 :

$$F_1 = \{A_{j^*}^{(n)} = A_j^{(n)} \mid \forall j, n : w(A_j^{(n)}) \geq \min w, B_{j^*} = B_j \mid \forall j : w(B_j) \geq \min w\}, \quad (149)$$

gdzie $A_{j^*}^{(n)}$ stanowi j^* -te 1-elementowe rozmyte zdarzenie częste dla zmiennej wejściowej x_n , $j^* \in \{1, \dots, |A^{(n)}|\}$, B_{j^*} stanowi j^* -te 1-elementowe rozmyte zdarzenie częste dla zmiennej wyjściowej y , $j^* \in \{1, \dots, |B|\}$.

Krok 5: Jeżeli zbiór F_1 nie jest zbiorem pustym to przeliczenie wartości każdego pomiaru w D^i na stopnie przynależności zaktywowych zdarzeń rozmytych:

- dla zmiennych wejściowych, na stopnie przynależności ich wartości $\{x_1^i, \dots, x_N^i\}$, do zbiorów rozmytych $A_{j_i^*}^{(n)}, n=1, \dots, N$,
- dla zmiennej wyjściowej na stopnie przynależności jej wartości y^i do zbiorów rozmytych $B_{j_i^*}$.

$$\left\{ \frac{\mu_{A_{j_i^*}^{(1)}}(a_{k_{1i}}^{(1)})}{x_1^i}, \dots, \frac{\mu_{A_{j_i^*}^{(N)}}(a_{k_{Ni}}^{(N)})}{x_N^i}, \frac{\mu_{B_{j_i^*}}(b_{k_{N+1i}})}{y^i} \right\}_{j_i^*} \text{ dla } x_1^i \in a_{k_{1i}}^{(1)}, \dots, x_N^i \in a_{k_{Ni}}^{(N)}, y^i \in b_{k_{N+1i}}, \quad (150)$$

gdzie: $n=1, \dots, N$, $i=1, \dots, I$, $j_i^* \in \{1, \dots, |A^{(n)}|\}$ dla $A_{j_i^*}^{(n)}$ lub $j_i^* \in \{1, \dots, |B|\}$ dla $B_{j_i^*}$, natomiast J_i^{**} stanowi liczbę utworzonych kombinacji zdarzeń rozmytych z i -tego zbioru wartości doświadczalnych D^i .

Wówczas, na podstawie I pomiarów, otrzymujemy zbiór C_{N+1}^{**} – zbiór $J^{**} = \sum_{i=1}^I J_i^{**}$ kombinacji z powtórzeniami $N+1$ zdarzeń rozmytych $\{A_{j_1^*}^{(1)}, \dots, A_{j_i^*}^{(N)}, B_{j_c^*}\}$ (wraz z niezerowymi stopniami przynależności poszczególnych zmiennych dla określonego zdarzenia rozmytego), których iloczyn może okazać się $N+1$ -elementowym zdarzeniem częstym. Zatem:

$$C_{N+1}^{**} = \left\{ A_{j_1^*}^{(1)}, \dots, A_{j_i^*}^{(N)}, B_{j_c^*} \right\}_{J^{**}}, i = 1, \dots, I. \quad (151)$$

Jeżeli zbiór F_1 jest pusty należy zakończyć algorytm – brak reguł rozmytych spełniających warunek minimalnego wsparcia $\min w$.

Krok 6: Wyznaczenie J^* niepowtarzających się kombinacji pojedynczych zdarzeń rozmytych dla zmiennych w przesłance i konkluzji reguły:

$$C_{N+1}^* = \left\{ A_{j_c^*}^{(1)}, \dots, A_{j_c^*}^{(N)}, B_{j_c^*} \right\}_{J^*}, C_{N+1}^* \subset C_{N+1}^{**}. \quad (152)$$

Krok 7: Obliczenie miary wsparcia w dla każdego $N+1$ -elementowego zdarzenia rozmytego $A_{j_c^*}^{(1)} \cap \dots \cap A_{j_c^*}^{(N)} \cap B_{j_c^*}$, utworzonego na podstawie elementów zbioru C_{N+1}^* :

$$\begin{aligned} \forall_{c=1, \dots, J^*} w(A_{j_c^*}^{(1)}, \dots, A_{j_c^*}^{(N)}, B_{j_c^*}) &= P(A_{j_c^*}^{(1)} \cap \dots \cap A_{j_c^*}^{(N)} \cap B_{j_c^*}) = \\ \sum_{k_1=1}^K \dots \sum_{k_{N+1}=1}^K [p_{k_1 \dots k_{N+1}}(x_1, \dots, x_N, y) \cdot T(\mu_{A_{j_c^*}^{(1)}}(a_{k_1}^{(1)}), \dots, \mu_{A_{j_c^*}^{(N)}}(a_{k_N}^{(N)}), \mu_{B_{j_c^*}}(b_{k_{N+1}}))], \end{aligned} \quad (153)$$

gdzie: $p_{k_1 \dots k_{N+1}}(x_1, \dots, x_N, y)$ wyznaczane jest na podstawie (146), natomiast symbol T oznacza operatora t-normy wg tabeli III-1.

Krok 8: Wybranie ze zbioru C_{N+1}^* rozmytych zdarzeń częstych $N+1$ -elementowych, które spełniają założenie minimalnego wsparcia i utworzenie z nich zbioru F_{N+1}^* :

$$F_{N+1}^* = \left\{ \{A_{j_c^*}^{(1)}, \dots, A_{j_c^*}^{(N)}, B_{j_c^*}\} \mid \forall c : w(A_{j_c^*}^{(1)}, \dots, A_{j_c^*}^{(N)}, B_{j_c^*}) \geq \min w \right\}. \quad (154)$$

Krok 9: Jeżeli zbiór F_{N+1}^* nie jest zbiorem pustym to utworzenie bazy reguł probabilistyczno-rozmytego modelu wiedzy na podstawie rozmytych zdarzeń częstych zbioru F_{N+1}^* . Otrzymujemy wówczas zbiór reguł elementarnych w postaci:

$$(w(A_{j_c^*}^{(1)}, \dots, A_{j_c^*}^{(N)}, B_{j_c^*})) \text{ JEŻELI } x_1 \text{ jest } A_{j_c^*}^{(1)} \text{ I } \dots \text{ I } x_N \text{ jest } A_{j_c^*}^{(N)} \text{ TO } y \text{ jest } B_{j_c^*}. \quad (155)$$

Bazę reguł w postaci (109) otrzymujemy łącząc reguły elementarne (155) o tym samym poprzedniku reguły – analogicznie do kroku 10 w algorytmie opisanym w III-3.2.2.

W przypadku, gdy zbiór F_{N+1}^* jest zbiorem pustym, należy powtórzyć kroki 4-9 algorytmu dla mniejszej wartości parametru minimalnego wsparcia reguł rozmytych ($\min w$).

Algorytmy opisane w rozdziałach III-3.2.2 i III-3.2.3 wymagają znajomości rozkładu prawdopodobieństwa zdarzenia nierożmytego $p_{k_1 \dots k_N k_{N+1}}(x_1, \dots, x_N, y)$, w oparciu o który obliczane jest prawdopodobieństwo odpowiedniego zdarzenia rozmytego $P(A_j^{(1)} \cap \dots \cap A_j^{(N)} \cap B_j)$ zgodnie ze wzorem (50) (na podstawie wzorów (142), (153)). Zastosowanie w tym celu wielowymiarowych macierzy pociąga za sobą złożoność pamięciową algorytmów. Założenie, iż każdy pomiar danych jest jednakowo prawdopodobny, pozwala na mierzenie prawdopodobieństwa zdarzeń rozmytych z wykorzystaniem mocy zbiorów rozmytych (def. 8). Obliczenia wówczas nie dotyczą pełnego zakresu przestrzeni zmiennych $X_1 \times \dots \times X_N \times Y$, a jedynie tych rozłącznych przedziałów wartości zmiennych, do których należą badane dane empiryczne.

Opisany powyżej zmodyfikowany algorytm FP-Growth zostanie przedstawiony w wersji z wyliczaniem prawdopodobieństw zdarzeń rozmytych na podstawie mocy zbioru rozmytego (stąd skrót FP-Growth (P), gdzie P ang. *Power* oznacza moc zbioru rozmytego).

Zmodyfikowany algorytm FP-Growth (P)

Krok 1: Krok 1 zmodyfikowanego algorytmu FP-Growth.

Krok 2: Zdefiniowanie kandydatów zbioru częstego 1-elementowego C_1 . Obliczenie wartości wsparcia dla zmiennych wejściowych:

$$w(A_j^{(n)}) = P(A_j^{(n)}) = \frac{\sum_{i=1}^I \mu_{A_j^{(n)}}(a_{k_{ni}}^{(n)})}{I}, \quad j=1, \dots, |A^{(n)}|, \quad n=1, \dots, N, \quad (156)$$

oraz zmiennej wyjściowej:

$$w(B_j) = P(B_j) = \frac{\sum_{i=1}^I \mu_{B_j}(b_{k_{N+1i}})}{I}, \quad j=1, \dots, |B|, \quad (157)$$

przy czym $x_n^i \in a_{k_{ni}}^{(n)}$ oraz $y^i \in b_{k_{N+1i}}$, gdzie x_n^i, y^i stanowią wartości i -tych pomiarów zmiennych x_n, y .

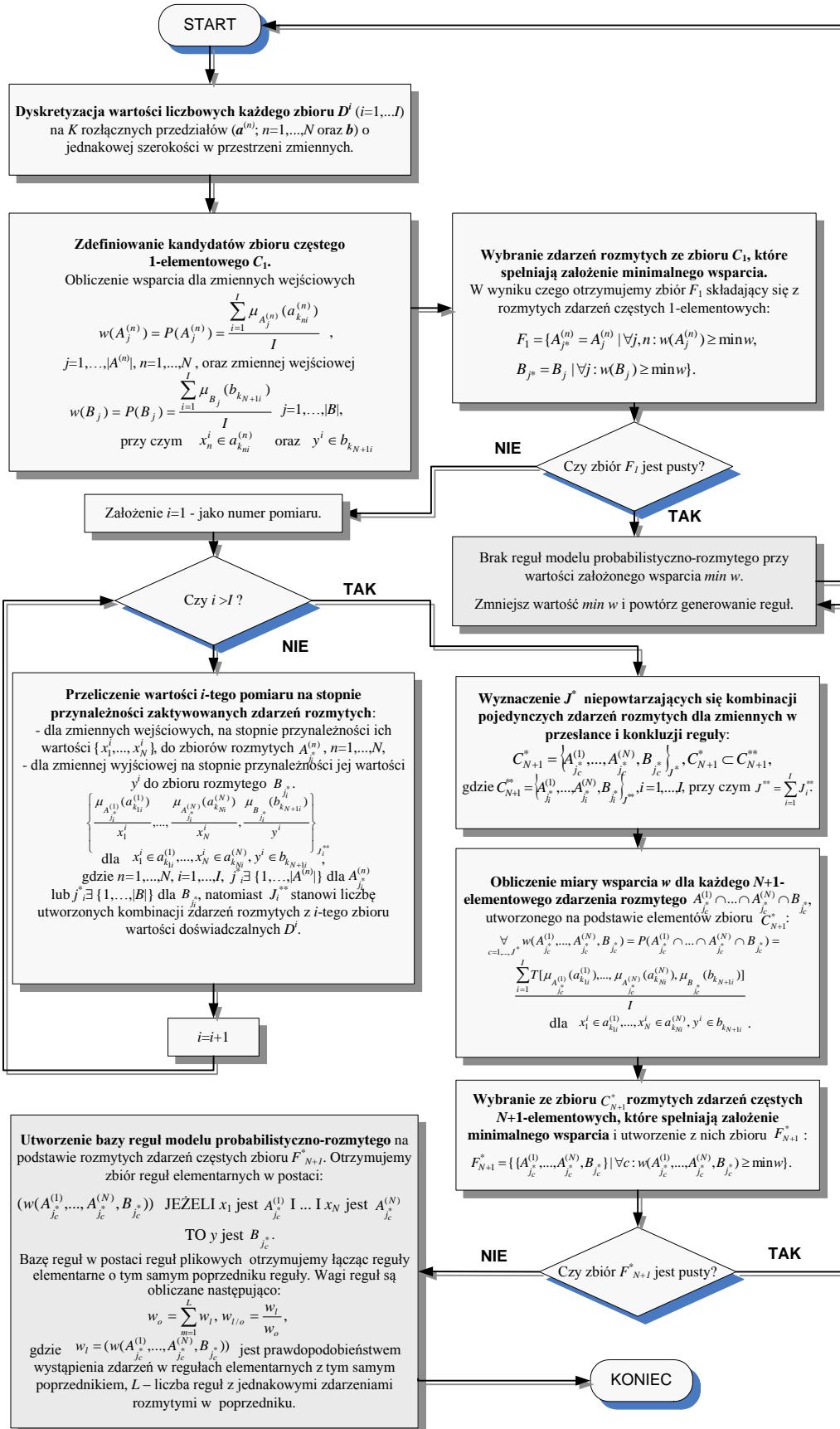
Krok 3-5: Krok 4-6 zmodyfikowanego algorytmu FP-Growth.

Krok 6: Obliczenie miary wsparcia w dla każdego $N+1$ -elementowego zdarzenia rozmytego $A_{j_c^*}^{(1)} \cap \dots \cap A_{j_c^*}^{(N)} \cap B_{j_c^*}$, utworzonego na podstawie elementów zbioru C_{N+1}^* :

$$\begin{aligned} \forall_{c=1,\dots,J^*} w(A_{j_c^*}^{(1)}, \dots, A_{j_c^*}^{(N)}, B_{j_c^*}) &= P(A_{j_c^*}^{(1)} \cap \dots \cap A_{j_c^*}^{(N)} \cap B_{j_c^*}) = \\ \frac{\sum_{i=1}^I T[\mu_{A_{j_c^*}^{(1)}}(a_{k_{1i}}^{(1)}), \dots, \mu_{A_{j_c^*}^{(N)}}(a_{k_{Ni}}^{(N)}), \mu_{B_{j_c^*}}(b_{k_{N+1i}})]}{I}, \text{ dla } x_1^i &\in a_{k_{1i}}^{(1)}, \dots, x_N^i \in a_{k_{Ni}}^{(N)}, y^i \in b_{k_{N+1i}}. \quad (158) \end{aligned}$$

Krok 7-8: Krok 8-9 zmodyfikowanego algorytmu FP-Growth.

Schemat blokowy zmodyfikowanego algorytmu FP-Growth (P) został przedstawiony na rysunku III-17.



Rys. III-17. Algorytm generowania reguł modelu probabilistyczno-rozmytego bazujący na założeniach algorytmu FP-Growth (P)

3.3. Porównanie algorytmów generowania bazy reguł

Efektywność algorytmów jest określana umiejętnością generowania bazy wiedzy w stosunku do nakładów czasowych, zużytych do realizacji zadania. Zatem, w celu wybrania efektywnego algorytmu służącego do budowy probabilistyczno-rozmytej bazy wiedzy przedstawiona zostanie analiza czasu generowania bazy reguł za pomocą opisanych algorytmów: algorytmu stosującego pełny, empiryczny rozkład prawdopodobieństwa, zmodyfikowanego algorytmu Apriori oraz zmodyfikowanego algorytmu FP-Growth i FP-Growth (P). Aby móc porównać wyniki z rozdziału III-3.1, do analizy zastosowano te same dane pomiarowe udziału frakcji lekkiej węgla x w ilości 491 próbek. Z danych zbudowano ciągi uczące dla modeli dyskretnych procesów autoregresji AR(n):

$$\hat{x}_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-n}), \quad (159)$$

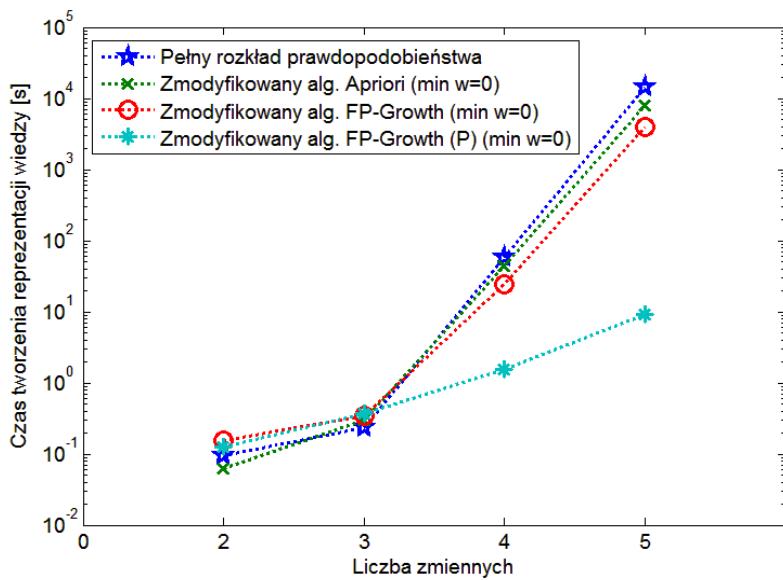
z krokiem $\tau=1$, gdzie n jest rzędem regresji. Wówczas model takiego procesu zawiera $n+1$ zmiennych. Przestrzeń wartości zmiennej $x(t) \in [0.24, 0.57]$ została początkowo podzielona na 35 (ozn. a_i , gdzie $i=1, \dots, 35$) rozłącznych przedziałów o szerokości 0.00943 każdy. Podobnie, jak w przykładzie omówionym w rozdziale III-3.1, dla każdej zmiennej zdefiniowano k równomiernie rozłożonych zbiorów rozmytych A_j o stopniach przynależności $\mu_{A_j}(a_i)$ przeliczanych z funkcji trójkątnych, gdzie $k=3, \dots, 9$, $j=1, \dots, k$. Zbiory rozmyte dla pozostałych zmiennych modelu rozmytego są definiowane analogicznie. W celu utworzenia iloczynu zbiorów rozmytych, podczas generowania reguł, w przesłance rozmytej reguły użyto iloczynu algebraicznego, jako operatora t-normy.

3.3.1. Analiza czasu generowania i rozmiarów bazy reguł z pełnym rozkładem prawdopodobieństwa

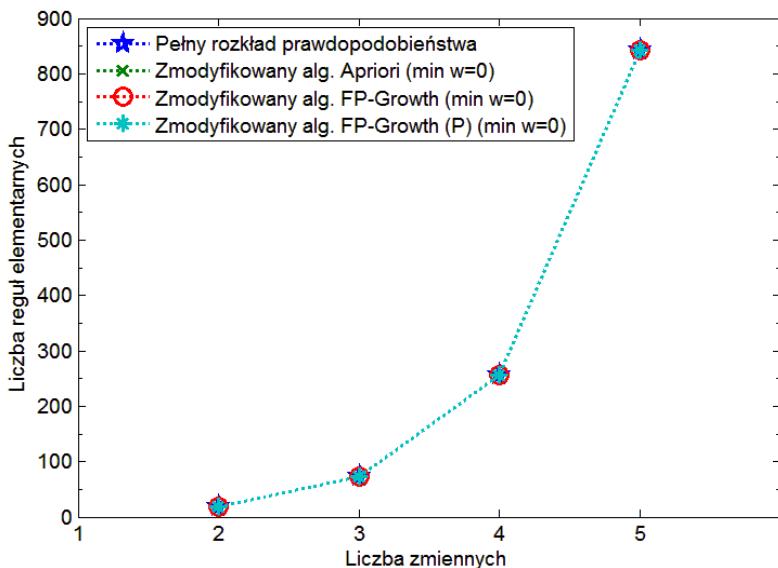
W przypadku generowania bazy reguł za pomocą zmodyfikowanego algorytmu Apriori oraz algorytmu FP-Growth (FP-Growth (P)), wagi otrzymanych reguł rozmytych modelu (109) zawierają pełny rozkład prawdopodobieństwa zdarzeń rozmytych w przesłance reguły oraz pełny rozkład warunkowy zdarzeń znajdujących się w następniku reguły, wówczas, gdy wartość minimalnego wsparcia dla wyszukiwanych rozmytych reguł asocjacji jest równy zero. W niektórych przypadkach wartości minimalnego wsparcia bliskie zeru mogą również dawać w wyniku reguły odzwierciedlające pełny rozkład prawdopodobieństwa zdarzeń znajdujących się w regule. Wówczas, zastosowanie wszystkich wymienionych algorytmów, dla takich samych założeń wejściowych, daje rezultat w postaci identycznej bazy reguł probabilistyczno-rozmytego modelu wiedzy. Różny jest natomiast czas dochodzenia do tego samego wyniku.

Rysunek III-18 przedstawia zależność czasu tworzenia reprezentacji wiedzy od liczby zmiennych modelu rozmytego procesu dynamicznego autoregresji AR(n) ($n=1, 2, 3, 4$), dla czterech porównywanych algorytmów. W modelu dla każdej zmiennej zdefiniowano jednakowo po 5 zbiorów rozmytych. Jak pokazuje wykres, przy mniejszej liczbie zmiennych ($N=2$) najszybszym algorymem okazuje się zmodyfikowany algorytm Apriori, najdłużej zaś odnajdywał rozmytą bazę reguł zmodyfikowany algorytm FP-Growth. W przypadku większej liczby zmiennych, czas utworzenia reprezentacji wiedzy za pomocą algorytmu generującego model wiedzy z wagami o pełnym rozkładzie prawdopodobieństwa jest najdłuższy i najszybciej wzrasta wraz ze zwiększeniem się liczby zmiennych. Najkrótszy czas generowania reguł otrzymano dla zmodyfikowanego algorytmu FP-Growth (P).

Zależność liczby otrzymanych reguł elementarnych od ilości zmiennych modelu przedstawia rysunek III-19. Otrzymano poprawne identyczne wyniki dla wszystkich algorytmów.

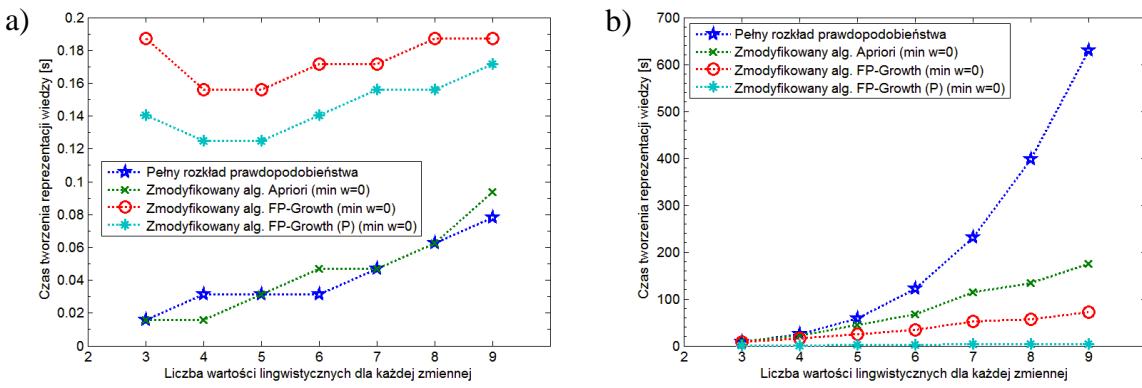


Rys. III-18. Zależność czasu tworzenia reprezentacji wiedzy od liczby zmiennych modelu dla porównywanych algorytmów



Rys. III-19. Zależność liczby reguł elementarnych modelu wiedzy od liczby zmiennych modelu dla porównywanych algorytmów

Dla modeli wiedzy procesu autoregresji AR(1) i autoregresji AR(3) udziału frakcji lekkiej węgla zmieniano liczbę zbiorów rozmytych w zakresie 3-9. Jedynie w przypadku dwóch zmiennych można zauważać przewagę algorytmu naiwnego i zmodyfikowanego algorytmu Apriori. Na rysunku III-18 obserwujemy już inną kolejność czasów wykonywania algorytmów dla AR(3) i 5 zbiorów rozmytych (najszybszy zmodyfikowany FP-Growth (P), kolejno FP-Growth, Apriori, najwolniejszy algorytm naiwny). Z rysunku III-20 wynika, że kolejność ta utrzymuje się dla każdej liczby zbiorów rozmytych. Wraz ze wzrostem liczby zbiorów rozmytych, czas generowania reguł wzrasta najszybciej dla naiwnego algorytmu, najwolniej dla zmodyfikowanego algorytmu FP-Growth (P).

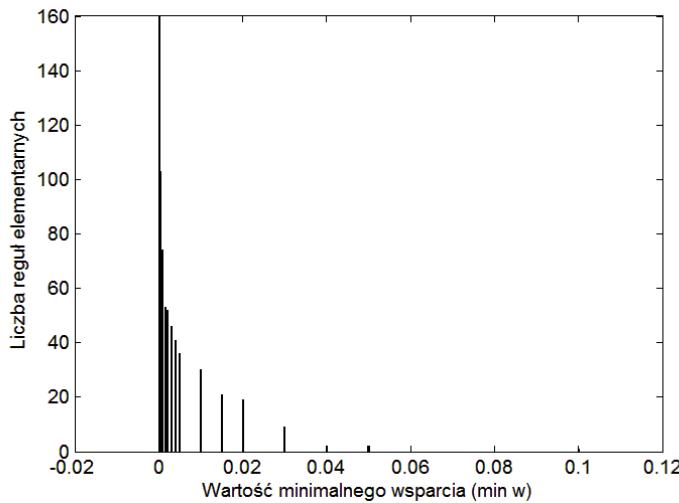


Rys. III-20. Zależność czasu tworzenia reprezentacji wiedzy od liczby zbiorów rozmytych definiowanych dla każdej zmiennej w przypadku: a) procesu AR(1) udziału frakcji lekkiej węgla,
b) procesu AR(3) udziału frakcji lekkiej węgla

3.3.2. Analiza czasu generowania i rozmiarów bazy reguł z niepełnym rozkładem prawdopodobieństwa

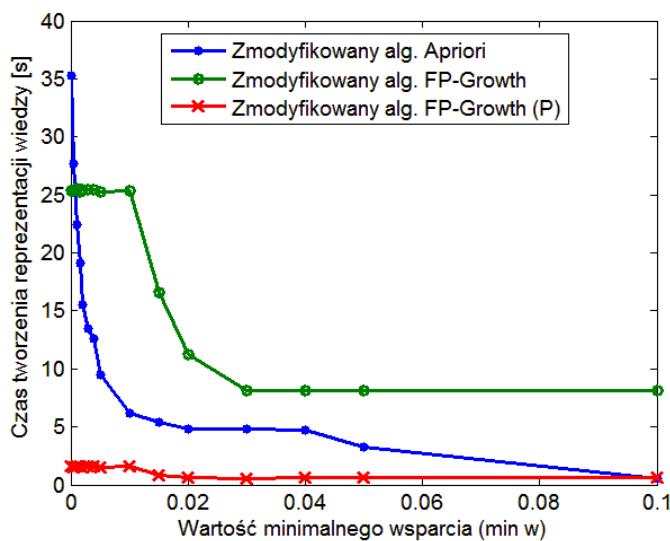
W celu zbadania możliwości ograniczenia złożoności modelu wiedzy (ilość reguł elementarnych) oraz czasu jego generowania zostaną przedstawione zależności różnych parametrów modelu od wartości minimalnego wsparcia, przy generowaniu reguł z zastosowaniem zmodyfikowanego algorytmu Apriori i zmodyfikowanego algorytmu FP-Growth (FP-Growth (P)).

Wstępnie, analizie poddano rozmyty model dynamiczny autoregresji AR(3) dla udziału frakcji lekkiej węgla, w którym dla 35 rozłącznych przedziałów wartości zmiennych zdefiniowano po 5 zbiorów rozmytych dla każdej zmiennej. Przy zastosowaniu algorytmów z jednakowymi założeniami wejściowymi, w tym również z jednakowymi wartościami minimalnego wsparcia otrzymujemy tą samą probabilistyczno-rozmytą bazę wiedzy. Zależność liczby reguł elementarnych modelu w stosunku do wartości minimalnego wsparcia przedstawia rysunek III-21. Wraz z ograniczeniem wartości prawdopodobieństwa wystąpienia jednoczesnego zdarzeń w przesłance i konkluzji, liczba reguł elementarnych z niezerowymi wagami maleje – początkowo bardzo szybko, następnie wolniej. Wobec tego, im większa jest wartość minimalnego wsparcia, tym zmniejsza się złożoność modelu wiedzy. Należy jednak określić, rozsądny poziom złożoności modelu, przy którym otrzymujemy wystarczającą zdolność modelu do dokładnego odwzorowania systemu rzeczywistego. Zagadnienie to jest zadaniem identyfikacji, opisanej w rozdziale III-4 niniejszej pracy.

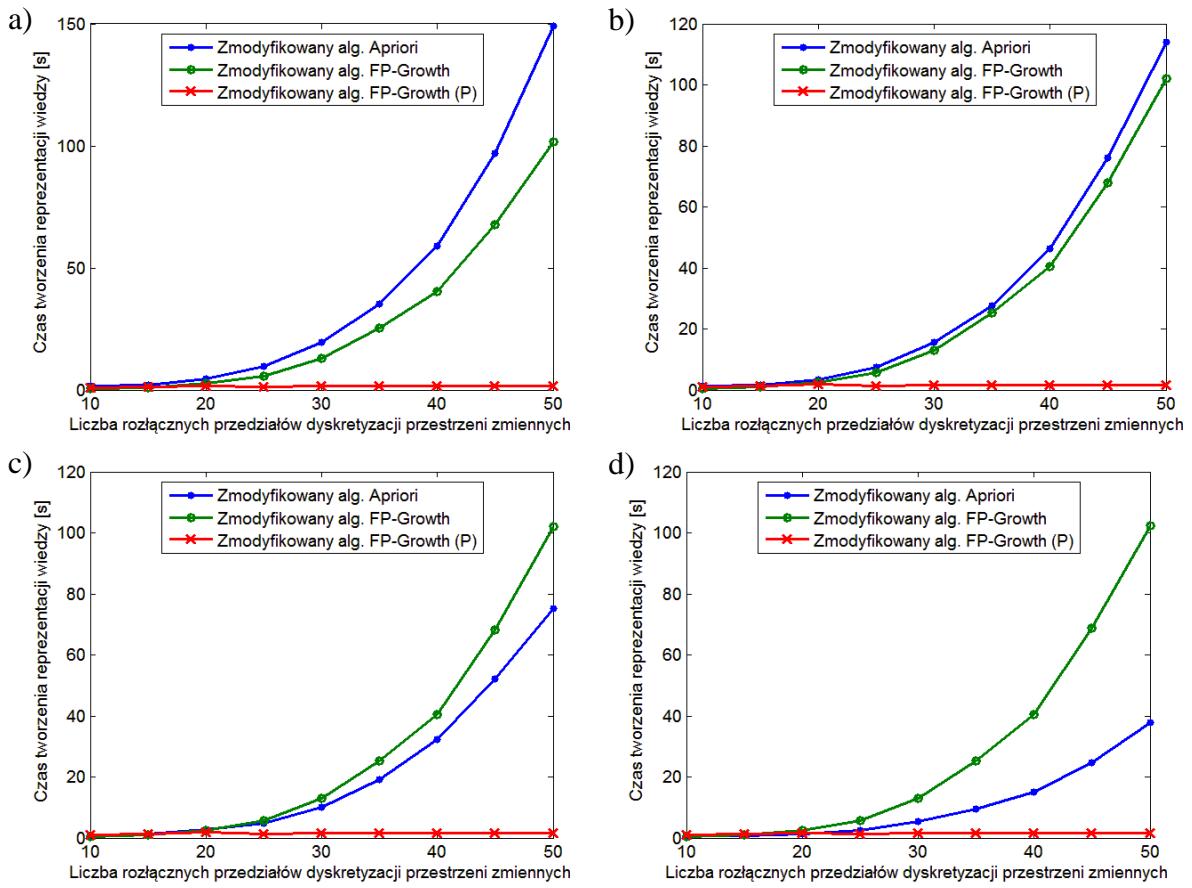


Rys. III-21. Zależność liczby reguł elementarnych modelu wiedzy od wartości minimalnego wsparcia

Zależność czasu tworzenia probabilistyczno-rozmytej bazy wiedzy od wartości minimalnego wsparcia w regułach dla porównywanych algorytmów przedstawia rysunek III-22. Okazuje się, że zmodyfikowany algorytm FP-Growth jest efektywniejszy od zmodyfikowanego algorytmu Apriori jedynie przy generowaniu reguł o wagach stanowiących pełne lub bliskie pełnego rozkładu prawdopodobieństwa zdarzeń rozmytych w regule. Najbardziej efektywny, dla każdej wartości minimalnego wsparcia różnej od zera, jest zmodyfikowany algorytm FP-Growth (P). Algorytm ten nie wylicza wag reguł na bazie niepełnego (a przy $min w=0$ - pełnego) rozkładu prawdopodobieństwa zdarzeń nieroźmytych, ale zlicza wartość mocy przecięcia zbiorów rozmytych, uzyskanej na podstawie każdego pomiaru. Stąd też jego złożoność czasowa jest uzależniona głównie od liczby rekordów danych uczących, co przy podanych założeniach wejściowych daje szybsze rezultaty obliczeń. Ponieważ liczba danych uczących jest stała, stąd przy określonej strukturze danych doświadczalnych otrzymujemy stały poziom czasu wykonania algorytmu dla danego zakresu wartości minimalnego wsparcia, przy którym otrzymujemy jednakowe pojedyncze zdarzenia częste.



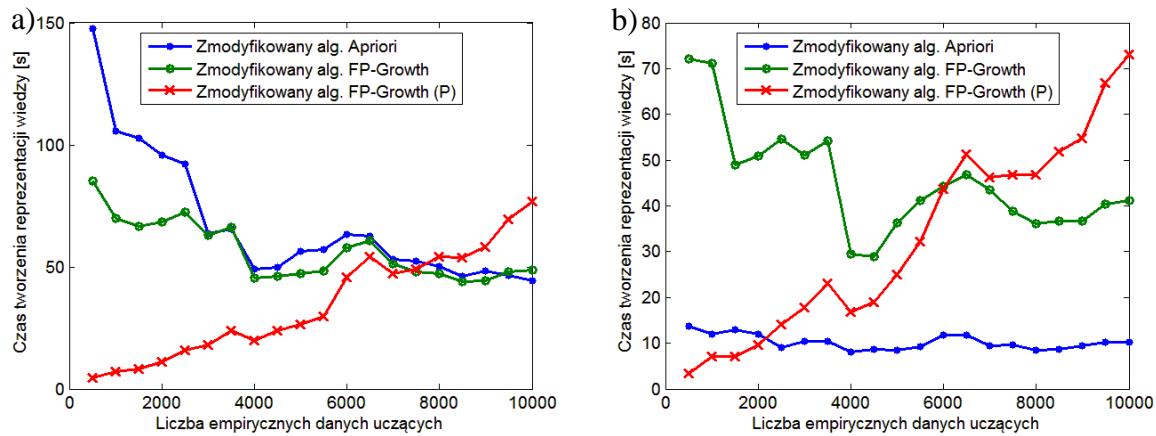
Rys. III-22. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia



Rys. III-23. Zależności czasu tworzenia reprezentacji wiedzy od liczby rozłącznych przedziałów dyskretyzacji przestrzeni wartości zmiennych dla porównywanych algorytmów oraz dla minimalnego wsparcia o wartościach równych odpowiednio: a) $\min w=0.0001$, b) $\min w=0.0005$, c) $\min w=0.0016$, d) $\min w=0.005$

Rysunek III-23 zawiera porównanie algorytmów na podstawie wykresów przedstawiających zależności czasu tworzenia reprezentacji reguł od liczby rozłącznych przedziałów dyskretyzacji przestrzeni wartości zmiennych, dla wybranych wartości minimalnego wsparcia. W przypadku zmodyfikowanych algorytmów Apriori i FP-Growth, wzrost liczby rozłącznych przedziałów powoduje niemal wykładniczy wzrost czasu trwania obliczeń. Natomiast w przypadku zastosowania zmodyfikowanego algorytmu FP-Growth (P) takiej zależności się nie zauważa.

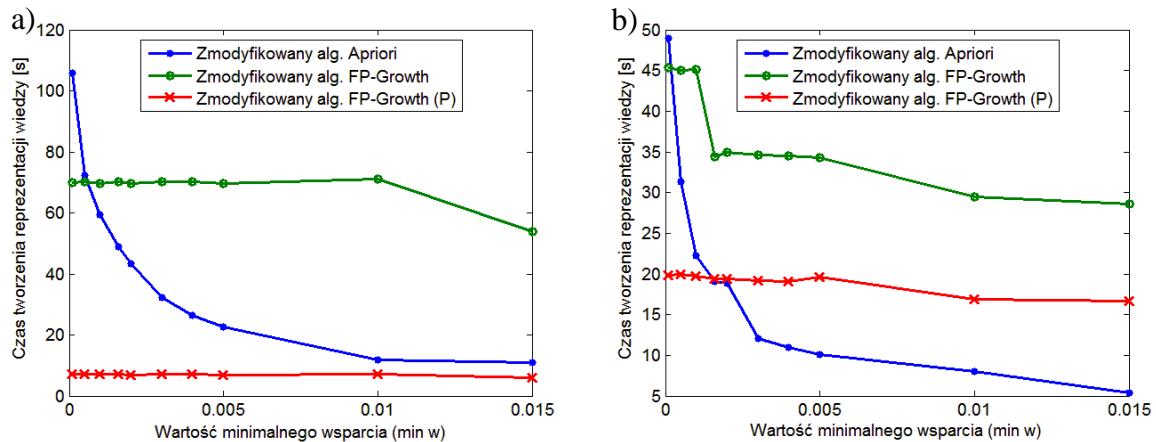
Analizowano również wpływ liczby rekordów danych empirycznych, służących do uczenia systemu, na czas generowania bazy reguł. W tym celu zbudowano model wiedzy dla procesu AR(3) prędkości wiatru [km/s]. Dla każdej zmiennej modelu utworzono po 7 równomiernie rozłożonych zbiorów rozmytych w kształcie funkcji trójkątnej. Model uczyono kolejno 500, 1000, ..., 10000 rekordami pomiarów prędkości wiatru, zmierzonymi w okresie od 01.01.2010 do 08.01.2010 w odstępach minutowych. Wyniki analiz zamieszczono na wykresach III-24÷III-25.



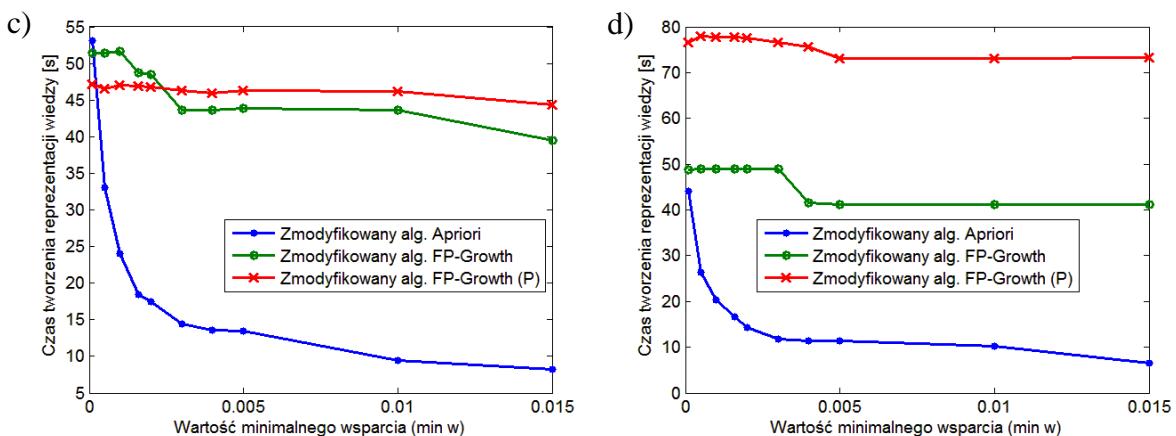
Rys. III-24. Zależność czasu tworzenia reprezentacji wiedzy od liczby danych uczących dla różnych algorytmów, przy założeniach: a) $\text{min } w=0.0001$, b) $\text{min } w=0.01$

Zauważono, iż wzrost liczby danych empirycznych powoduje, że czas tworzenia reprezentacji wiedzy, w przypadku zmodyfikowanego algorytmu Apriori i FP-Growth maleje, zwłaszcza przy wartości minimalnego wsparcia bliskiej zeru. Jest to związane z malejącymi wartościami prawdopodobieństw zdarzeń mniej częstych w przypadku zwiększenia próby danych uczących. Natomiast dla zmodyfikowanego algorytmu FP-Growth (P) zwiększenie liczby empirycznych danych powoduje wydłużenie czasu trwania obliczeń.

Biorąc pod uwagę wybór algorytmu, można powiedzieć, że przy określonych założeniach modelu wiedzy (łącznie z wartością minimalnego wsparcia), do pewnej wartości liczby danych empirycznych, chcąc najszybciej uzyskać wynik w postaci reguł, należy zastosować zmodyfikowany algorytm PF-Growth (P), powyżej tej wartości – zmodyfikowany algorytm Apriori.



Rys. III-25. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia dla różnych algorytmów, przy wykorzystaniu: a) 1 tys. rekordów uczących, b) 4 tys. rekordów uczących



Rys. III-25. c,d. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia dla różnych algorytmów, przy wykorzystaniu: c) 7 tys. rekordów uczących, d) 10 tys. rekordów uczących

W dodatku A pracy zamieszczono analogiczne wykresy analiz dla przykładowych modeli wiedzy z uwzględnieniem procesów AR(3) i AR(4) prędkości wiatru.

3.4. Wnioski

Aby wybrać efektywny algorytm służący do generowania reprezentacji wiedzy w analizowanym systemie wnioskującym, zostały przebadane zależności czasu i rozmiaru baz wiedzy od różnych parametrów struktury systemu. Ponieważ liczba otrzymanych reguł elementarnych jest jednakowa dla wszystkich algorytmów, o efektywności metody świadczy czas generowania reguł rozmytych. Algorytm generujący reguły z pełnym rozkładem prawdopodobieństwa, zmodyfikowany algorytm Apriori oraz zmodyfikowany algorytm FP-Growth, wykorzystują wielowymiarowe macierze do budowy bazy wiedzy. Powoduje to, przyśpieszenie ich działania, jednakże w przypadku większej liczby zmiennych wejściowych systemu (powyżej 4) algorytmy wymagają dużych zasobów pamięciowych, co często uniemożliwia ich użycie. Zatem, w takich przypadkach sprawdza się jedynie zmodyfikowany algorytm FP-Growth (P) a porównania algorytmów ograniczono jedynie do generowania baz wiedzy dla systemów typu MISO o mniej niż 5 wejściach.

Wykresy zależności pokazały, że prawie dla wszystkich założeń wejściowych najszybszym algorymem okazał się zmodyfikowany algorytm Apriori bądź zmodyfikowany algorytm FP-Growth (P). Wyjątkami były sytuacje dla systemu typu SISO lub systemu MISO o dwóch wejściach (rys. III-18, III-20), gdzie wytycznym było otrzymanie pełnego rozkładu prawdopodobieństwa dla zdarzeń w regułach. Wówczas sprawdzanie warunków wiarygodności reguł rozmytych powodowało dodatkową złożoność czasową więc metoda generująca reguły o pełnym rozkładzie prawdopodobieństwa zdawała się być bardziej efektywna. Jednakże z uwagi na znakomą różnicę w czasach trwania obliczeń oraz niemożność ograniczania złożoności modelu, metoda generująca reguły o pełnym rozkładzie prawdopodobieństwa nie będzie brana pod uwagę podczas tworzenia modułu narzędziowego do tworzenia systemu z probabilistyczno-rozmytą bazą wiedzy.

Zmodyfikowany algorytm FP-Growth okazał się efektywnym algorymem tylko w jednym przypadku (rys. III-24), przy założeniach: systemu o trzech wejściach, siedmiu zbiorów rozmytych dla każdego wejścia, minimalnego wsparcia na poziomie zero lub bliskim zero oraz większej liczbie danych pomiarowych. Jednakże, podobnie jak wyżej, z uwagi na znakomą różnicę w czasach trwania obliczeń możemy go z powodzeniem zastąpić zmodyfikowanym algorymem Apriori.

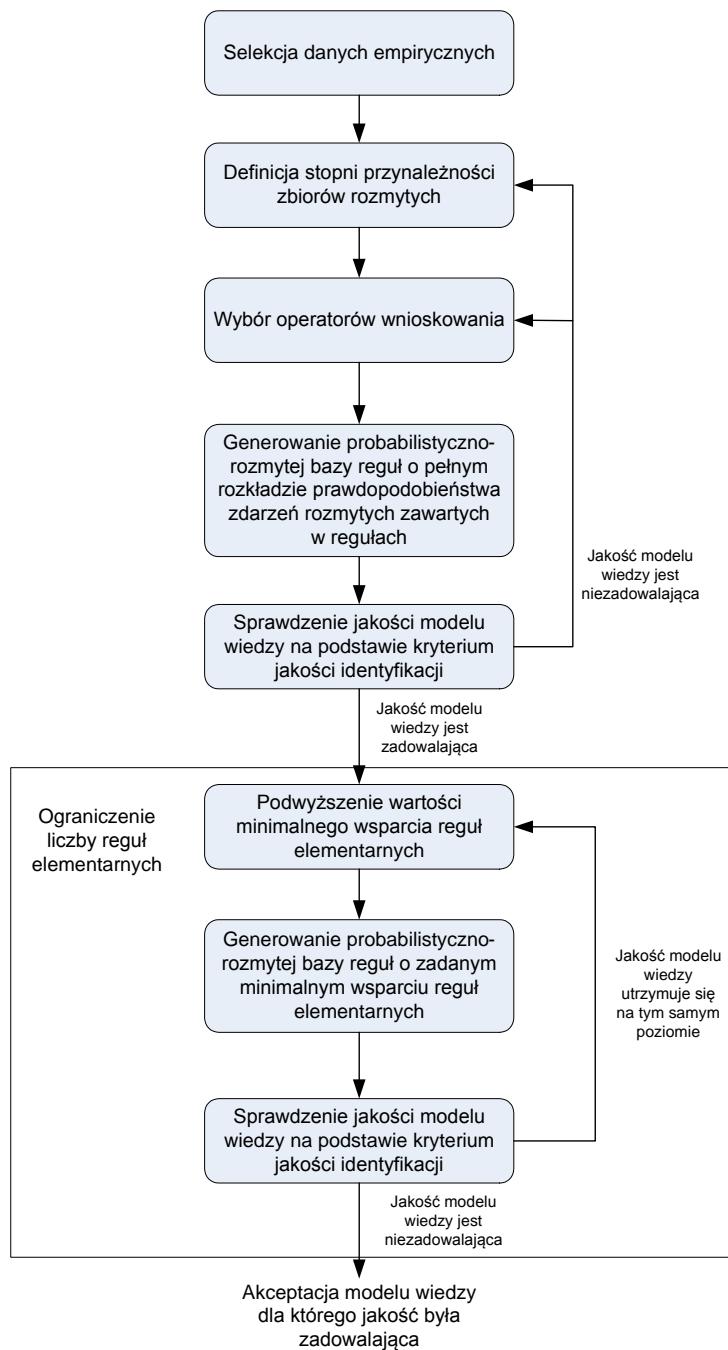
W przypadku algorytmu FP–Growth (P) analiza wykazała ścisłą zależność czasu generowania reprezentacji wiedzy od liczby rekordów danych uczących. W przypadku większej liczby danych wykorzystanie tej metody staje się mniej efektywne, jednakże dla systemu o wielu wejściach jej wykorzystanie jest jedyną rozpatrywaną możliwością tworzenia probabilistyczno-rozmytej bazy reguł. Zauważa się również, niezależność algorytmu FP–Growth (P) od wartości minimalnego wsparcia oraz liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych.

W przypadku zmodyfikowanego algorytmu Apriori czas generowania reprezentacji wiedzy skraca się wraz ze wzrostem wartości minimalnego wsparcia reguł. Zauważa się jednak, szybki wzrost czasu generowania reguł, wraz ze zwiększeniem się liczby zmiennych, liczby wartości lingwistycznych zdefiniowanych dla każdej zmiennej oraz liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych.

Można zatem wnioskować, iż na wybór najefektywniejszego algorytmu mają wpływ następujące parametry: liczba zmiennych systemu, liczba zdefiniowanych wartości rozmytych dla każdej zmiennej wejściowej i wyjściowej, liczba rozłącznych przedziałów dyskretyzacji wartości zmiennych, liczba rekordów danych uczących oraz wartość minimalnego wsparcia dla reguł rozmytych. Z uwagi na trudność jednoznacznego stwierdzenia, która z metod najszybciej utworzy probabilistyczno-rozmytą bazę wiedzy przy danych założeniach, postanowiono zbudować system podejmowania decyzji, który przy ww. parametrach wejściowych określi, który z algorytmów (zmodyfikowany algorytm Apriori czy zmodyfikowany algorytm FP–Growth (P)) powinno się wziąć pod uwagę. Bazę wiedzy systemu podejmującego decyzję zbudowano w oparciu o liczne symulacje komputerowe, których częściowe wyniki zawarto na wykresach w rozdziale III-3.3 oraz w dodatku A pracy. Szczegóły utworzonego systemu zamieszczono w rozdziale dotyczącym zastosowań systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy (rozdział IV-2).

4. Metoda identyfikacji modelu wiedzy w systemie z probabilistyczno-rozmytą bazą reguł

Mając na uwadze algorytmy generowania reguł probabilistyczno-rozmytej bazy wiedzy, które zostały opisane w rozdziałach III-2.2 i III-2.3, zaproponowano metodę identyfikacji modelu wiedzy. Kolejne etapy metody identyfikacji przedstawiono na rysunku III-26.



Rys. III-26. Identyfikacja modelu wiedzy w systemie z probabilistyczno-rozmytą bazą wiedzy

Jako kryterium jakości identyfikacji można uznać odległość między wyjściem obiektu (na podstawie empirycznych danych uczących lub/i danych testujących) i wyjściem utworzonego systemu dla tego samego wejścia. W niniejszej dysertacji jako kryterium jakości identyfikacji, czyli kryterium dopasowania modelu wiedzy do danych empirycznych zastosowano minimalizację wartości pierwiastka błędu średniokwadratowego RMSE lub procentu błędnych odpowiedzi systemu (w zależności od zastosowania analizowanego systemu wnioskującego).

5. Wnioskowanie w oparciu o probabilistyczno-rozmytą bazę wiedzy

Aby móc uzyskać ilościowy wynik na wyjściu systemu z probabilistyczno-rozmytą bazą wiedzy, niezbędne staje się określenie operacji logicznych w bloku wnioskowania, których realizacja pozwoli na wnioskowanie w oparciu o zadane fakty i otrzymaną bazę wiedzy. Szczegółowe informacje na temat uogólnionych reguł wnioskowania rozmytego opisano w rozdziale II-3.4. Mechanizm wnioskujący w oparciu o bazę wiedzy z regułami stanowiącymi prawdopodobieństwa zdarzeń zapisanych w regułach została opisany w [wb07] oraz [wb10].

W niniejszej dysertacji, dla systemu typu MISO z probabilistyczno-rozmytą bazą wiedzy w postaci następujących reguł plikowych³:

$$\left\{ \text{Jezeli } \left(\bigwedge_{n=1}^N x_n \text{ jest } A_o^{(n)}, z \text{ wagą } w_o \right) \text{ To } \left(\bigwedge_{l=1}^L y \text{ jest } B_{l/o}, z \text{ wagą } w_{l/o} \right) \right\}_{o=1}^O \quad (160)$$

stosuje się metodę wnioskowania zgodnie z *uogólnioną regułą wnioskowania modus ponendo ponens*. Konkluzję wyprowadza się z pojedynczych reguł a następnie tworzy się zagregowane wyjście rozmyte (metoda WA, przedstawiona w rozdziale III-3.4).

Rozważmy fakt, taki że $x_1 \text{ jest } A^{(1)}$... $x_N \text{ jest } A^{(N)}$, przy czym zbiory rozmyte $A^{(1)}, \dots, A^{(N)}$ są zbliżone do zbiorów $A_o^{(1)}, \dots, A_o^{(N)}$, $o=1, \dots, O$.

Śledząc proces wnioskowania w oparciu o rozmytą bazę wiedzy postaci (160), niezbędne staje się określenie relacji rozmytej $R_{l/o}$ dla pojedynczej reguły elementarnej:

$$\text{Jezeli } \left(\bigwedge_{n=1}^N x_n \text{ jest } A_o^{(n)}, z \text{ wagą } w_o \right) \text{ To } y \text{ jest } B_{l/o}, z \text{ wagą } w_{l/o}. \quad (161)$$

Przy założeniu logicznej interpretacji reguł (patrz II-3.2.1), relacja rozmyta $R_{l/o}$ obliczana jest następująco:

$$\mu_{R_{l/o}}(x_1, \dots, x_N, y) = I \left\{ T_1(\mu_{A_o^{(1)}}(x_1), \dots, \mu_{A_o^{(N)}}(x_N)), \mu_{B_{l/o}}(y) \right\}, \quad (162)$$

gdzie funkcja I jest wybranym operatorem implikacji rozmytej (tab. II-5), natomiast T_1 oznacza wybrany operator t-normy, odzwierciedlający operację obliczania stopnia prawdziwości przesłanki, złożonej z przesłanek prostych, połączonych ze sobą spójnikiem logicznym I (\wedge).

Stosując szczególną interpretację reguł typu Mamdaniego, relacja rozmyta $R_{l/o}$ przyjmuje postać:

³ oznaczenia wzoru są analogiczne do wzoru (109)

$$\mu_{R_{l/o}}(x_1, \dots, x_N, y) = \text{MIN} \{ T_1(\mu_{A_o^{(1)}}(x_1), \dots, \mu_{A_o^{(N)}}(x_N)), \mu_{B_{l/o}}(y) \}. \quad (163)$$

Natomiast, gdy zastosujemy interpretację rozmytych reguł JEŻELI-TO wg Larsena, relację rozmytą otrzymuje się z zależności:

$$\mu_{R_{l/o}}(x_1, \dots, x_N, y) = T_1(\mu_{A_o^{(1)}}(x_1), \dots, \mu_{A_o^{(N)}}(x_N)) \cdot \mu_{B_{l/o}}(y). \quad (164)$$

Wówczas rozmyty wniosek dla reguły elementarnej (161) o indeksie l/o wynosi:

$$\mu_{B_{l/o}}(y) = \sup_{x_1, \dots, x_N \in X_1 \times \dots \times X_N} T_2 \{ T_1(\mu_{A_o^{(1)}}(x_1), \dots, \mu_{A_o^{(N)}}(x_N)), \mu_{R_{l/o}}(x_1, \dots, x_N, y) \}, \quad (165)$$

przy czym zbiory rozmyte $A^{(1)}, \dots, A^{(N)}$, zbliżone do zbiorów $A_o^{(1)}, \dots, A_o^{(N)}$, stanowią wartości wejściowe modelu rozmytego.

Przedstawiany w dysertacji system, pozwoli na wnioskowanie jedynie na podstawie ostrych (numerycznych) wartości wejść x^*_1, \dots, x^*_N . Wówczas proces wnioskowania można uprościć. Uwzględniając powyższe założenie, mechanizm wnioskowania będziemy rozważać jako kolejne etapy tj.:

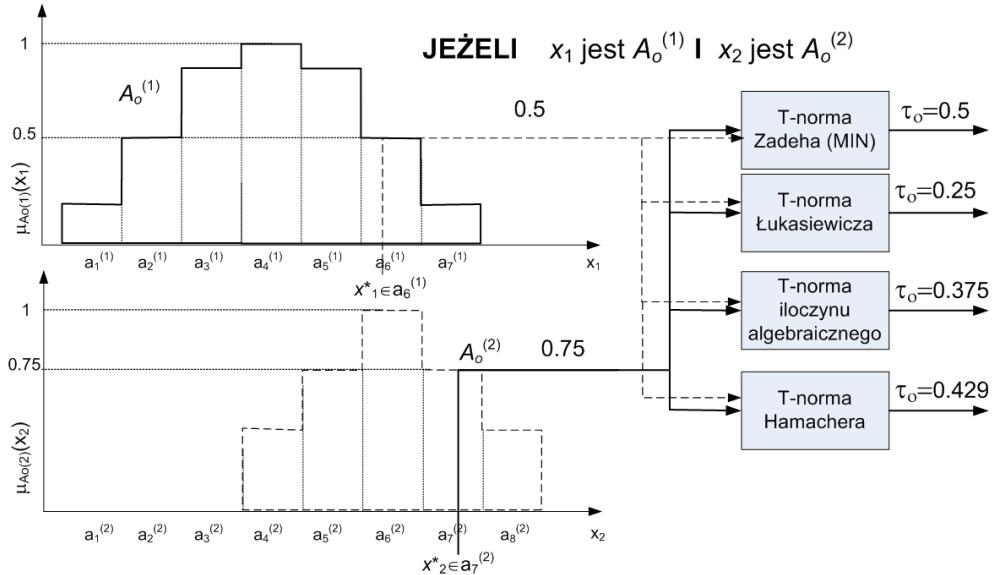
- a) wyprowadzenie stopnia aktywacji reguł plikowych,
- b) wyprowadzenie rozmytego wniosku reguł elementarnych,
- c) wyprowadzenie rozmytego wniosku reguł plikowych,
- d) wyprowadzenie rozmytego wniosku rozmytej bazy wiedzy.

5.1. Wyprowadzenie stopnia aktywacji reguł plikowych

W przypadku ostrych wartości wejściowych $x_1 = x^*_1, \dots, x_N = x^*_N$, **stopień aktywacji o -tej reguły plikowej** (161) określający stopień prawdziwości przesłanki tej reguły, złożonej z przesłanek prostych połączonych spójnikiem logicznym I , jest równy:

$$\tau_o = T_1(\mu_{A_o^{(1)}}(x^*_1), \dots, \mu_{A_o^{(N)}}(x^*_N)), \quad (166)$$

gdzie: operatorem t-normy T_1 może być dowolny operator, brany pod uwagę przy tworzeniu bazy wiedzy (tab. III-1), przy czym nie jest wymagane, aby były to identyczne operatory. Stopień aktywacji przyjmuje wartości z przedziału 0-1. Jeżeli stopień aktywacji danej reguły elementarnej jest równy 0 to reguła ta nie zostaje zaktywizowana i nie bierze udziału w procesie wnioskowania. Natomiast, im stopień przesłanki jest wyższy, tym większy jest również udział danej reguły elementarnej w wyprowadzeniu rozmytych wniosków. Rysunek III-27 przedstawia schemat wyprowadzenia stopnia aktywacji dla przykładowej o -tej reguły plikowej o przesłance złożonej z uwzględnieniem różnych operatorów t-normy: t-normy Zadeha (minimum), t-normy Łukasiewicza, t-normy iloczynu algebraicznego oraz t-normy Hamachera.



Rys. III-27. Przykład wyprowadzenia stopnia aktywacji o -tej reguły plikowej dla wybranych operatorów t-normy

5.2. Wyprowadzenie rozmytego wniosku reguł elementarnych

Ostre wartości wejść systemu podlegają rozmywaniu typu singleton (tab. II-2). Wówczas, aby otrzymać **rozmyty wniosek reguły elementarnej o indeksie l/o** zamiast operacji (165), wystarczy uwzględnić zależność:

$$\mu_{B'_{l/o}}(y) = \mu_{R_{l/o}}(x^*_1, \dots, x^*_N, y). \quad (167)$$

Zatem, rozmyty wniosek reguły elementarnej wynosi:

- dla interpretacji logicznej reguł:

$$\mu_{B'_{l/o}}(y) = I\{T_1(\mu_{A_o^{(1)}}(x^*_1), \dots, \mu_{A_o^{(N)}}(x^*_N)), \mu_{B_{l/o}}(y)\} = I(\tau_o, \mu_{B_{l/o}}(y)), \quad (168)$$

- dla interpretacji reguł typu Mamdaniego:

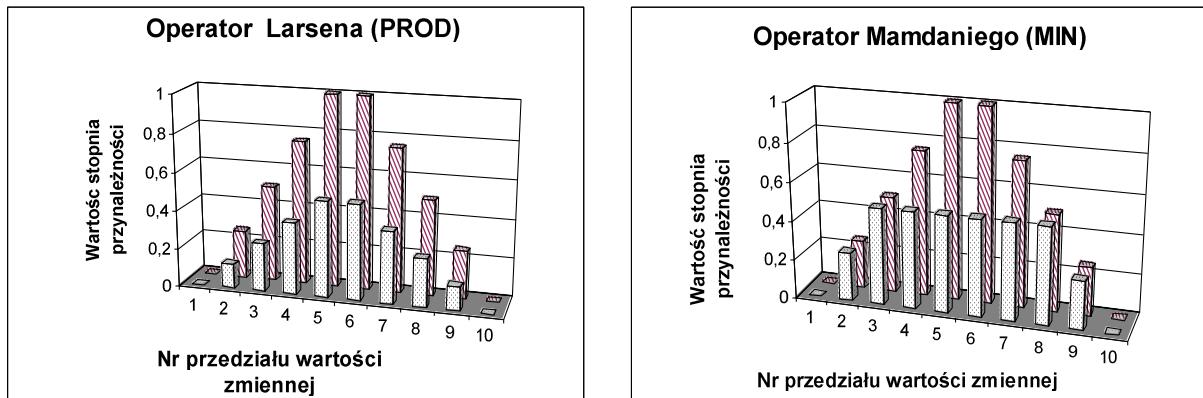
$$\mu_{B'_{l/o}}(y) = MIN\{T_1(\mu_{A_o^{(1)}}(x^*_1), \dots, \mu_{A_o^{(N)}}(x^*_N)), \mu_{B_{l/o}}(y)\} = MIN(\tau_o, \mu_{B_{l/o}}(y)), \quad (169)$$

- dla interpretacji reguł wg Larsena:

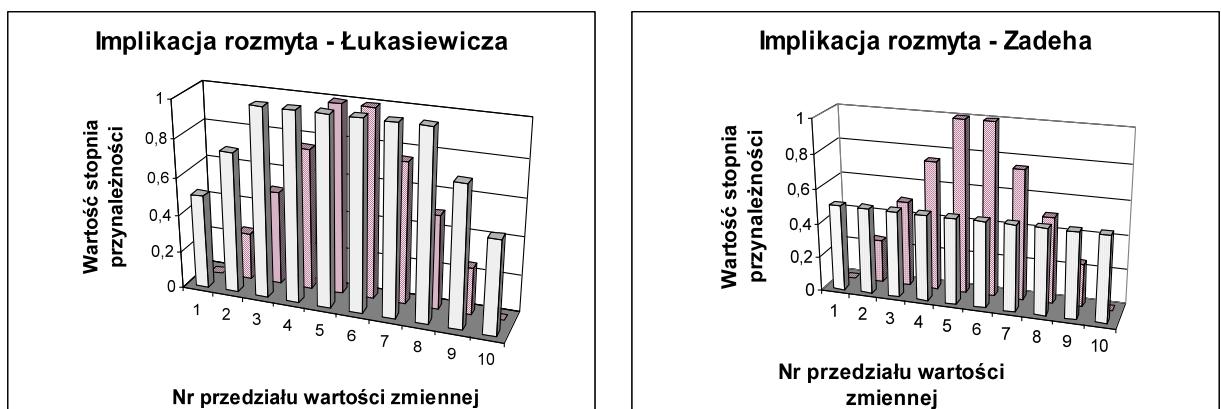
$$\mu_{B'_{l/o}}(y) = T_1(\mu_{A_o^{(1)}}(x^*_1), \dots, \mu_{A_o^{(N)}}(x^*_N)) \cdot \mu_{B_{l/o}}(y) = \tau_o \cdot \mu_{B_{l/o}}(y). \quad (170)$$

Dla interpretacji logicznej reguł, w tworzonym systemie dopuszcza się wykorzystanie różnych operatorów implikacji rozmytych tj.: Łukasiewicza, Goguena, Gödela, Kleene-Dienesa, Reichenbacha oraz Zadeha. Wzory wyżej wymienionych implikacji zamieszczono w tabeli II-5. Wybór odpowiednich operatorów implikacji (t-normy bądź implikacji rozmytej) ma zasadnicze znaczenie w procesie wnioskowania. Niezbędne staje się poznanie własności matematycznych tych operatorów przed dokonaniem ich właściwego wyboru. Na rysunkach III-28 ÷ III.31. przedstawione zostały różne wyniki rozmytego wniosku przykładowej reguły elementarnej o określonych stopniach przynależności zbioru $B_{l/o}$ w konkluzji oraz stopniu

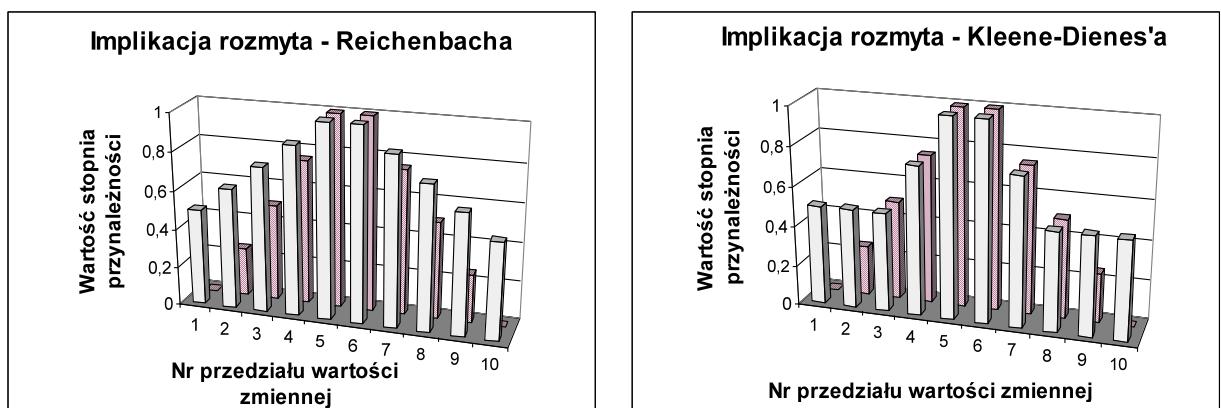
aktywacji o -tej reguły $\tau_o = 0.5$. Stopnie przynależności dla rozmytego zbioru uzyskanego w procesie wnioskowania reprezentują na rysunkach słupki o deseniu w kropki. Dla przejrzystości rysunków zmniejszono szerokości słupków – nie oznacza to zmniejszenia zakresu wartości zmiennych dla poszczególnych przedziałów.



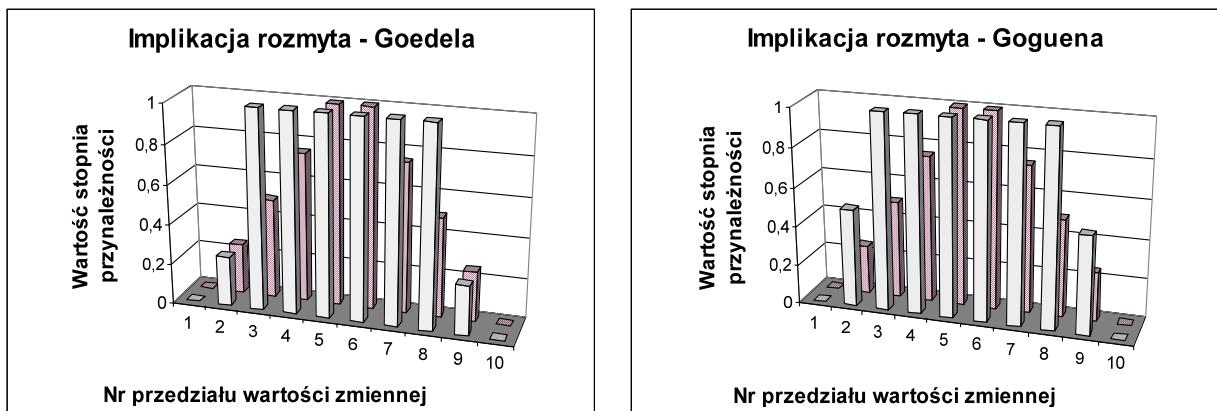
Rys. III-28. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla operatora: a) Larsena b) Mamdaniego



Rys. III-29. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla implikacji rozmytej: a) Łukasiewicza b) Zadeha



Rys. III-30. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla implikacji rozmytej: a) Reichenbacha b) Kleene-Dienes'a



Rys. III-31. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla implikacji rozmytej: a) Gödela b) Goguena

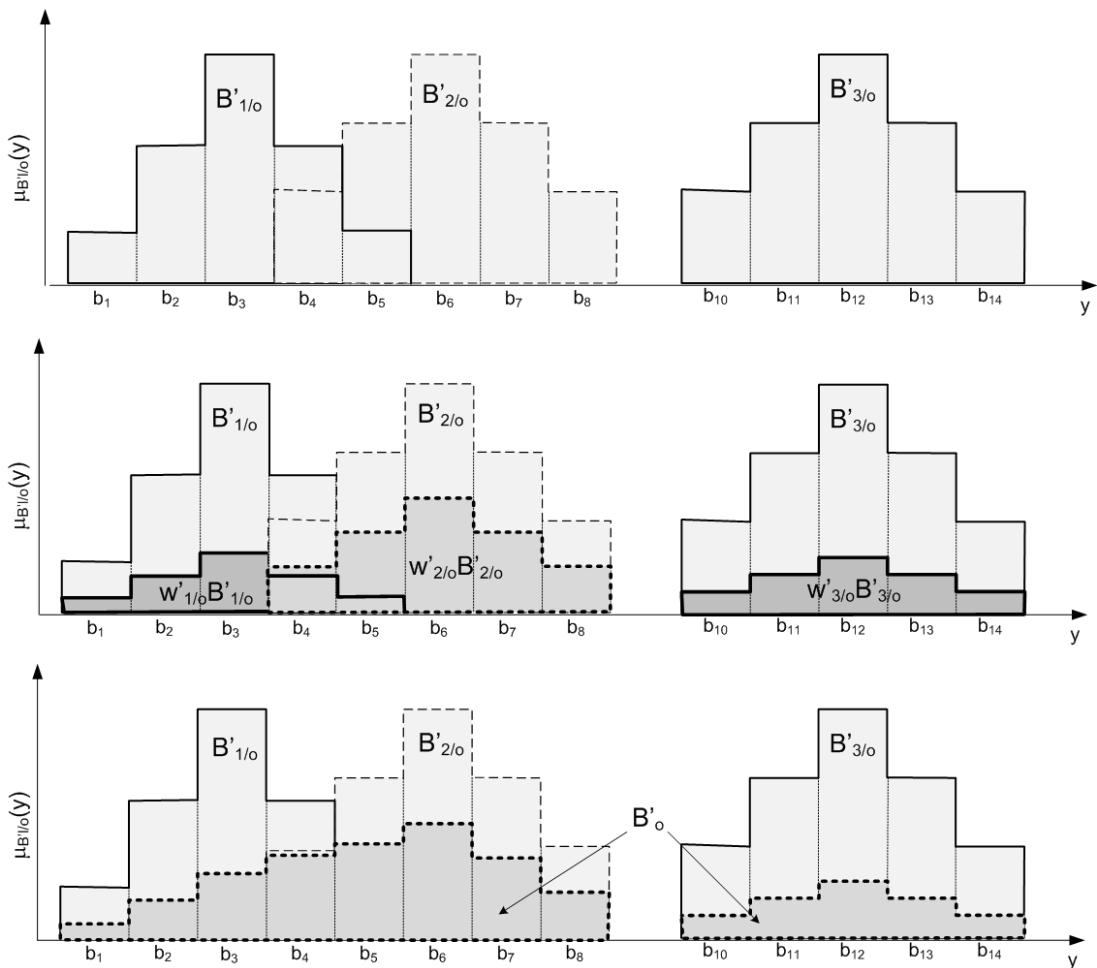
5.3. Wyprowadzenie rozmytego wniosku reguł plikowych

Wyprowadzając rozmyty wniosek dla o -tej reguły plikowej (160) należy uwzględnić wagę poszczególnych reguł elementarnych $w_{l/o}$ ($l=1, \dots, L$), stanowiące prawdopodobieństwa zajścia zdarzeń rozmytych (y jest $B_{l/o}$) w konkluzji pod warunkiem zajścia zdarzeń rozmytych ($\bigwedge_{n=1}^N x_n$ jest $A_o^{(n)}$) w przesłance reguły. **Rozmyty wniosek o -tej reguły plikowej** jest zatem obliczany za pomocą wzoru:

$$\mu_{B'_o}(y) = \sum_{l=1}^L w_{l/o} \mu_{B'_{l/o}}(y). \quad (171)$$

Rysunek III-32 ilustruje sposób wyprowadzenia rozmytego wniosku o -tej reguły plikowej B'_o , o następującej konkluzji:

$$\begin{aligned} TO \ y \ jest \ B_{1/o} \ z \ wagą \ w_{1/o} &= 0.25 \\ TAKŻE \ y \ jest \ B_{2/o} \ z \ wagą \ w_{2/o} &= 0.5 \\ TAKŻE \ y \ jest \ B_{3/o} \ z \ wagą \ w_{3/o} &= 0.25. \end{aligned} \quad (172)$$



Rys. III-32. Ilustracja wyprowadzenia rozmytego wniosku o -tej reguły plikowej

5.4. Wyprowadzenie rozmytego wniosku bazy wiedzy

Wyjście rozmyte modelu (wynikowy, rozmyty wniosek bazy wiedzy) wyznaczone w oparciu o wszystkie aktywne reguły plikowe bazy wiedzy, stanowi wartość oczekiwana rozmytego wyniku ([wb10]) i jest obliczane jako:

$$\mu_{B^o}(y) = \sum_{o=1}^O w'_o \mu_{B'_o}(y), \quad (173)$$

gdzie: waga w'_o wynosi

$$w'_o = \frac{w_o}{\sum_{o=1}^O w_o}. \quad (174)$$

Normalizacja wag w_o , jako prawdopodobieństw brzegowych zajścia zdarzeń w poprzedniku reguł plikowych, jest wymagana z uwagi na zastosowanie warunku minimalnego wsparcia w regułach podczas generowania bazy wiedzy z użyciem algorytmów opartych na regułach asocjacji.

5.5. Wyznaczenie wartości numerycznej na wyjściu systemu

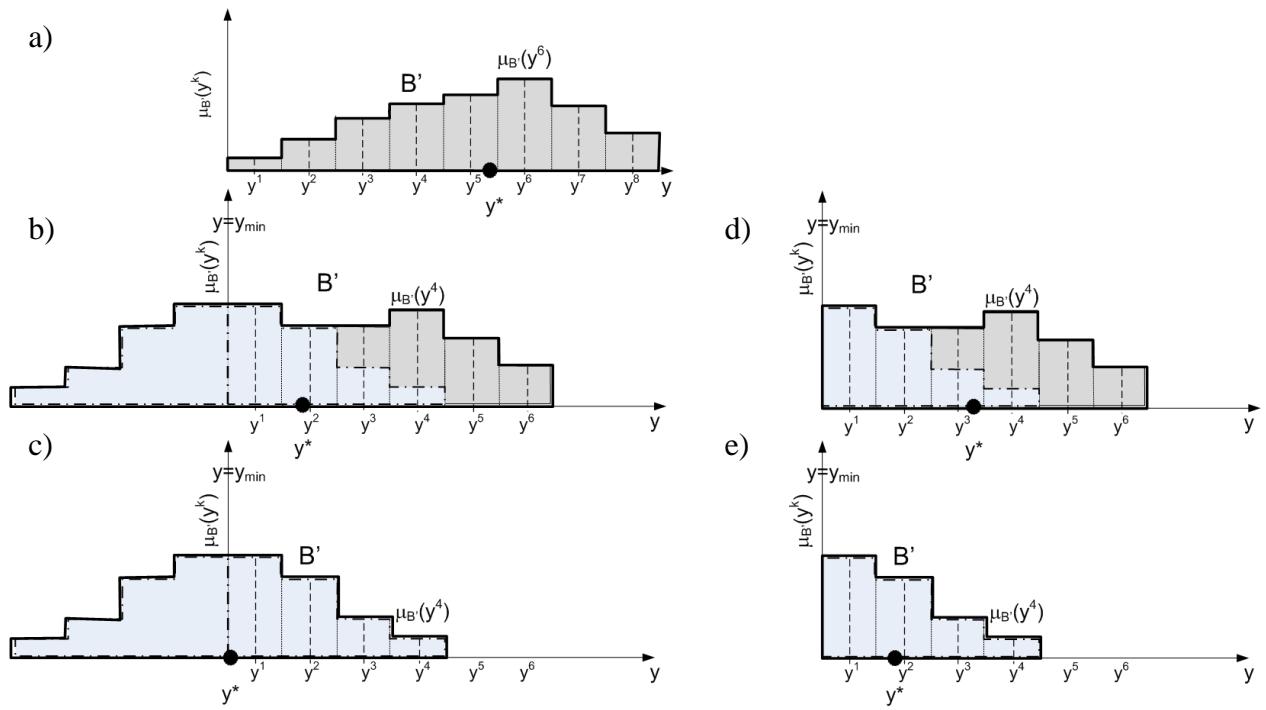
Numeryczna wartość na wyjściu y^* dla systemu wnioskującego typu MISO z probabilistyczno-rozmytą bazą wiedzy jest wyznaczana w oparciu o wybraną metodę wyostrzania na zbiorze rozmytym B' . Przegląd dostępnych w literaturze metod wyostrzania zamieszczono w podrozdziale II-3.5. Każda z metod posiada zalety, jak również i wady. Zatem wybór odpowiedniej metody opiera się na zasadzie wyboru mniejszego złego. W pozycjach [wb07] oraz [wb10] w celu realizacji wyostrzania proponuje się wykorzystanie **metody środka ciężkości COG** (zwanej COA) (tab. II-6). W przypadku definicji zbiorów rozmytych z wykorzystaniem stopni przynależności do rozłącznych przedziałów wartości zmiennych, metoda ta nie wiąże się z dużym nakładem obliczeniowym. Ponadto, pozwala na uwzględnienie wszystkich zbiorów rozmytych, dzięki czemu otrzymujemy płynne zmiany wartości wyjścia systemu (ciągłość sterowania). Wadą metody środka ciężkości jest brak reakcji na stopień aktywacji w przypadku zaktywowania tylko jednej reguły z jednym zbiorem rozmytym w konkluzji. Jednakże, taka sytuacja praktycznie nie ma miejsca w przypadku modelu z probabilistyczno-rozmytą bazą wiedzy. Kolejną wadą omawianej metody jest zawężenie zakresu wyostrzania [pie03], co oznacza, że nie można uzyskać minimalnych (maksymalnych) wartości zakresu zmiennej wejściowej mimo aktywacji brzegowych zbiorów rozmytych. Proponuje się zatem, w celu uzyskania numerycznych wartości na wyjściu omawianego systemu, wykorzystanie również tzw. **rozszerzonej metody środka ciężkości**. W rozszerzonej metodzie środka ciężkości wyjście systemu y^* oblicza się analogicznie do standardowej metody środka ciężkości:

$$y^* = \frac{\sum_{k=1}^K y^k \mu_{B'}(y^k)}{\sum_{k=1}^K \mu_{B'}(y^k)}, \quad (175)$$

gdzie: K jest liczbą rozłącznych przedziałów dyskretyzacji przestrzeni zmiennej wyjściowej, y^k stanowi wartość środkową k -tego przedziału:

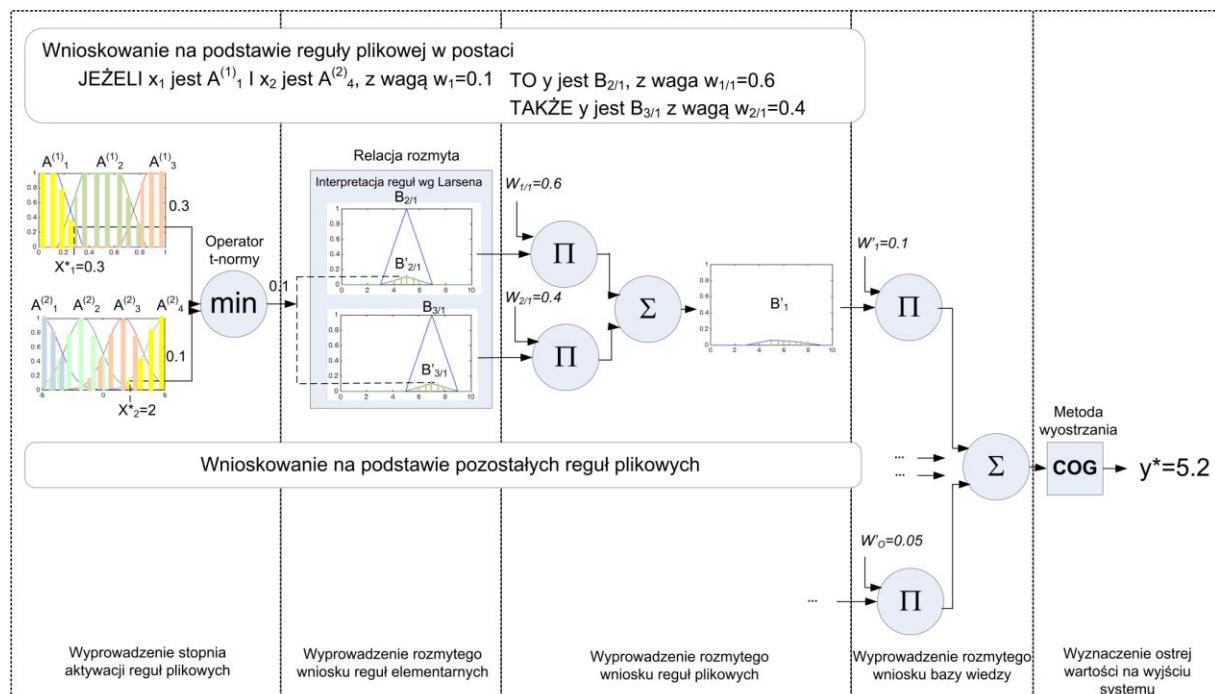
$$y^k = y_{\min} + \left(\frac{y_{\max} - y_{\min}}{K} \right) \cdot (k - 0.5) \quad (176)$$

a $\mu_{B'}(y^k)$ – wartość stopnia przynależności y^k do rozmytego wyniku bazy wiedzy B' . Wielkości y_{\min} i y_{\max} oznaczają zakres rozpiętości wartości zmiennej wyjściowej. Różnicę w metodzie stanowi zwiększenie zakresu brzegowych zbiorów rozmytych zmiennej wejściowej, aby móc uzyskać minimalne lub maksymalne wartości wyjścia. Na rysunku II-33 pokazano sposób wyznaczania wartości y^* rozszerzoną metodą środka ciężkości, a także wskazano na różnicę pomiędzy wynikami otrzymanymi rozszerzoną metodą środka ciężkości i standardową metodą środka ciężkości. Wybór metody wyostrzania jest uzależniony od struktury danych uczących i zdefiniowanej bazy danych. W przypadku brzegowych zbiorów rozmytych o dużym nośniku, metoda środka ciężkości powoduje wyznaczenie wartości ostrej, która znacznie wybiega spoza zakresu wyjściowych danych uczących.



Rys. III-33. Wyznaczanie wartości y^* za pomocą: a) metody środka ciężkości (rozszerzonej metody środka ciężkości) b)-c) rozszerzonej metody środka ciężkości d)-e) metody środka ciężkości

Graficzna interpretacja całości rozmytego wnioskowania w oparciu o probabilistyczno-rozmytą bazę wiedzy, odzwierciedlającą pełny rozkład prawdopodobieństwa zdarzeń rozmytych znajdujących się regułach ($w_o=w'_o$), została zamieszczona na rysunku III-34. W przykładzie uwzględniono interpretację reguł w postaci Larsena (170) oraz minimum, jako operatora t-normy łączącego proste przesłanki reguły rozmytej.



Rys. III-34. Graficzna interpretacja rozmytego wnioskowania w oparciu o probabilistyczno-rozmytą bazę wiedzy (na podstawie [rud10])

6. Implementacja systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy w środowisku Matlab

Do przeprowadzenia badań wykorzystano środowisko MATLAB w wersji 7.6 (R2008a). Wybór środowiska obliczeniowego podyktowany był faktem szerokiego rozpowszechnienia programu zarówno wśród naukowców (w zastosowaniach badawczych i dydaktycznych), jak i inżynierów (przy projektowaniu prac rozwojowych) [mat]. Daje to możliwość późniejszego wykorzystania tworzonego narzędzia przez różnych użytkowników w celu głębszej analizy i testowania omawianego systemu.

6.1. Założenia oraz wymaganie funkcjonalne implementacji modulu narzędziowego PFIS

Dając do sprawdzenia słuszności postawionej w dysertacji tezy utworzono w środowisku MATLAB moduł narzędziowy (ang. *toolbox*) o nazwie PFIS, który implementuje system wnioskujący z probabilistyczno-rozmytą bazą wiedzy, zgodnie z założeniami przedstawionymi w częściach III-1÷III-5 pracy. Implementacja modułu narzędziowego została wykonana z wykorzystaniem programowania strukturalnego w oparciu o pliki funkcyjne środowiska Matlab. Aspektem przemawiającym za wykorzystaniem w Matlabie programowania strukturalnego, zamiast programowania obiektowego, jest znaczne spowolnienie obliczeń wersji obiektowej, w stosunku do wersji operującej na wielowymiarowych macierzach [cze04]. Założeniem jest również integracja budowanego modułu z istniejącym w programie Matlab narzędziem *Fuzzy Logic Toolbox*. Pozwoli to na wykorzystanie gotowych rozwiązań modelowania w przestrzeni zbiorów rozmytych za pomocą standardowych funkcji przynależności. W szczególności, przed tworzonym modułem narzędziowym PFIS, stawia się następujące wymagania funkcjonalne:

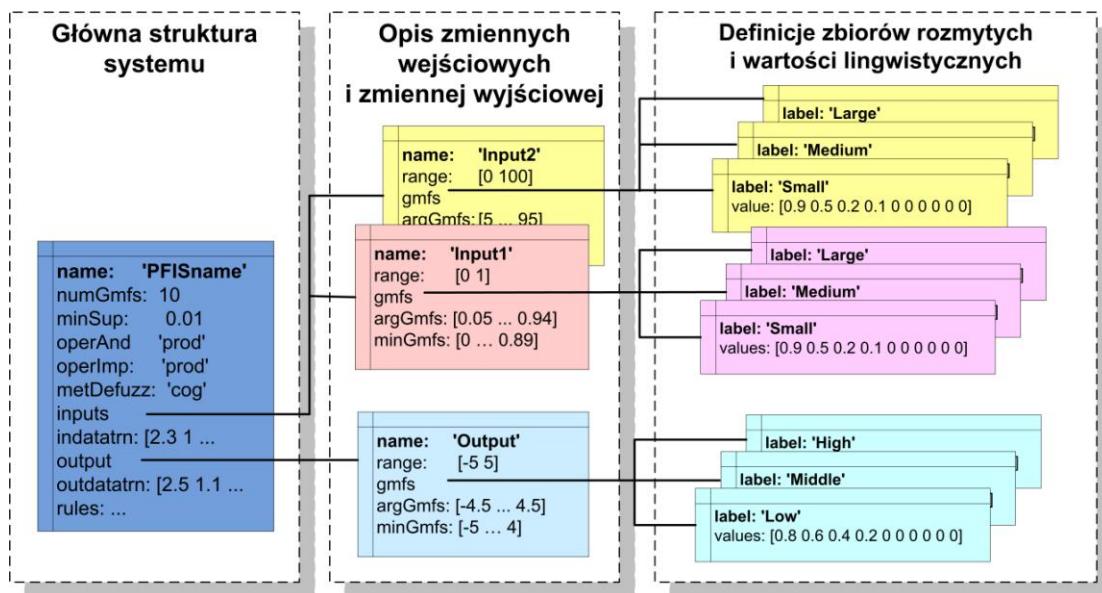
- możliwość budowy systemu typu MISO z probabilistyczno-rozmytą bazą wiedzy w oparciu o różne parametry jego struktury i parametry wnioskowania rozmytego (tj. liczba wejść systemu, liczba rozłącznych przedziałów dyskretyzacji przestrzeni zmiennych rozmytych wejścia/wyjścia, liczba wartości lingwistycznych dla zmiennych, stopnie przynależności dla wartości lingwistycznych, operator t-normy jako interpretacja spójnika logicznego AND w przesłance reguły, metoda interpretacji reguł, metoda wnioskowania),
- możliwość definiowania przez użytkownika własnych wartości lingwistycznych i zbiorów rozmytych zmiennych wejścia (wyjścia),
- możliwość generowania probabilistyczno-rozmytej bazy wiedzy (109) w oparciu o dane pomiarowe z wykorzystaniem określonego operatora t-normy oraz na poziomie zadanej wartości minimalnego prawdopodobieństwa łącznego występowania wartości rozmytych zmiennych wejściowych i zmiennej wyjściowej,
- wybór optymalnego (pod kątem minimalizacji czasu trwania obliczeń) algorytmu, w celu generowania probabilistyczno-rozmytej bazy reguł,
- możliwość przeglądu i analizy utworzonej bazy reguł wraz z jej wagami,
- możliwość wnioskowania dla danych uczących i testujących w oparciu o bazę wiedzy i ustalone parametry wnioskowania systemu,
- dostęp do wyżej wymienionych opcji poprzez przystępny i intuicyjny interfejs graficzny GUI.

6.2. Budowa omawianego systemu w kodzie programu Matlab

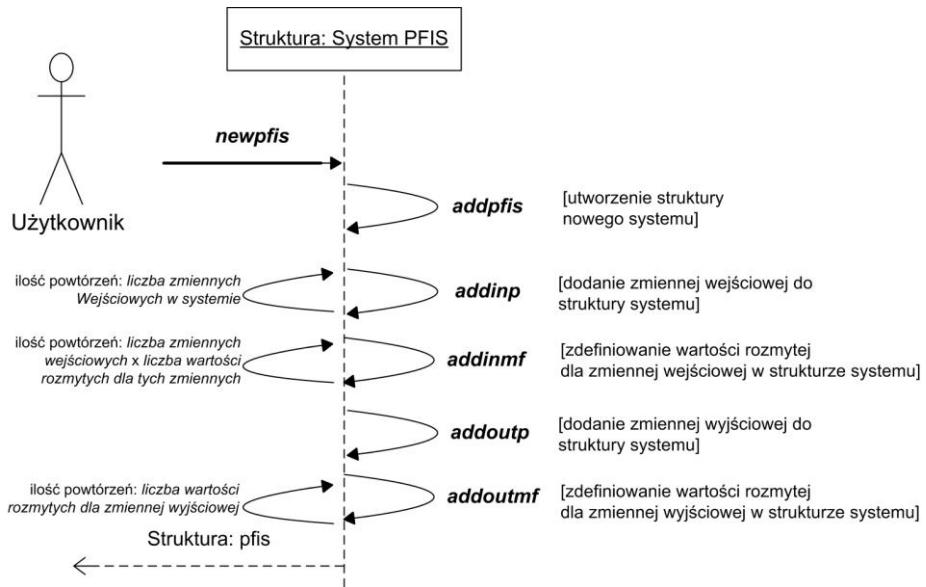
Moduł narzędziowy PFIS oparto o trzy główne funkcje, które pozwolą na utworzenie nowego systemu, wygenerowanie probabilistyczno-rozmytej bazy wiedzy z wykorzystaniem danych empirycznych oraz umożliwią wnioskowanie rozmyte w oparciu o utworzony model wiedzy.

6.2.1. Tworzenie nowego systemu jako obiektu struktury

Utworzenie nowego systemu jest możliwe dzięki wywołaniu funkcji *newpfis* (argumenty wywołania funkcji opisane zostały w dodatku B pracy). Generowany jest wówczas obiekt systemu o strukturze zamieszczonej na rysunku III-35. Struktura przechowuje w polach niezbędne informacje dotyczące zmiennych wejścia i wyjścia systemu, wartości lingwistycznych dla ww. zmiennych i utożsamianymi z nimi zbiorami rozmytymi, oraz informacje na temat interpretacji bazy reguł i metody wyostrzania. Na rysunku III-36 został przedstawiony uproszczony diagram sekwencji wywołania funkcji z modułu narzędziowego PFIS, które zostają wykorzystywane przy okazji tworzenia nowego systemu z probabilistyczno-rozmytą bazą wiedzy. W celach przejrzystości rysunku diagram nie zawiera opcji alternatywnych oraz argumentów wywołań funkcji.

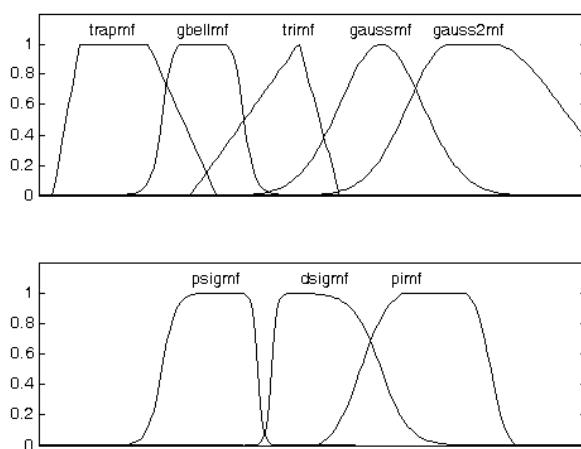


Rys. III-35. Przykład obiektu systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy jako struktury w Matlabie (na podstawie [bla08])



Rys. III-36. Diagram sekwencji wywołania funkcji z modułu narzędziowego PFIS, przy tworzeniu nowego systemu z probabilistyczno-rozmytą bazą wiedzy

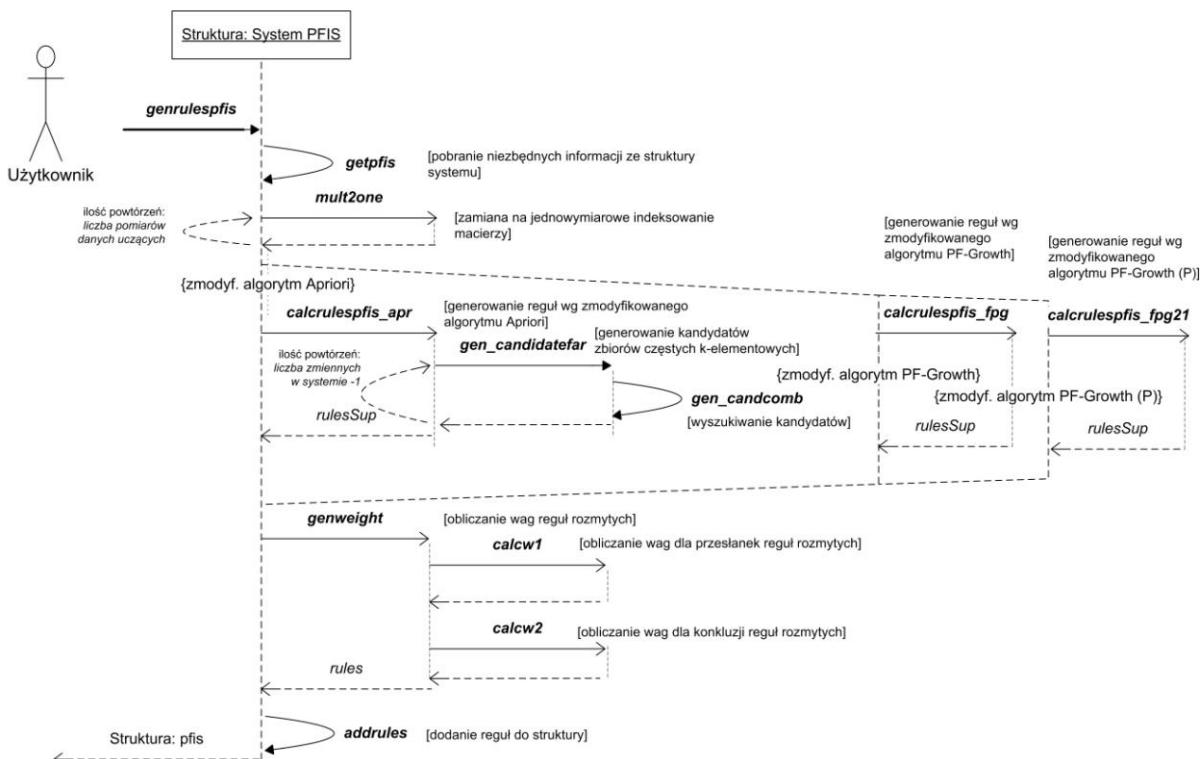
Prawidłowy dobór funkcji przynależności zależy od wiedzy i doświadczenia ekspertów. Metoda pozwala na zdefiniowanie własnych stopni przynależności dla stałych przedziałów wartości zmiennych lub daje możliwość definiowania je poprzez standardowe funkcje przynależności, jakie są dostępne w *Toolbox Fuzzy Logic* (rys. III-37) [fuz08]. Funkcja *newpfis* tworzy stałe stopnie przynależności (ozn. *gmfs*) w oparciu o jednakowe funkcje przynależności dla każdej zmiennej modelu. Chcąc dowolnie zróżnicować parametry można zastosować kolejno wywołanie funkcji: *addpfis*, *addinp*, *addinmf*, *addoutp*, *addoutmf* (por. rys. III-36). Doświadczenia naukowców wykazują, że korzystne jest zastosowanie najprostszych, wielokątnych funkcji przynależności, które ułatwiają proces strojenia modelu rozmytego i dają wysoką dokładność [pie03]. Sposób przekształcania funkcji przynależności w stałe stopnie przynależności dla rozłącznych przedziałów zmienności zmiennych został omówiony w rozdziale III-2.



Rys. III-37. Funkcje przynależności dostępne w *Toolbox Fuzzy Logic* (na podstawie [fuz08])

6.2.2. Generowanie probabilistyczno-rozmytej bazy reguł

Utworzony za pomocą funkcji *newpfis* system jest pozbawiony głównego składnika struktury – probabilistyczno-rozmytej bazy reguł. Funkcja *genrulespfis* pozwala na wygenerowanie reguł w postaci (109), na podstawie danych doświadczalnych oraz z użyciem iloczynu algebraicznego jako operatora t-normy (tab. III-1). W tym celu funkcja umożliwia wykorzystanie następujących algorytmów: zmodyfikowanego algorytmu Apriori (rozdział III-3.2.2) oraz zmodyfikowanego algorytmu FP-Growth lub FP-Growth (P) (rozdział III-3.2.3). Argumenty wywołania funkcji opisane zostały w dodatku B pracy. Na rysunku III-38 przedstawiony jest uproszczony diagram sekwencji wywołania funkcji podczas generowania probabilistyczno-rozmytej bazy wiedzy.



Rys. III-38. Diagram sekwencji wywołania funkcji z modułu narzędziowego PFIS, przy generowaniu probabilistyczno-rozmytej bazy wiedzy

6.2.3. Wnioskowanie w oparciu o utworzony model wiedzy

Mechanizm wnioskowania rozmytego modelu MISO oblicza na podstawie numerycznych wartości danych wejściowych, wynikowy zbiór rozmyty, a w konsekwencji numeryczną wartość wyjścia modelu. Dla systemu z bazą reguł w postaci (109), możliwe są dwa sposoby odnajdywania ostrych wyników y^* [wb07]. Jeden ze sposobów wnioskowania został omówiony w rozdziale III-5. Mechanizm ten jest wykorzystany przez funkcję *infermodpfis* do określania wartości wyjściowej systemu o strukturze jak na rysunku III-35. Argumenty wywołania funkcji opisane zostały w dodatku B pracy.

Poniższy przykład przedstawia budowę oraz analizę systemu z probabilistyczno-rozmytą bazą wiedzy o dwóch wejściach i jednym wyjściu, gdzie $x_1 \in [0,100]$, $x_2 \in [0,1]$, $y \in [-5,5]$. W tym celu wykorzystano funkcje modułu narzędziowego PFIS.

```
%utworzenie nowego modelu wiedzy(liczba rozłącznych przedziałów dyskretyzacji przestrzeni zmiennych: 40, typ f. przynależności: trapmf, liczba wartości lingwistycznych: 7, operator And: prod, interpretacja reguł: prod, metoda wyostrzania: cog, macierz zmiennych wejściowych: inX, macierz zmiennych wyjściowych: outX)
mod=newpfis('modelPFIS',inX,outX,40,['trapmf' 7],...
{'prod' 'prod' 'cog'},[0 0 -5;100 1 5]);
%generowanie probabilistyczno-rozmytej bazy wiedzy (t-norm: prod, wartość minimalnego wsparcia: 0.001)
mod=genrulespfis(mod,'apriori',inX,outX,'prod',0.001);
%wnioskowanie dla danych uczących w oparciu o utworzony model wiedzy
y=infermodpfis(mod, inX);
%określenie prawdopodobieństwa zdarzeń rozmytych na wyjściu przy zadanych wartościach numerycznych wejść
yprob=inferprobpfis(mod, [15.5 0.29]);
%wyświetlenie informacji o systemie wnioskującym
getpfis(mod,'all')
%wyświetlenie stopni przynależności dla zbiorów rozmytych
getpfis(mod,'plotgmfs',1)
getpfis(mod,'plotgmfs',2)
getpfis(mod,'plotgmfs',3)
%wyświetlenie probabilistyczno-rozmytej bazy wiedzy w postaci reguł plikowych
getpfis(mod,'rules')
%wyświetlenie probabilistyczno-rozmytej bazy wiedzy w postaci reguł elementarnych
getpfis(mod,'rules','elem')
```

Omówienie pozostałych funkcji moduły narzędziowego PFIS, utworzonych na potrzeby analiz prowadzonych w pracy doktorskiej, zostało zamieszone w dodatku B pracy.

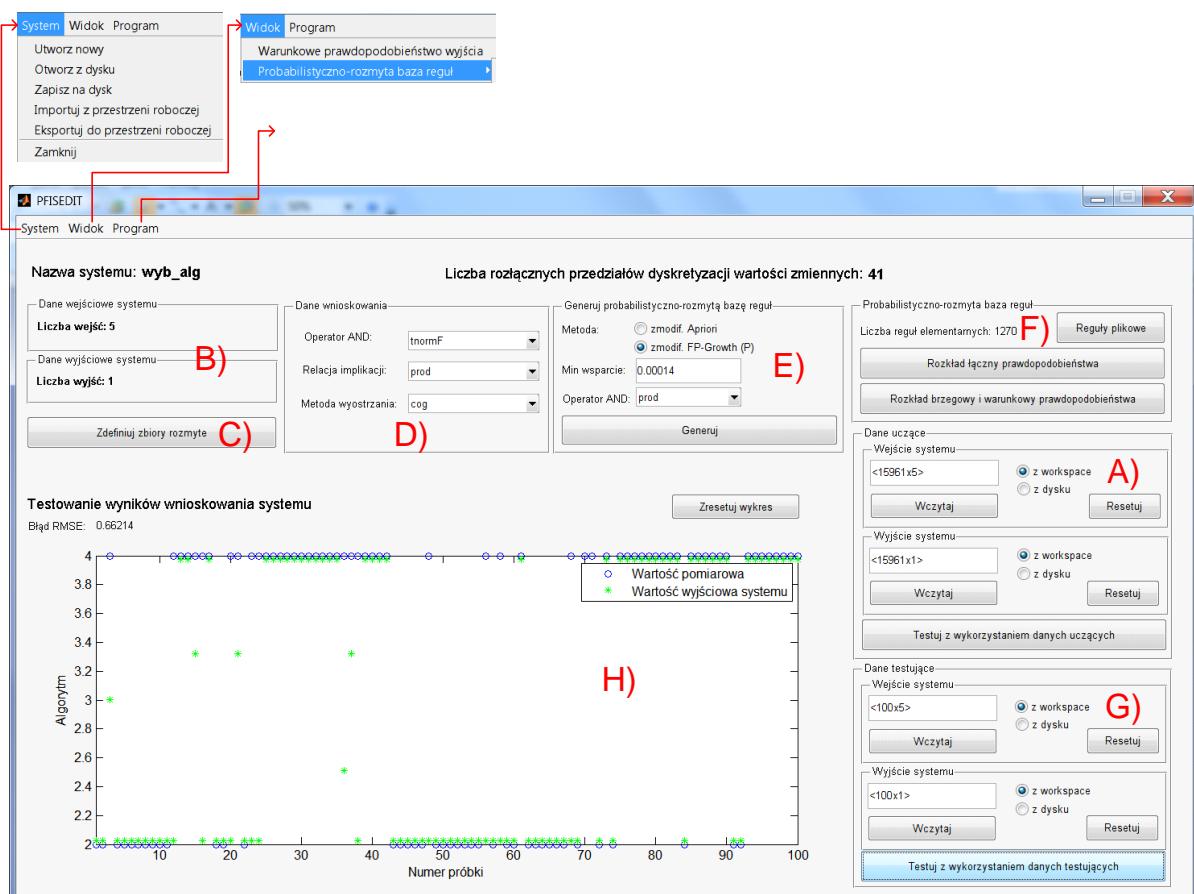
6.3. Graficzny interfejs użytkownika - PFISEDIT

Do modułu narzędziowego PFIS został także dołączony interfejs graficzny GUI o nazwie PFISEDIT, który umożliwia, osobom niewtajemniczonym w szczegóły implementacji, utworzenie systemu z probabilistyczno-rozmytą bazą wiedzy oraz jego testowanie z uwzględnieniem danych empirycznych. Główne okno programu przedstawia rysunek III-39, gdzie jako przykład został otworzony omawiany w rozdziale IV-2 rozmyty system decyzyjny.

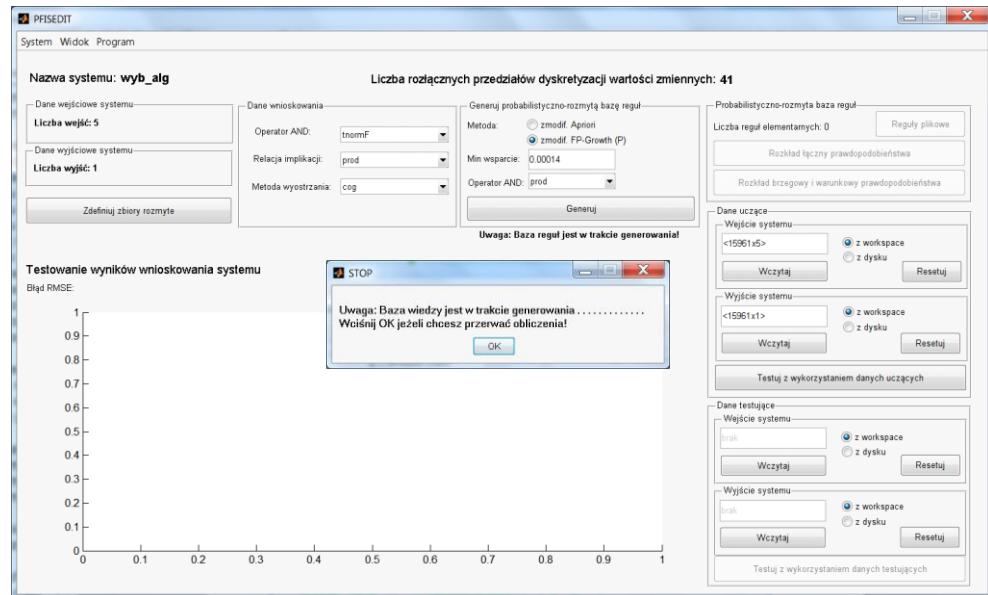
Narzędzie umożliwia utworzenie nowego systemu z probabilistyczno-rozmytą bazą wiedzy i zapisanie go na dysku komputera bądź w przestrzeni roboczej Matlaba. Pozwala również na otwarcie, wcześniej zapisanego systemu z dysku komputera bądź również z przestrzeni roboczej Matlaba.

Aby zbudować nowy system (menu programu: *System=>Utwórz nowy*) w pierwszej kolejności należy wczytać dane uczące (funkcja A). Po wczytaniu danych, pojawi się informacja o liczbie wejść i wyjść systemu (funkcja B) oraz uaktywnią się parametry

wnioskowania (funkcja D). Wówczas będzie można przystąpić do definicji zbiorów rozmytych (funkcja C). Gdy wartości lingwistyczne i utożsamiane z nimi zbiory rozmyte dla wszystkich zmiennych wejściowych i zmiennej wyjściowej zostaną zdefiniowane przez użytkownika, program pozwoli na generowanie probabilistyczno-rozmytej bazy reguł (opcja E). W tym celu wykorzystywany jest rozmyty system decyzyjny, omówiony w rozdziale IV-2 pracy. Na jego podstawie program sam wybiera efektywny algorytm do budowy bazy reguł. Na wybór algorytmu wpływają parametry ustalone wcześniej pośrednio lub bezpośrednio przez użytkownika takie jak: liczba zmiennych systemu, średnia liczba zdefiniowanych wartości rozmytych dla każdej zmiennej wejściowej i wyjściowej, liczba rozłącznych przedziałów dyskretyzacji wartości zmiennych, wartość minimalnego wsparcia reguł modelu, liczba rekordów danych uczących. W konsekwencji użycia systemu decyzyjnego, baza reguł jest generowana za pomocą zmodyfikowanego algorytmu Apriori lub zmodyfikowanego algorytmu FP-Growth (P). Z uwagi na długi czas trwania obliczeń dla większych wartości ww. parametrów, zniecierpliwiony użytkownik może w każdej chwili przerwać obliczenia w celu korekty parametrów modelu wiedzy (rys. III-40).



Rys. III-39. Interfejs PFISEDIT – okno główne interfejsu z zaznaczonymi zakresami funkcjonalnymi



Rys. III-40. Interfejs PFISEDIT – generowanie probabilistyczno-rozmytej bazy reguł

Interfejs PFISEDIT pozwala na szczegółową analizę informacji zawartych w utworzonej bazie wiedzy (funkcja F lub menu *Widok*). W szczególności umożliwia:

- przegląd reguł plikowych bazy wiedzy modelu w postaci tekstuowej (rys. III-41) (menu programu: *Widok=>Reguły plikowe* lub funkcja F: przycisk *Reguły plikowe*),
- podgląd, w postaci tabelarycznej, rozkładu łącznego prawdopodobieństwa zdarzeń rozmytych wejścia i wyjścia w regułach (rys. III-42) (menu programu: *Widok=>Rozkład łączny prawdopodobieństwa* lub funkcja F: przycisk *Rozkład łączny prawdopodobieństwa*),
- podgląd, w postaci tabelarycznej, wartości prawdopodobieństw brzegowych zdarzeń znajdujących się w przesłankach reguł oraz wykresów prawdopodobieństw warunkowych dla poszczególnych reguł plikowych bazy wiedzy (rys. III-43) (menu programu: *Widok=>Rozkład brzegowy i warunkowy prawdopodobieństwa* lub funkcja F: przycisk *Rozkład brzegowy i warunkowy prawdopodobieństwa*).

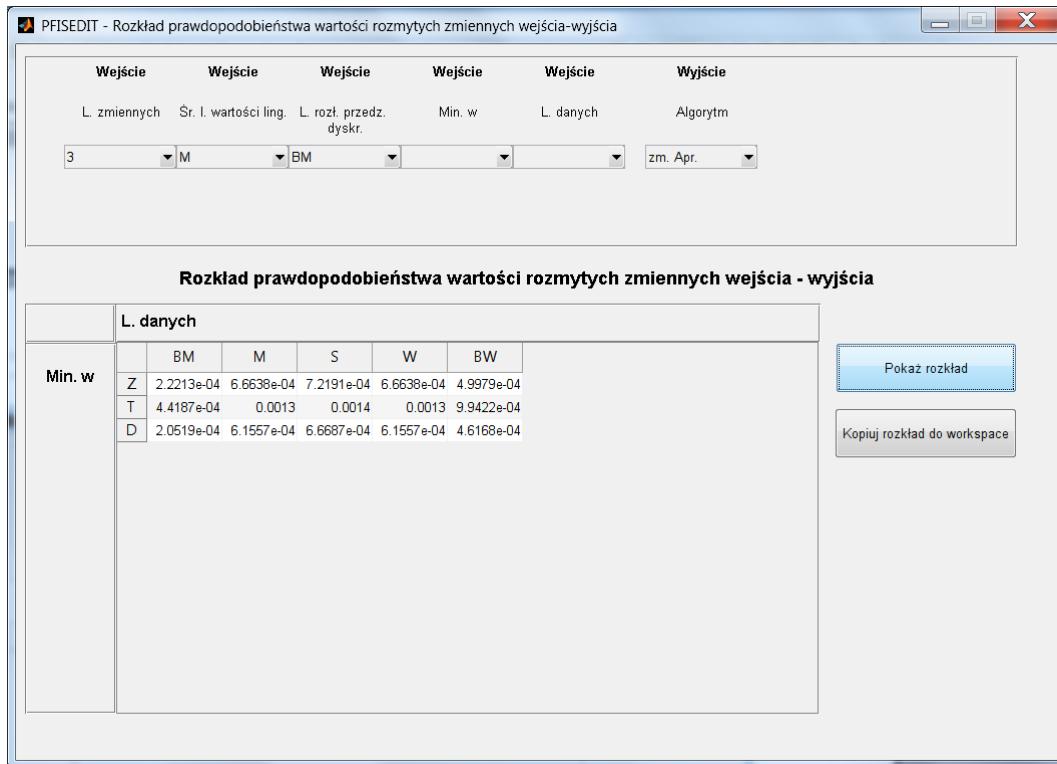
The screenshot shows a MATLAB editor window titled 'Editor - C:\MATLAB\R2008a\work\rules.txt'. The code area contains a series of numbered IF-THEN rules in a plain text format. Each rule consists of an IF part (with conditions like 'IS 2', 'AND', 'Min.', 'IS T') and a THEN part (containing 'Algorytm IS zm. Apr.' and a value like '[1.0000]').

```

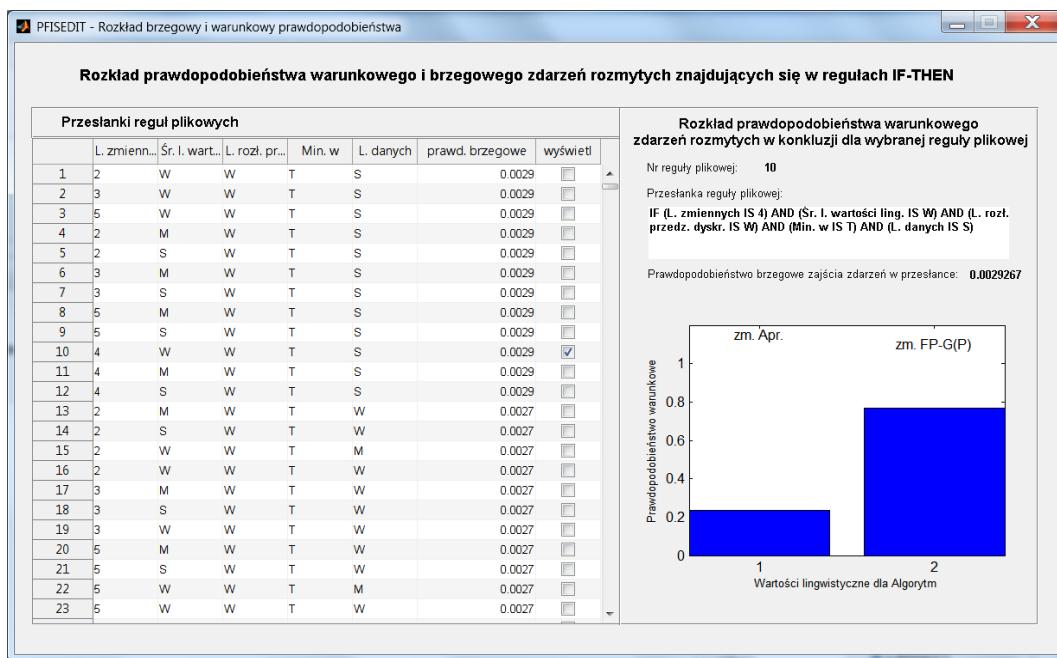
1: IF (L. zmiennych IS 2) AND (Sr. l. wartości ling. IS W) AND (L. rozl. przedz. dyskr. IS W) AND
2: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. Apr.) [1.0000]
3: IF (L. zmiennych IS 3) AND (Sr. l. wartości ling. IS W) AND (L. rozl. przedz. dyskr. IS W) AND
4: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. Apr.) [1.0000]
5: IF (L. zmiennych IS 5) AND (Sr. l. wartości ling. IS W) AND (L. rozl. przedz. dyskr. IS W) AND
6: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. FP-G(P)) [1.0000]
7: IF (L. zmiennych IS 2) AND (Sr. l. wartości ling. IS M) AND (L. rozl. przedz. dyskr. IS W) AND
8: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. Apr.) [1.0000]
9: IF (L. zmiennych IS 2) AND (Sr. l. wartości ling. IS S) AND (L. rozl. przedz. dyskr. IS W) AND
10: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. Apr.) [1.0000]
11: IF (L. zmiennych IS 3) AND (Sr. l. wartości ling. IS M) AND (L. rozl. przedz. dyskr. IS W) AND
12: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. Apr.) [1.0000]
13: IF (L. zmiennych IS 3) AND (Sr. l. wartości ling. IS S) AND (L. rozl. przedz. dyskr. IS W) AND
14: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. Apr.) [1.0000]
15: IF (L. zmiennych IS 5) AND (Sr. l. wartości ling. IS M) AND (L. rozl. przedz. dyskr. IS W) AND
16: |(Min. w IS T) AND (L. danych IS S) [0.0029] THEN (Algorytm IS zm. FP-G(P)) [1.0000]
17: IF (L. zmiennych IS 5) AND (Sr. l. wartości ling. IS S) AND (L. rozl. przedz. dyskr. IS W) AND (Min.

```

Rys. III-41. Przegląd reguł plikowych probabilistyczno-rozmytej bazy wiedzy



Rys. III-42. Interfejs PFISEDIT – rozkład prawdopodobieństwa wartości rozmytych zmiennych wejścia-wyjścia



Rys. III-43. Interfejs PFISEDIT – rozkład prawdopodobieństwa warunkowego i brzegowego zdarzeń rozmytych znajdujących się w regułach probabilistyczno-rozmytej bazy wiedzy

Interfejs PFISEDIT pozwala również na wnioskowanie rozmyte w oparciu o ustawione przez użytkownika parametry (funkcja D), z zastosowaniem metod omówionych w rozdziale III-5. Nowe fakty, na podstawie których system wnioskuje, mogą stanowić dane

uczące (funkcja A) lub wczytane dane testujące (funkcja G). Wynik wnioskowania zostaje przedstawiony w postaci:

- wykresu porównania wartości rzeczywistych zmiennej wyjściowej oraz wartości uzyskanych na podstawie wnioskowania dla kolejnych rekordów danych (funkcja H),
- pierwiastka błędów średniokwadratowego RMSE, wyznaczonego ze wzoru:

$$RMSE = \sqrt{\frac{\sum_{n=1}^N [y(n) - \hat{y}(n)]^2}{N}}, \quad (177)$$

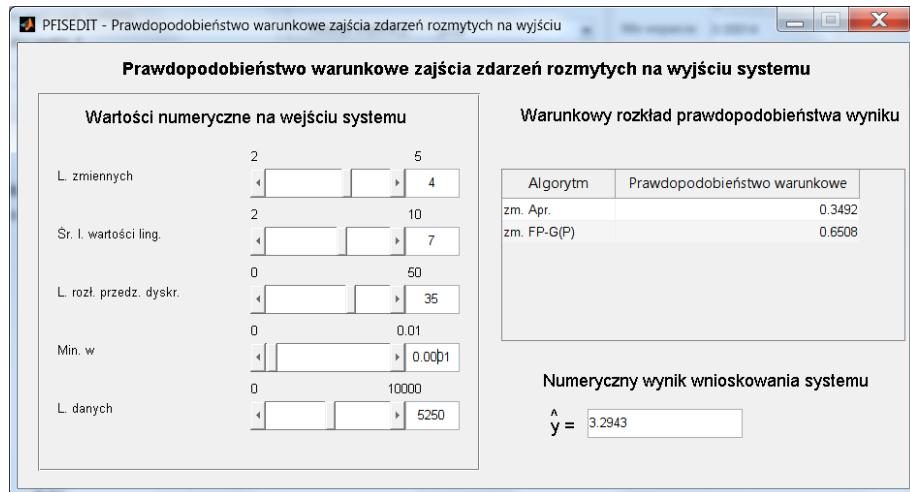
gdzie:

N – liczliwość zbioru,

$\hat{y}(n)$ – wartość wyjściowa dla rekordu n danych, wyznaczana w oparciu o utworzony model wiedzy,

$y(n)$ – rzeczywista wartość wyjściowa dla rekordu n danych.

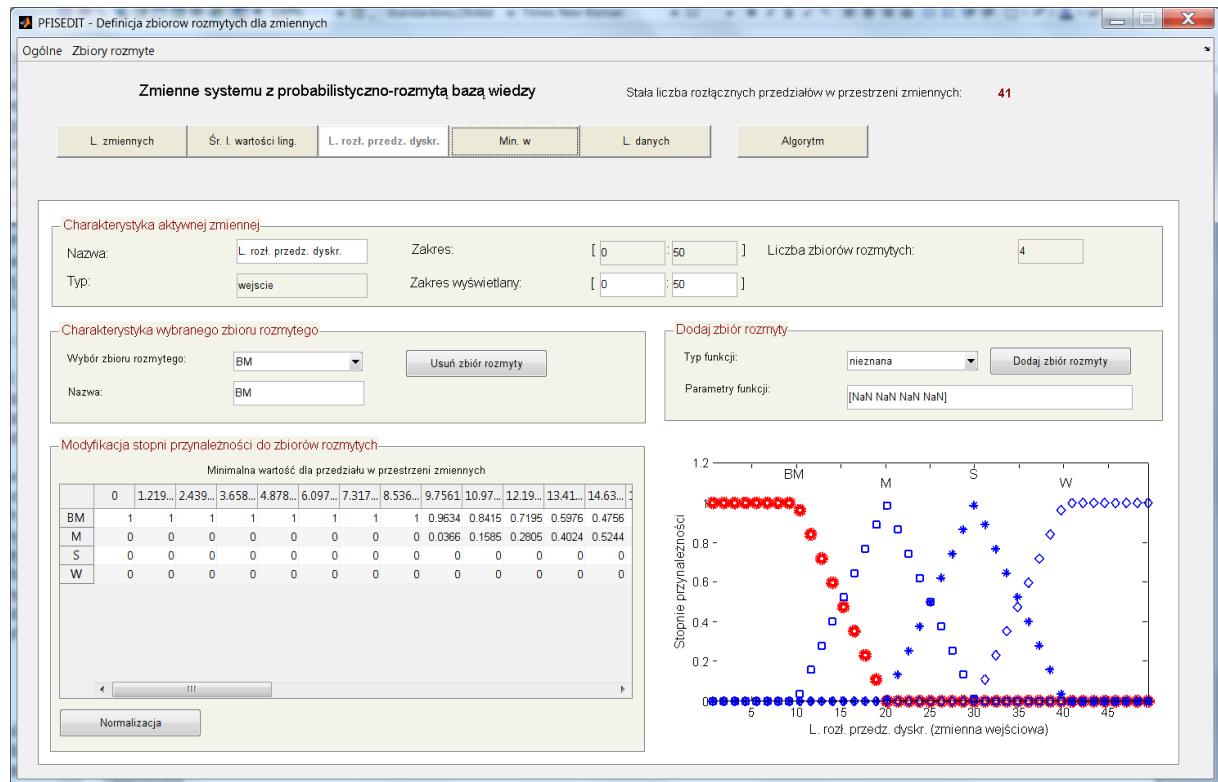
Mając na uwadze zastosowanie systemu wnioskującego do zagadnień podejmowania decyzji lub diagnostyki, narzędzie umożliwia również określenie prawdopodobieństw warunkowych zajścia zdarzeń rozmytych na wyjściu, przy faktach określonych za pomocą wartości numerycznych (menu programu: *Widok=>Warunkowe prawdopodobieństwo wyjścia*) (rys. III-44). W praktyce oznacza to, że otrzymujemy nie tylko wskazania odnośnie rozwiązania danego problemu (np. diagnozy), ale także określone jest ryzyko danego rozwiązania w ujęciu probabilistycznym.



Rys. III-44. Interfejs PFISEDIT – warunkowy rozkład prawdopodobieństwa zajścia zdarzeń rozmytych na wyjściu systemu

Rysunek III-45 przedstawia interfejs PFISEDIT, który umożliwia określenie charakterystyki zmiennych modelu, a w szczególności ich wartości lingwistycznych (zbiorów rozmytych). Program pozwala użytkownikowi na dodanie pojedynczego zbioru rozmytego (przycisk: *Dodaj zbiór rozmyty*) lub wielu zbiorów jednocześnie (menu programu: *Zbiory rozmyte=> Dodaj wiele zbiorów*), zdefiniowanych jako stopnie przynależności dla przedziałów wartości zmiennych, wynikające z przeliczeń na podstawie określonej przez

użytkownika funkcji przynależności. Istnieje możliwość dowolnej modyfikacji przez użytkownika stopni przynależności przedziałów wartości do zbioru rozmytego. Opcja normalizacji umożliwia wówczas dopasowanie stopni przynależności zbiorów rozmytych, aby spełniały warunek dopełnienia do jedności (107).

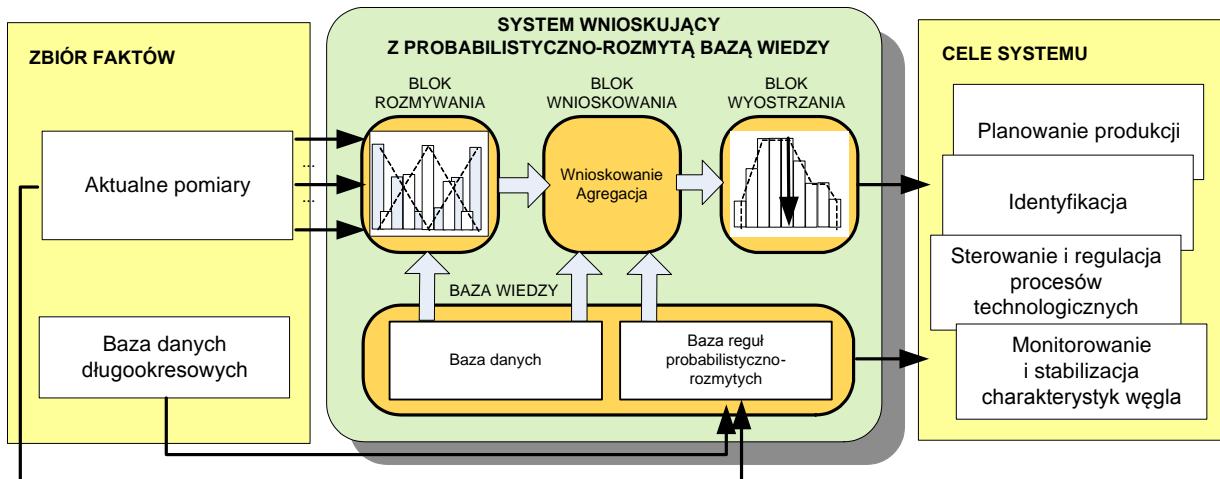


Rys. III-45. Interfejs PFISEDIT – definiowanie zbiorów rozmytych dla zmiennych systemu

IV. Zastosowania systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy

1. Modelowanie własności węgla

W niniejszej dysertacji proponuje się zastosowanie systemu z probabilistyczno-rozmytą bazą wiedzy do modelowania własności węgla. Idea utworzenia bazy wiedzy dla systemów informatycznych wspomagających procesy planowania, sterowania i identyfikację w zakładach procesów przeróbki węgla została zarysowana w projekcie badawczym [wb97]. Zakłady przeróbki węgla posiadają strukturę złożoną z wielu pojedynczych procesów (rozdrabnianie i rozdziały materiału, wzmacnianie składnika użytecznego, mieszanie materiałów o różnej zawartości składnika użytecznego, uśrednianie materiału ze względu na dany pomiar, układy regulacji dla osadzarek, filtrów i wiele innych), łączących się w ciągi technologiczne o określonych zadaniach [wb97]. Mając na uwadze cele strategiczne zakładu przeróbki surowców mineralnych, jak i bieżące sterowanie poszczególnymi procesami, można określić zarys możliwych zastosowań utworzonego systemu wnioskującego oraz jego bazy wiedzy (rys. IV-1).



Rys. IV-1. Zarys możliwych zastosowań systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy (na podstawie [wb97])

Zmienność parametrów węgla dla optymalnego sterowania w zakładzie mechanicznej przeróbki węgla była przedmiotem analiz w wielu pracach. W pracy [gor05] autorka przewiduje, m.in. ilość, jakość oraz wartość możliwych produktów handlowych uzyskanych z węgla surowego, posługując się tzw. formułą sprzedaży [bla02]. Natomiast prace [wb08b], [wb08c], [bla08] poświęcone zostały metodyce utworzenia rozmytej bazy wiedzy zmienności cech badanego paliwa z zastosowaniem reguł asocjacji.

1.1. Wielkości charakteryzujące węgiel jako materiał uziarniony

Węgiel, jako paliwo kopalne, stanowi materiał uziarniony, niejednorodny. Pomiar parametrów węgla – materiału surowego, nie jest sprawą trywialną. Dodatkowym czynnikiem utrudniającym przewidywalność pomiarów jest czynnik losowy wpływający na wybór próbki poddawanej pomiarowi. Jest to czynnik generujący błędy zarówno w przypadku pomiarów dyskretnych, jak i w pomiarach ciągłych [wb04].

Podstawowymi charakterystykami materiału uziarnionego, którymi posługują się automatycy i technologodzy przeróbki, są [wb97]:

- charakterystyka granulometryczna, podająca udziały klas ziarnowych o danej średnicy,
- charakterystyka densymetryczna, podająca frakcji ziarn o danej gęstości,
- charakterystyka jakości węgla.

Do podstawowych parametrów charakteryzujących jakość węgla należą (na podstawie [cie05], [wb08b]):

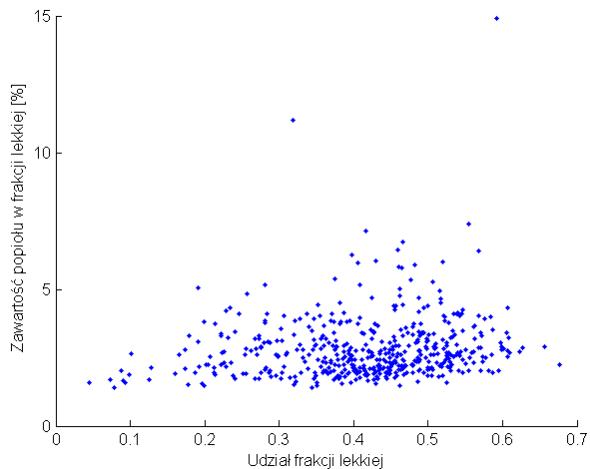
- wartość opałowa (MJ/kg) decydująca o ilości możliwej energii uzyskanej z jednego kilograma masy węgla,
- zawartość wilgoci (%),
- zawartość popiołu (%), wyrażająca udział masy substancji niepalnej w całkowitej masie węgla,
- zawartość siarki (%) i inne.

Pierwsze dwa parametry decydują o przydatności węgla jako paliwa dla elektrowni, dwa kolejne parametry wskazują na zagrożenie obciążenia środowiska naturalnego przez odpady powstające przy produkcji energii elektrycznej (lub cieplnej) [wb08b]. Powyższe wielkości są zatem najważniejszymi elementami kontraktów handlowych zawieranych pomiędzy kopalnią a odbiorcą produktu (elektrownią). Stąd też powinny być pod stałą kontrolą podczas wykorzystywania ich w układach automatycznej stabilizacji czy też regulacji procesów technologicznych w zakładach wzbogacania węgla [cie05].

Analizie zostanie poddana zależność udziałów masowych frakcji gęstościowych węgla oraz zawartości popiołu w tych frakcjach. Zawartość popiołu jest to wyrażona w procentach masa stałej pozostałości po wypróżnieniu próbki węgla o masie do $2 \cdot 10^{-3}$ kg i uziarnieniu mniejszym niż $2 \cdot 10^{-4}$ m w warunkach ustalonych z normą (PN-80/G-04512). Węgiel jest tym lepszy jakościowo, im posiada mniejszą zawartość popiołu. Jednakże wyniki metod pomiarów zawartości popiołu, które są wykonywane w warunkach laboratoryjnych, uzyskuje się z dużym opóźnieniem. Zatem, w celu sterowania procesem technologicznym stosowane są metody korelacji gęstości węgla i zawartości popiołu w próbce węgla, bądź też metody oddziaływanego promieniowania jądrowego z węglem [cie05].

1.2. Reprezentacja probabilistyczno-rozmyta dwóch zmiennych charakteryzujących węgiel dla danych surowych

W celu przejrzystości analiz dotyczących identyfikacji probabilistycznych własności zmiennych rozmytych rozpatruje się zależności tylko dwóch zmiennych. Badaniu podlega masowy udział gęstościowej frakcji lekkiej węgla (w skrócie nazywany udziałem frakcji lekkiej węgla), w której gęstość węgla jest mniejsza od $1,3 \cdot 10^3$ kg/m³ – stanowiący wejście systemu z probabilistyczno-rozmytą bazą wiedzy, oraz zawartość popiołu w tej frakcji, stanowiącej wyjście systemu. Zależność danych empirycznych wyżej wymienionych parametrów została przedstawiona na rysunku IV-2.



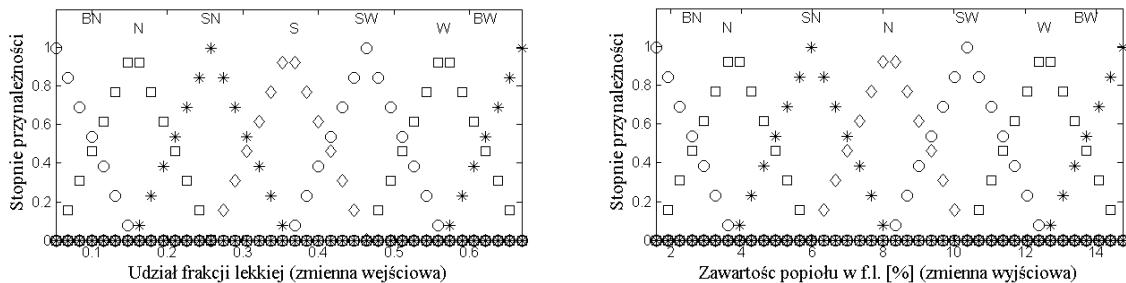
Rys. IV-2. Zależność zawartości popiołu w frakcji lekkiej [%] od masowego udziału tej frakcji

W badaniach empirycznych, przestrzeń rozważań zmiennych została podzielona na 40 rozłącznych przedziałów o równej długości. Przyjęto po 7 następujących wartości lingwistycznych zmiennych:

$$L(Udział\ frakcji\ lekkiej) = \{'Bardzo\ Niski', 'Niski', 'Średnio\ Niski', 'Średni', 'Średnio\ Wysoki', 'Wysoki', 'Bardzo\ Wysoki'\} = \{'BN', 'N', 'SN', 'S', 'SW', 'W', 'BW'\},$$

$$L(Zawartość\ popiołu) = \{'Bardzo\ Niska', 'Niska', 'Średnio\ Niska', 'Średnia', 'Średnio\ Wysoka', 'Wysoka', 'Bardzo\ Wysoka'\} = \{'BN', 'N', 'SN', 'S', 'SW', 'W', 'BW'\}.$$

Jednocześnie, dla każdej wartości lingwistycznej został określony zbiór rozmyty (rys. IV-3).



Rys. IV-3. Definicja zbiorów rozmytych dla wartości lingwistycznych zmiennych: udziału frakcji lekkiej i zawartości popiołu w tej frakcji

Do budowy modelu wiedzy, w celu obliczenia wartości łącznego prawdopodobieństwa zdarzeń rozmytych zmiennej wejściowej i wyjściowej, użyto iloczynu algebraicznego jako operatora t-normy. W przypadku pełnego rozkładu prawdopodobieństwa zdarzeń rozmytych, wyniki rozkładów są jednakowe bez względu na to, który algorytm użyjemy do generowania reguł (rozdział III-3).

Otrzymany model wiedzy składa się z 7 reguł plikowych (30 reguł elementarnych). Na podstawie wag modelu możemy otrzymać łączny rozkład prawdopodobieństwa zmiennych wejścia-wyjścia. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych: udziału frakcji lekkiej węgla oraz zawartości popiołu w tej frakcji został przedstawiony w tabeli IV-1.

Można zaobserwować najczęstsze współwystępowanie średnio wysokich wartości udziału frakcji lekkiej oraz bardzo niskich (18,73% przypadków) i niskich (15,22% przypadków)

wartości popiołu w tej frakcji. Nieco mniej częste jest współwystępowanie średnich wartości udziału frakcji lekkiej oraz bardzo niskich (14,03% przypadków) i niskich (11,08% przypadków) wartości popiołu w tej frakcji. Ponadto, dla każdej wartości udziału frakcji lekkiej, największe jest prawdopodobieństwo wystąpienia bardzo niskiej i niskiej zawartości popiołu w tej frakcji, co jest efektem pożdanym z punktu widzenia przydatności węgla, jako paliwa.

Tab. IV-1. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych wejścia-wyjścia

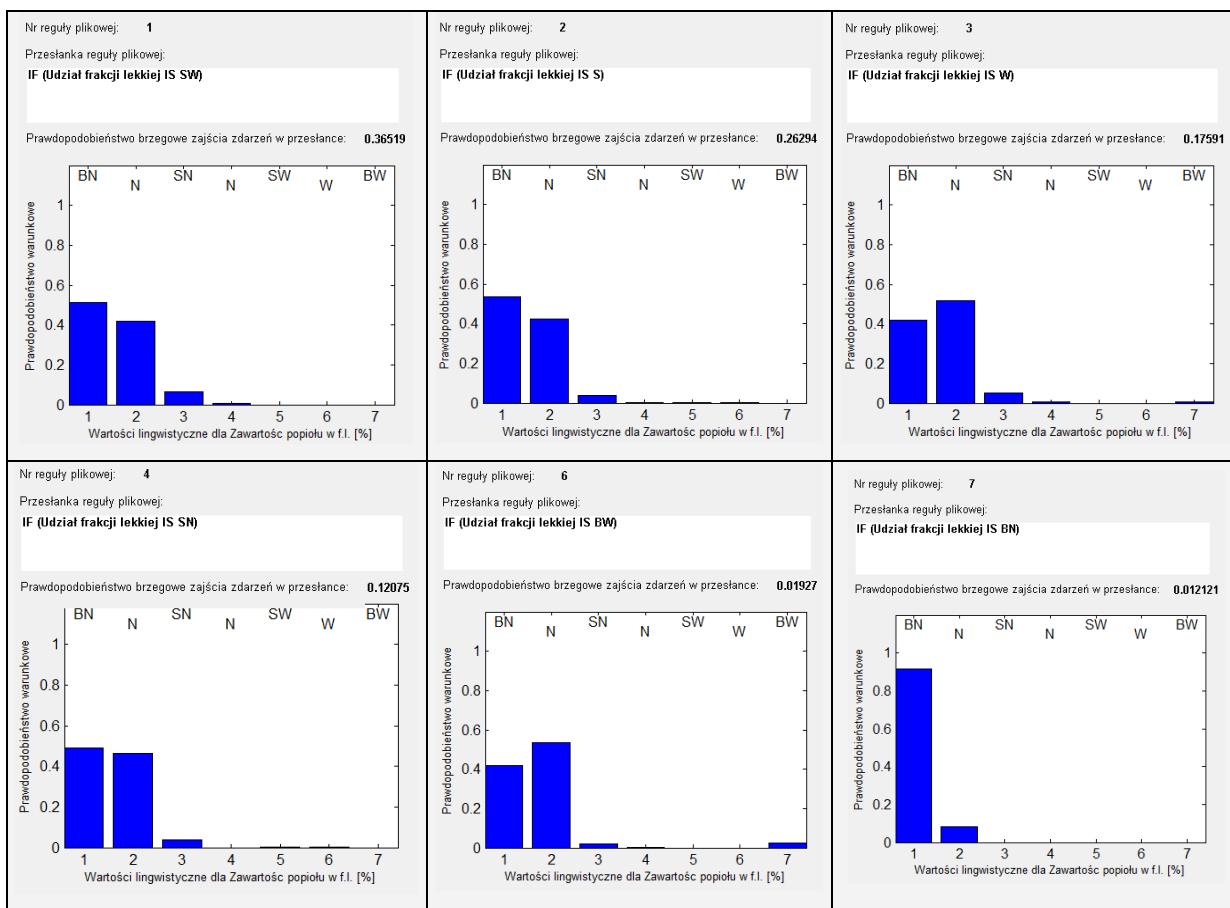
		Zawartość popiołu w frakcji lekkiej [%]						
		1 'BN'	2 'N'	3 'SN'	4 'S'	5 'SW'	6 'W'	7 'BW'
		Udział frakcji lekkiej węgla						
1 'BN'	1 'BN'	0,0110932	0,0010280	0	0	0	0	0
	2 'N'	0,0263941	0,0162094	0,0012193	0	0	0	0
	3 'SN'	0,0594465	0,0561353	0,0043871	0	0,0005379	0,0002391	0
	4 'S'	0,1402905	0,1107884	0,0099934	0,0006216	0,0008607	0,0003825	0
	5 'SW'	0,1872452	0,1521726	0,0241229	0,0016496	0	0	0
	6 'W'	0,0733011	0,0906102	0,0090132	0,0014345	0	0	0,0015540
	7 'BW'	0,0080569	0,0103520	0,0003706	2,39E-05	0	0	0,0004662

Tabela IV-2 przedstawia rozkład prawdopodobieństwa brzegowego wartości lingwistycznych udziału frakcji lekkiej. Wartości te stanowią bezpośrednio wagę modelu wiedzy. Można zauważyc, że dla udziału frakcji lekkiej, najczęstsze jest występowanie wartości średnio wysokich (36,52% przypadków), następnie średnich (26,29% przypadków), wysokich (17,59% przypadków) i średnio niskich (12,07% przypadków).

Tab. IV-2. Rozkład prawdopodobieństwa brzegowego wartości lingwistycznych przesłanki (udziału frakcji lekkiej)

Lp.	Wartość lingwistyczna	Prawdopodobieństwo brzegowe
1	'SW'	0,3652
2	'S'	0,2629
3	'W'	0,1759
4	'SN'	0,1207
5	'N'	0,0438
6	'BW'	0,0193
7	'BN'	0,0121

Rysunek IV-3 przedstawia rozkład prawdopodobieństwa warunkowego wartości lingwistycznych zawartości popiołu w frakcji lekkiej dla wybranych reguł plikowych. W przypadku najważniejszej reguły, której przesłanka mówi, że jeżeli udział frakcji lekkiej jest średnio wysoki to najbardziej prawdopodobnym zdarzeniem będzie otrzymanie wartości bardzo niskiej (51,27% przypadków) i niskiej (41,67% przypadków) zawartości węgla w analizowanej frakcji. Inaczej rozkładają się proporcje dla najmniej ważnej reguły, jeżeli udział frakcji lekkiej będzie bardzo niski to aż 91,52% przypadków wskazuje, że zawartość popiołu będzie bardzo wysoka. Nie należy się tym jednak zbyt bardzo sugerować, gdyż wartość bardzo niska udziału frakcji lekkiej zdarza się stosunkowo rzadko (1,21% przypadków).

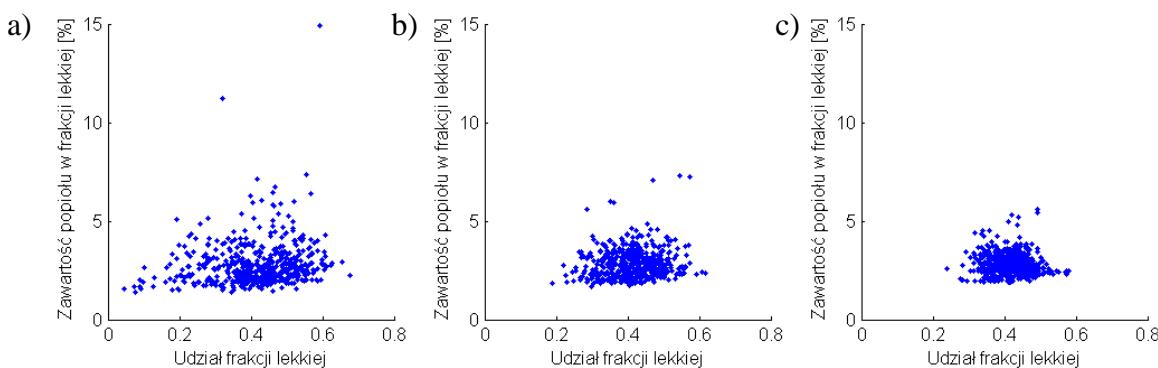


Rys. IV-4. Rozkład prawdopodobieństwa warunkowego wartości lingwistycznych w konkluzji (zawartości popiołu we frakcji lekkiej) dla wybranych reguł plikowych modelu

1.3. Reprezentacja probabilistyczno-rozmyta dwóch zmiennych charakteryzujących węgiel dla danych uśrednionych

W celu pokazania i porównania zmienności rozkładów zmiennych lingwistycznych dla danych o różnym stopniu rozproszenia ich wartości, przedstawiono odpowiednie rozkłady dla uśrednionych wartości udziału frakcji lekkiej i zawartości popiołu w tej frakcji. Uśrednień dokonano odpowiednio z trzech oraz sześciu wartości pomiarowych. Z powodu losowości wyboru badanych próbek w procesie technologicznym pomiary są również najczęściej uśredniane. Zależności danych empirycznych oraz ich uśrednień zamieszczono na rysunku IV-5.

Definicji zbiorów rozmytych oraz wartości lingwistycznych dokonano analogicznie jak w rozdziale IV-1.3. Tabele IV-3÷IV-4 oraz rysunek IV-6 przedstawiają odpowiednio rozkład prawdopodobieństwa łącznego wartości lingwistycznych obu zmiennych, rozkład prawdopodobieństwa brzegowego wartości lingwistycznych udziału frakcji lekkiej oraz rozkład prawdopodobieństwa warunkowego wartości lingwistycznych zawartości popiołu w frakcji lekkiej, dla danych uśrednionych na podstawie 3 kolejnych wartości pomiarowych.



Rys. IV-5. Zależność zawartości popiołu we frakcji lekkiej od masowego udziału węgla tej frakcji dla: a) danych empirycznych, b) danych uśrednionych (średnia z 3 wartości), c) danych uśrednionych (średnia z 6 wartości)

Tab. IV-3. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych wejścia-wyjścia (udziału frakcji lekkiej i zawartości popiołu w tej frakcji)

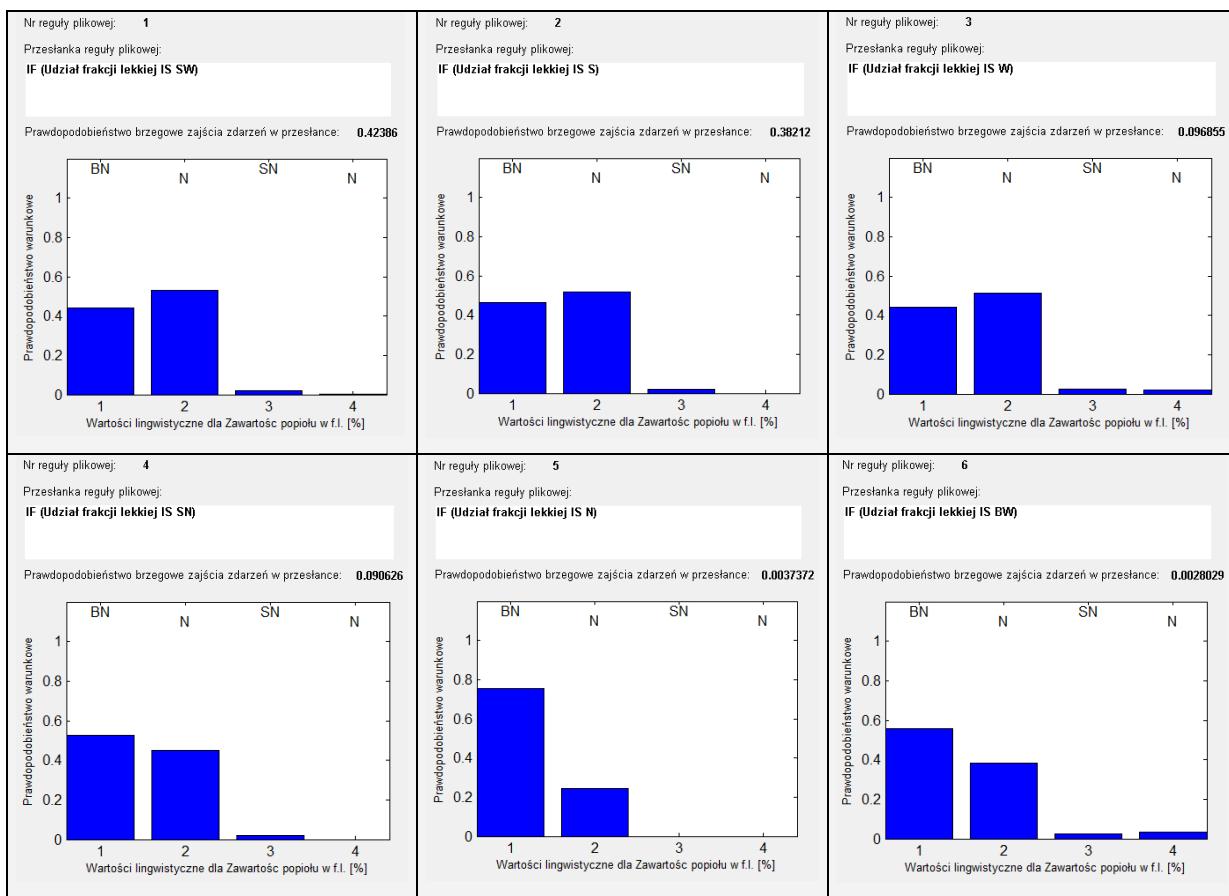
Udział frakcji lekkiej węgla	Zawartość popiołu w frakcji lekkiej [%]						
	1 'BN'	2 'N'	3 'SN'	4 'S'	5 'SW'	6 'W'	7 'BW'
	1 'BN'	0	0	0	0	0	0
2 'N'	0,0028268	0,0009103	0	0	0	0	0
3 'SN'	0,0478164	0,0408212	0,0019884	0	0	0	0
4 'S'	0,1764607	0,1972546	0,0084086	5,05E-19	0	0	0
5 'SW'	0,1875644	0,2256666	0,0094028	0,0012218	0	0	0
6 'W'	0,0426658	0,0496850	0,0023956	0,0021081	0	0	0
7 'BW'	0,0015571	0,0010780	7,19E-05	9,58E-05	0	0	0

Tab. IV-4. Rozkład prawdopodobieństwa brzegowego wartości lingwistycznych przesłanki (udziału frakcji lekkiej)

Lp.	Wartość lingwistyczna	Prawdopodobieństwo brzegowe
1	'SW'	0,4239
2	'S'	0,3821
3	'W'	0,0969
4	'SN'	0,0906
5	'N'	3,70E-03
6	'BW'	2,80E-03

Tabele IV-5÷IV-6 oraz rysunek IV-7 przedstawiają odpowiednio rozkład prawdopodobieństwa łącznego wartości lingwistycznych obu zmiennych, rozkład prawdopodobieństwa brzegowego wartości lingwistycznych udziału frakcji lekkiej oraz rozkład prawdopodobieństwa warunkowego wartości lingwistycznych zwartości popiołu we frakcji lekkiej, dla danych uśrednionych z 6 kolejnych wartości pomiarowych.

W przypadku uśrednień, wyższe wartości łącznego prawdopodobieństwa zajęcia zdarzeń rozmytych przesunęły się w kierunku średnio wysokiej i średniej wartości udziału frakcji lekkiej oraz niskiej zawartości popiołu w tej frakcji, natomiast zdarzenia rozmyte o małej częstości wystąpień zostały zredukowane do zdarzeń niemożliwych.



Rys. IV-6. Rozkład prawdopodobieństwa warunkowego wartości lingwistycznych w konkluzji (zawartości popiołu we frakcji lekkiej) dla wybranych reguł plikowych modelu

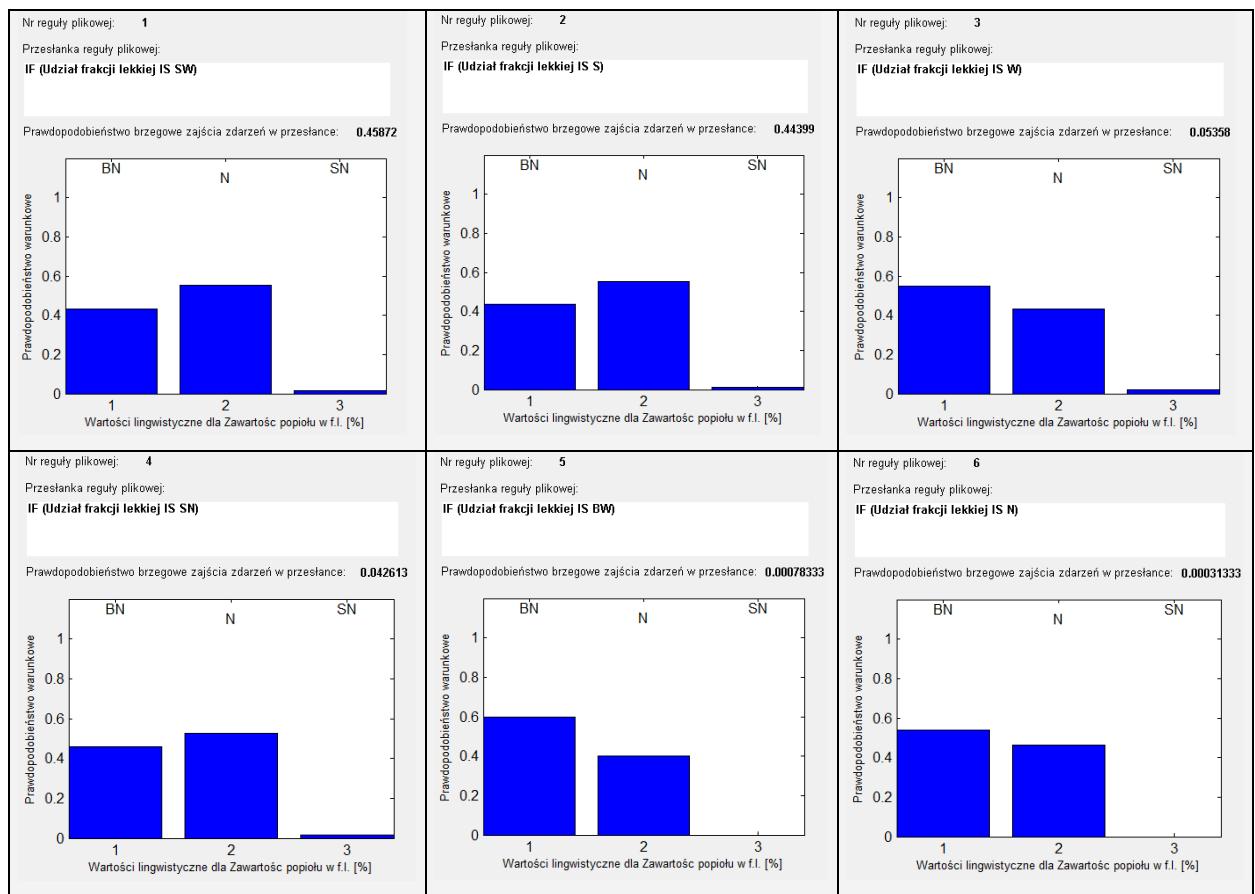
Zaobserwowano również wyższe prawdopodobieństwa brzegowe udziału frakcji lekkiej węgla dla wartości lingwistycznych "średnio wysoki" i "średni". Pozostałe wartości lingwistyczne zmiennej w przesłance pojawiały się znacznie rzadziej. W związku ze zmniejszeniem się ilości łącznych, rozmytych zdarzeń możliwych dla obu zmiennych, ilość konkluzji reguł w modelu uległa zmniejszeniu. Można to zaobserwować na rozkłach prawdopodobieństw warunkowych wartości lingwistycznych zawartości popiołu w frakcji lekkiej (rys. IV-6÷IV-7).

Tab. IV-5. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych wejścia-wyjścia (udziału frakcji lekkiej i zawartości popiołu w tej frakcji)

		Zawartość popiołu w frakcji lekkiej [%]						
		1 'BN'	2 'N'	3 'SN'	4 'S'	5 'SW'	6 'W'	7 'BW'
Udział frakcji lekkiej węgla	1 'BN'	0	0	0	0	0	0	0
	2 'N'	0,0001687	0,0001446	0	0	0	0	0
	3 'SN'	0,0195351	0,0224394	0,0006387	0	0	0	0
	4 'S'	0,1938322	0,2452669	0,0048928	0	0	0	0
	5 'SW'	0,1978693	0,2542089	0,0066402	0	0	0	0
	6 'W'	0,0294532	0,0231384	0,0009882	0	0	0	0
	7 'BW'	0,0004700	0,0003133	0	0	0	0	0

Tab. IV-6. Rozkład prawdopodobieństwa brzegowego wartości lingwistycznych przesłanki (udziału frakcji lekkiej)

Lp.	Wartość lingwistyczna	Prawdopodobieństwo brzegowe
1	'SW'	0,4587
2	'S'	0,444
3	'W'	0,0536
4	'SN'	0,0426
5	'BW'	7,83E-04
6	'N'	3,13E-04

**Rys. IV-7.** Rozkład prawdopodobieństwa warunkowego wartości lingwistycznych w konkluzji (zawartości popiołu w frakcji lekkiej) dla wybranych reguł plikowych modelu

1.4. Reprezentacja probabilistyczno-rozmyta dla filtracji danych dynamicznych

1.4.1. Filtracja sygnałów

Współczesne tendencje rozwoju algorytmów sterowania cyfrowego ukierunkowane są na osiąganie coraz to większej dokładności mimo niepewności pomiaru [zim05]. W celu eliminacji zakłóceń i zniekształceń sygnałów użytych do sterowania wykorzystuje się filtrację pomiarów, która poprawa własności sygnału wejściowego. Filtracja polega na wykonaniu operacji na zbiorze próbek wejściowych zgodnie z deterministyczną funkcją przejścia i podaniu na wyjście przetworzonego ciągu próbek wyjściowych. Wg teorii przetwarzania

układów analogowych i cyfrowych, wraz z przetworzeniem sygnałów następuje przekształcenie jego widma. Niektóre składowe częstotliwościowe widma sygnału mogą być przez filtr uwypuklone, inne stłumione bądź całkowicie usunięte, czyli odfiltrowane. Często też filtracja polega na przekształcaniu widma sygnału wejściowego w widmo sygnału wyjściowego o pożądanym kształcie [sza03].

Filtr wykonuje operację nie tylko na bieżącym sygnale w chwili n , ale dzięki sprzężeniom zwrotnym może również wykorzystywać sygnały wyjściowe opóźnione w czasie bądź też sygnały po chwili n . Ten drugi typ operacji może być jednak wykorzystywany jedynie w filtrach off-line.

Istnieją różne typy filtrów, najczęściej spotyka się następujące charakterystyki:

- filtry liniowe i nieliniowe,
- filtry przyczynowe i niepryczynowe,
- filtry niezmienne w czasie,
- filtry o skończonej i nieskończonej odpowiedzi impulsowej.

W teorii sygnałów dla układu liniowego niezmennego w czasie (LTI, ang. *Linear Time-Invariant*) filtrację jednego sygnału dyskretnego przez drugi można uzyskać dzięki operacji splotu dwóch funkcji $x(n)$ i $h(n)$:

$$y(n) = x(n) \otimes h(n) = \sum_{k=-\infty}^{\infty} x(k)h^*(n-k). \quad (178)$$

Sygnal opisany przez funkcję $x(n)$ jest sygnałem filtrowanym, natomiast sygnał $h(n)$ – filtrującym. Sygnal filtrujący jest zazwyczaj odpowiedzią impulsową filtra przez który przechodzi sygnal, czyli odpowiedzią na pobudzenie dyskretnym impulsem jednostkowym (deltą Kroneckera $\delta(n)$). W wyniku filtracji z sygnału $x(n)$ redukowane są wybrane składowe częstotliwościowe. Zdolność splotu do realizacji filtracji częstotliwościowej jest związana z własnością transformacji Fouriera [zim05].

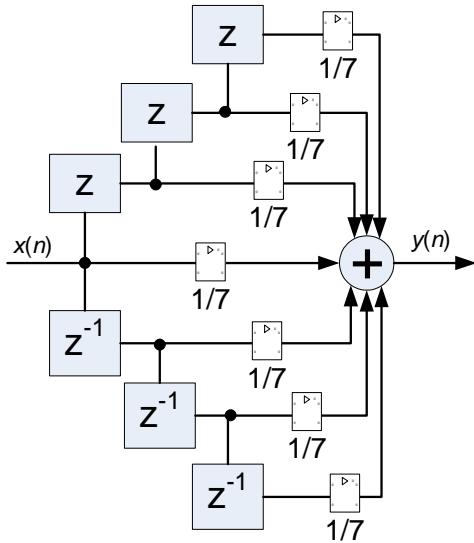
Filtry o skończonej odpowiedzi impulsowej (ang. *Finite Impulse Response FIR*), zwane także filtrami nierekursywnymi, działają w oparciu o bieżące i poprzednie wartości wejściowe. Filtry o nieskończonej odpowiedzi impulsowej (ang. *Finite Impulse Response IIR*) - filtry rekursywne (ang. *Auto-Regressive, AR*) - działają w oparciu o bieżące i poprzednie wartości wejściowe oraz poprzednie wartości wyjściowe.

1.4.2. Analiza właściwości systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy dla filtracji danych dynamicznych

W celu przetestowania właściwości systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy, została zbudowana baza wiedzy dla filtru typu FIR w oparciu o symulację filtru średniej ruchomej bieżących, 3 poprzednich i 3 przyszłych wartości wejściowych układu (rys. IV-8). Wykorzystanie nauczonej bazy wiedzy pozwala na zastosowanie w filtrze typu on-line próbek sygnałów pochodzących nie tylko z chwil poprzednich, ale również aproksymacji próbek sygnałów pochodzących z chwil kolejnych. Wpływa to pozytywnie na szybkość reakcji filtru na sygnał wejściowy.

Pojęcie filtracji w teorii sygnałów losowych często jest utożsamiane z prognozowaniem bieżącej wartości sygnału w chwili n , na podstawie chwil poprzednich. W przykładzie

zostanie opisana filtracja nie jako możliwość predykcji sygnału, ale możliwość zmiany własności sygnału wejściowego.



Rys. IV-8. Schemat blokowy filtru jako średniej ruchomej z 7 sygnałów

Zbiory rozmyte dla wejścia i wyjścia zostały zdefiniowane w oparciu o środki klasterów pozyskanych metodą grupowania k-srednich (rozdz. II-4.2.1). Środki 6 klasterów stanowią punkty charakterystyczne dla funkcji przynależności, na podstawie których otrzymano stopnie przynależności 50 rozłącznych przedziałów dyskretyzacji zmiennych do zbiorów rozmytych o następujących wartościach lingwistycznych:

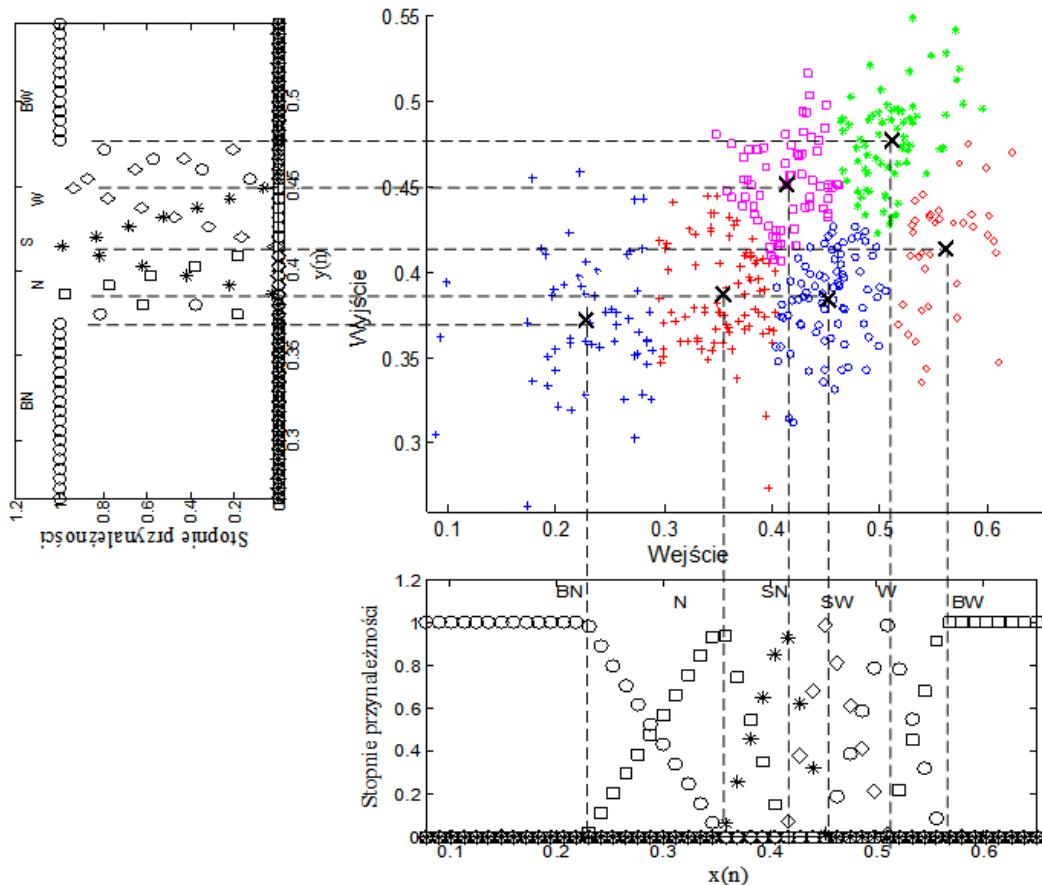
$$L(x)=\{\text{'Bardzo Niski', 'Niski', 'Średnio-Niski', 'Średnio-Wysoki', 'Wysoki', 'Bardzo Wysoki'}\}=\{\text{'BN', 'N', 'SN', 'SW', 'W', 'BW'}\},$$

$$L(y)=\{\text{'Bardzo Niski', 'Niski', 'Średni', 'Wysoki', 'Bardzo Wysoki'}\}=\{\text{'BN', 'N', 'S', 'W', 'BW'}\},$$

gdzie: x jest wejściem, y – wyjściem modelu rozmytego. Na wejście modelu użyto wielkości udziału frakcji lekkiej węgla o gęstości poniżej $1,3 \cdot 10^3 \text{ kg/m}^3$. Definicje zbiorów rozmytych reprezentowane przez powyższe wartości lingwistyczne oraz sposób ich pozyskania zobrazowano na rysunku IV-9.

Ciąg sygnałów pomiarowych $x(n)$ podzielono na dane uczące (400 pomiarów) oraz dane testujące (85 pomiarów). W oparciu o dane uczące zbudowano probabilistyczno-rozmytą bazę wiedzy o pełnym rozkładzie prawdopodobieństwa zdarzeń rozmytych. Analizowano wpływ interpretacji reguł (operatorów implikacji) na dopasowanie modelu wiedzy do matematycznego zapisu funkcji przejścia filtru. Dopasowanie modelu wiedzy było oceniane na podstawie pierwiastka błędu średniokwadratowego wyznaczonego ze wzoru (177). Wyniki błędów zamieszczono w tabeli IV-7. Można zauważyć, że błędy dla danych testujących są niewiele mniejsze od błędów dla danych uczących. Może to świadczyć o dobrych właściwościach uogólniających systemu bądź też jego małym dopasowaniu do danych uczących. Jednakże wartości błędów bliskie wartości odchylenia standardowego wyjściowych danych uczących (0,0496) wskazują na niedopasowanie modelu wiedzy. Przedstawiają to również wykresy IV-10 ÷ IV-11.

Mimo niedopasowania, analizę zagadnienia kontynuowano w celu zbadania właściwości utworzonego modelu.



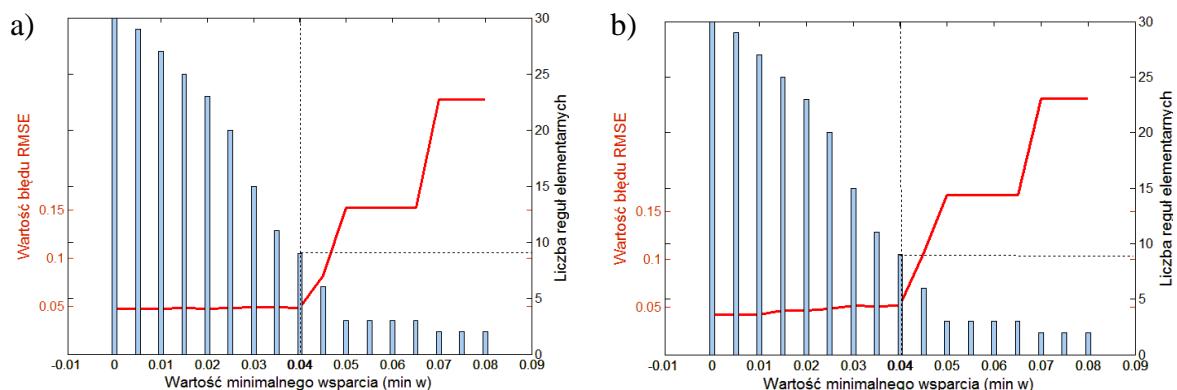
Rys. IV-9. Stopnie przynależności dla wartości lingwistycznych modelu wiedzy oraz sposób ich pozyskiwania na podstawie grupowania danych wejście-wyjście

Tab. IV-7. Dokładność działania systemu jako filtra - średniej ruchomej z 7 sygnałów

Operator interpretacji reguł	RMSE dla danych uczących	RMSE dla danych testujących
Iloczyn algebraiczny	0,0495	0,0481
Minimum	0,0481	0,0468
Implikacja Gödela	0,0485	0,0464
Implikacja Goguena	0,0479	0,0459
Implikacja Łukasiewicza	0,0471	0,0421
Implikacja Reichenbacha	0,0472	0,0421
Implikacja Kleene-Dienesa	0,0473	0,0422
Implikacja Zadeha	0,0480	0,0425

Na przykładzie implikacji rozmytej Łukasiewicza (najmniejszy błąd RMSE – tab. IV-7) badano wpływ wartości minimalnego wsparcia na błąd dopasowania oraz złożoność modelu wiedzy. Dla wartości minimalnego wsparcia równej 0,04, model składa się z 9 reguł elementarnych i wykazuje dopasowanie do symulowanych wartości, zarówno dla danych uczących, jak i danych testujących, na poziomie modelu wiedzy z pełnym rozkładem prawdopodobieństwa zdarzeń rozmytych. Błąd RMSE wskazanego modelu dla danych uczących wynosi 0,0485, dla danych testujących – 0,0516. Utworzona probabilistyczno-rozmyta baza wiedzy, zawiera następujące reguły plikowe:

- 1: IF $x(n)$ IS N [0.1331] THEN $(y(n))$ IS BN [0.6264]
ALSO $(y(n))$ IS S [0.3736]
- 2: IF $x(n)$ IS SN [0.0937] THEN $(y(n))$ IS S [0.5221]
ALSO $(y(n))$ IS BN [0.4779]
- 3: IF $x(n)$ IS SW [0.0911] THEN $(y(n))$ IS S [0.5317]
ALSO $(y(n))$ IS W [0.4683]
- 4: IF $x(n)$ IS BN [0.0841] THEN $(y(n))$ IS BN [1.0000]
- 5: IF $x(n)$ IS W [0.0660] THEN $(y(n))$ IS BW [1.0000]
- 6: IF $x(n)$ IS BW [0.0429] THEN $(y(n))$ IS BW [1.0000].



Rys. IV-10. Zależności błędów dopasowania modelu wiedzy i ilości reguł elementarnych w modelu od wartości minimalnego wsparcia dla: a) danych uczących, b) danych testujących

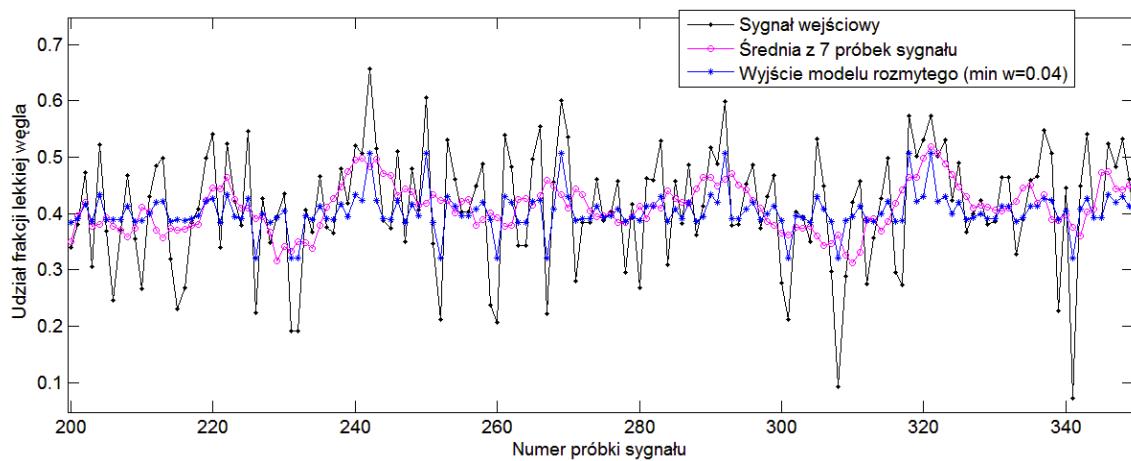
Sposób działania systemu, z modelem wiedzy zawierającym reguły o minimalnym wsparciu na poziomie 0,04, przedstawiono na rysunkach: IV-10 (dla danych uczących) oraz IV-11 (dla danych testujących). Przy analizowanych założeniach wejściowych, sygnał wyjściowy systemu nie naśladuje sygnału filtrowanego za pomocą średniej ruchomej, a kształtem przypomina rzeczywisty sygnał wejściowy stłumiony do wartości średniej, dla którego uwypuklone zostają sygnały o skrajnych wartościach. Jest to wynikiem charakterystyki systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy, który na wyjście podaje wartość oczekiwana wynikowych zbiorów rozmytych.

W celu porównania otrzymanych sygnałów, dokonano analizę ich funkcji autokorelacji (rys. IV-12). Ponadto, używając algorytmu szybkiej dyskretnej transformaty Fouriera utworzono widma przebiegu wejściowego i przebiegów odkształconych o częstotliwości 1/360 Hz (rys. IV-13).

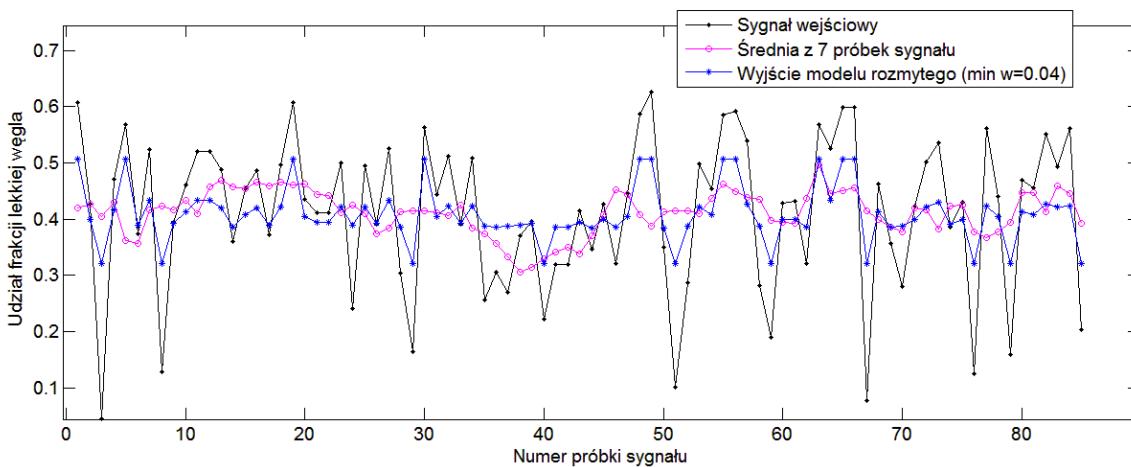
Z kształtu funkcji autokorelacji przebiegu sygnału wejściowego oraz wyjścia modelu rozmytego wynika, że sygnały są dobrze skorelowane jedynie dla przesunięcia równego zero (kolejne próbki nie są ze sobą skorelowane). Sygnały zatem mają właściwości szumu. Natomiast, funkcja autokorelacji średniej ruchomej wykazuje poprawnie korelację 7 pierwszych przesunięć sygnału, czego nie można zauważać na funkcji autokorelacji sygnału wejściowego modelu rozmytego.

Wartości modułu widma częstotliwościowego dla sygnału wejściowego odzwierciedlają właściwości losowe sygnału. Aby, jednak wskazać składowe sinusoidalne występujące w sygnale, należałoby dokonać estymaty funkcji gęstości widmowej mocy – nie jest to jednak celem analizy. Widmo częstotliwościowe dla danych uśrednionych ukazuje poprawnie, że obliczając średnią sygnału działamy na zasadzie filtra dolnoprzepustowego, wygładzając

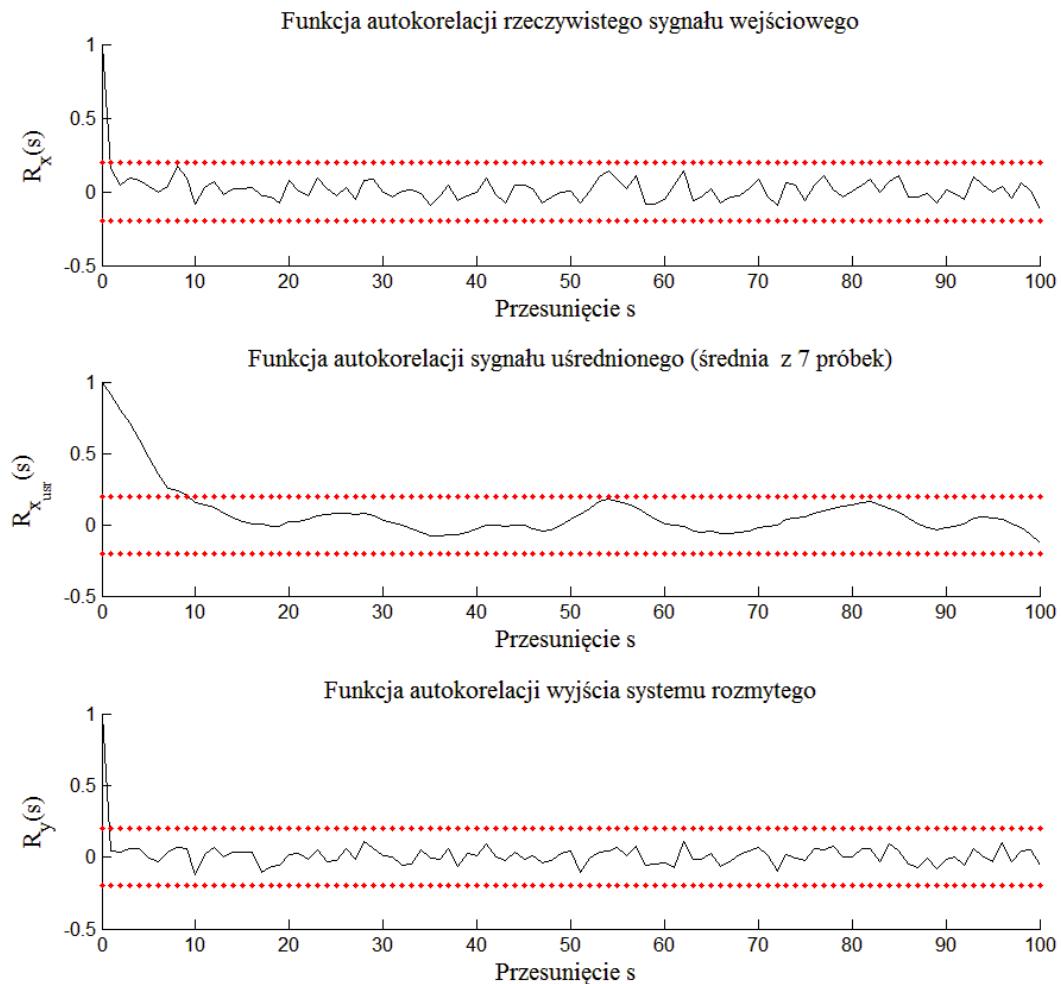
składowe sygnałów o wyższych częstotliwościach. Widmo częstotliwościowe dla wyjścia filtru w przypadku nauczonego modelu wiedzy wygładza składowe sygnałów w całym obserwowanym zakresie pasma częstotliwościowego.



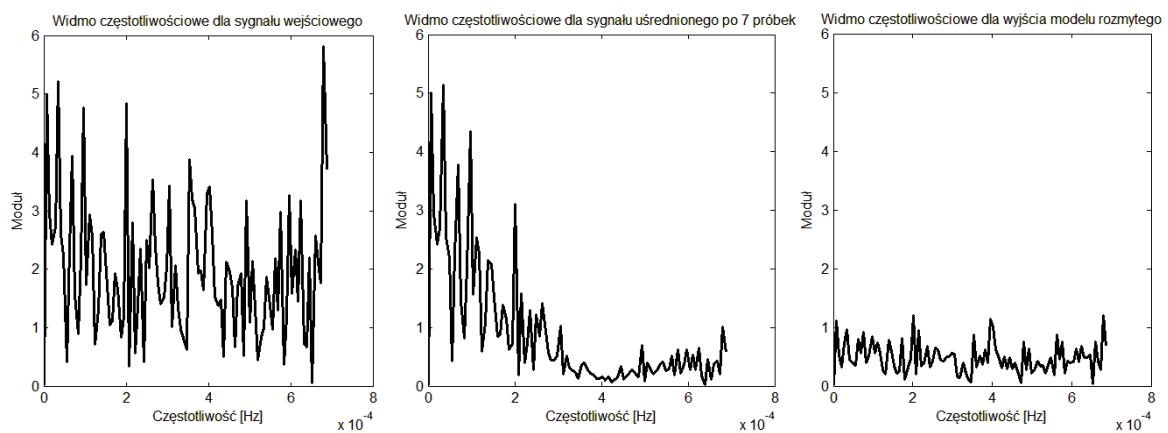
Rys. IV-11. Sygnał wejściowy uczący, sygnał będący średnią ruchomą z 7 próbek danych uczących oraz sygnał wyjściowy systemu rozmytego ($\min w=0,04$)



Rys. IV-12. Sygnał wejściowy testujący, sygnał będący średnią ruchomą z 7 próbek danych testujących oraz sygnał wyjściowy systemu rozmytego ($\min w=0,04$)



Rys. IV-13 Funkcja autokorelacji rzeczywistego sygnału wejściowego, sygnału uśrednionego oraz wyjścia systemu rozmytego



Rys. IV-14. Widmo częstotliwościowe Fouriera dla sygnału wejściowego, uśrednionego po 7 próbek i wyjścia modelu rozmytego

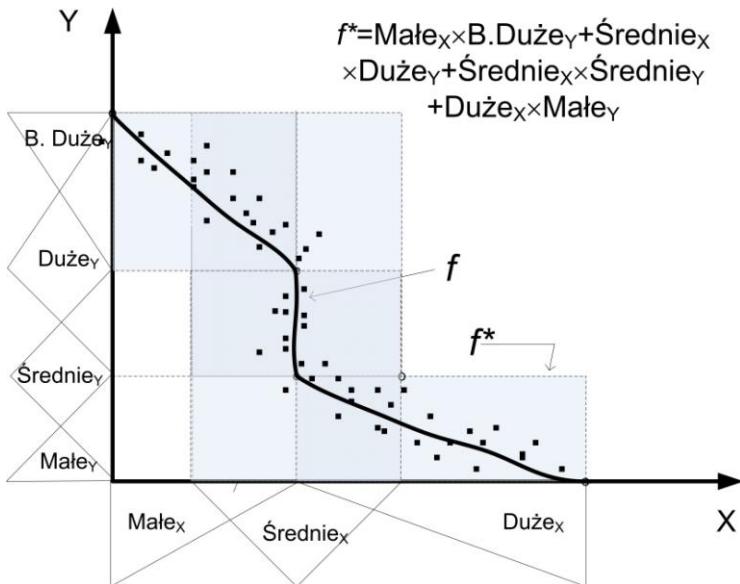
1.5. System wnioskujący z probabilistyczno-rozmytą bazą wiedzy dla pełnej charakterystyki węgla

1.5.1. Nawiązanie do idei *fuzzy graph*

Podstawowym aspektem rozważanym przy tworzeniu wielowymiarowych modeli jest poszukiwanie zależności i relacji pomiędzy analizowanymi zmiennymi. W przypadku zmiennych rozmytych zależności te można przedstawić w postaci tzw. *fuzzy graph*. Koncepcja *fuzzy graph* została wprowadzona przez Zadeh'a w 1971 roku [zad71], a później rozwijana w [zad75a-c], [zad97]. Zadeh, rozważając zmienne rozmyte X i Y oraz wartości lingwistyczne tychże zmiennych w postaci A_i i B_i , $i=1,\dots,n$, *fuzzy graph* przedstawia jako sumę mnogościową iloczynów kartezjańskich:

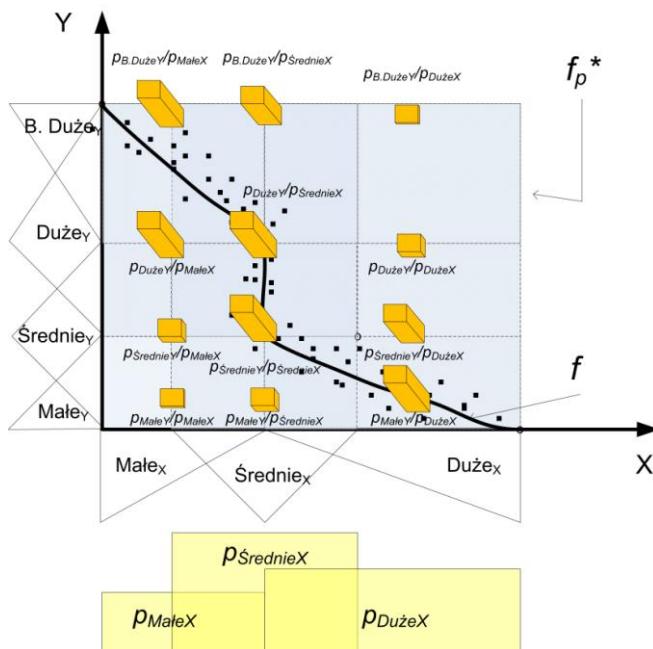
$$f^* = A_1 \times B_1 + \dots + A_n \times B_n = \sum_i A_i \times B_i, i = 1, \dots, n, \quad (179)$$

Właściwością kluczową *fuzzy graph* f^* jest możliwość aproksymacji pewnej funkcji f . Schematyczny przykład aproksymacji funkcji f za pomocą *fuzzy graph* został przedstawiony na rysunku IV-15.



Rys. IV-15. Aproksymacja funkcji f za pomocą *fuzzy graph* f^* (na podstawie [zad97])

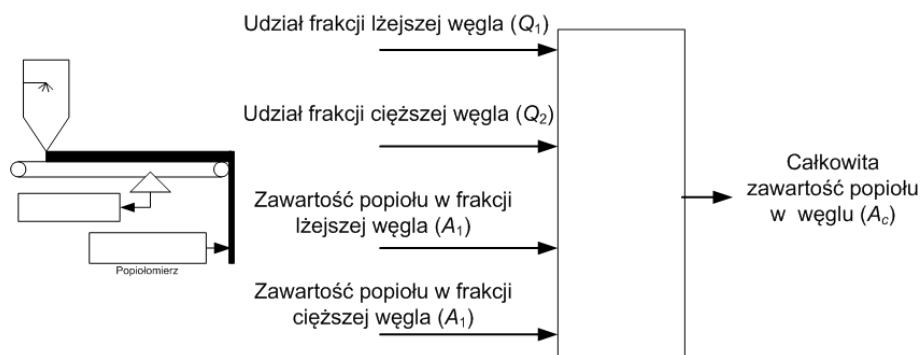
Przedstawiony w niniejszej dysertacji system z probabilistyczno-rozmytą bazą wiedzy pozwala na włączenie prawdopodobieństwa brzegowego i warunkowego zdarzeń rozmytych znajdujących się w regułach do modelu aproksymacji funkcji f , która może odzwierciedlać realizację współzależności zmiennych losowych. Schematyczny przykład aproksymacji funkcji f za pomocą *fuzzy graph* f_p^* z użyciem wspomnianych rozkładów prawdopodobieństw zdarzeń rozmytych został przedstawiony na rysunku IV-16. Włączenie rozkładu prawdopodobieństw pozwala na dopasowanie reguł rozmytych do danych empirycznych bez ingerencji eksperta.



Rys. IV-16. Aproksymacja funkcji f za pomocą fuzzy graph f_p^* z rozkładem prawdopodobieństwa brzegowego i warunkowego zdarzeń rozmytych

1.5.2. Dobór zmiennych lingwistycznych i ich wartości

Przykład fuzzy graph f_p^* zostanie przedstawiony na podstawie modelu statycznego zależności parametrów węgla, który może być wykorzystany do sterowania parametrami procesu technologicznego w zakładach wzbogacania węgla. Więcej informacji na temat węgla jako paliwa kopalnianego zostało zamieszczonych w rozdziale IV-1.1. Obecnie rozważa się zależności wielu zmiennych, tworząc bazę wiedzy jako wielowymiarowy fuzzy graph f_p^* , dla układu przedstawionego na rysunku IV-17.



Rys. IV-17. Schemat układu charakterystyki wzbogacania węgla

Na wejściu układu rozważa się następujące zmienne lingwistyczne:

- Q_1 - masowy udział lżejszej frakcji gęstościowej węgla (w skrócie nazywany udziałem frakcji lżejszej węgla) o gęstości węgla mniejszej od $1,5 \cdot 10^3 \text{ kg/m}^3$,
- Q_2 - masowy udział cięższej frakcji gęstościowej węgla (w skrócie nazywany udziałem frakcji cięższej węgla) o gęstości węgla większej od $1,5 \cdot 10^3 \text{ kg/m}^3$,
- A_1 - zawartość popiołu w frakcji lżejszej węgla [%],

- A_2 - zawartość popiołu w frakcji cięższej węgla [%].

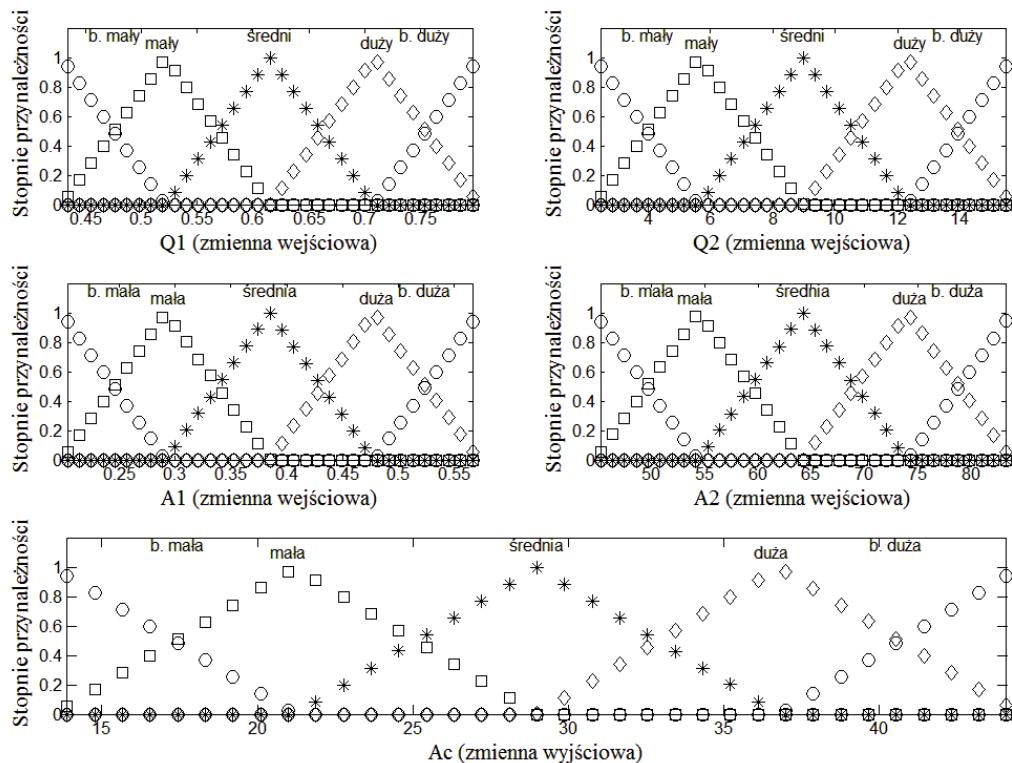
Charakterystyka wzbogacalności węgla, tj. masowy udział elementarnych frakcji ziarnowogęstościowych w węglu oraz wartości atrybutów jakościowych związanych z tymi frakcjami, stanowi o jakości węgla surowego [cie05]. W stosunku do parametru zawartości popiołu, węgiel jest tym lepszy jakościowo, im posiada mniejszą zawartość popiołu. Poprzez zależności powyższych parametrów możemy określić zawartość całkowitą popiołu A_c w próbce węgla. Powyższa wartość wyjściowa układu może posłużyć do podjęcia decyzji o zastosowaniach badanego surowca, wobec tego może być wykorzystywana w zadaniach klasyfikacji i diagnostyki sterowania procesami wzbogacania węgla.

W celu utworzenia bazy wiedzy, reprezentującej tworzony *fuzzy graph* f_p^* przedstawionego układu, dla każdej zmiennej zdefiniowano 5 zbiorów rozmytych (rys. IV-18) reprezentowanych przez następujące wartości lingwistyczne:

$$L(Q_1)=L(Q_2)=\{\text{'b. mały', 'mały', 'średni', 'duży', 'b. duży'}\},$$

$$L(A_1)=L(A_2)=L(A_c)=\{\text{'b. mała', 'mała', 'średnia', 'duża', 'b. duża'}\}.$$

Z uwagi na zastosowanie funkcji typu trójkątnego dobrano nieparzystą liczbę (35) zbiorów dyskretyzacji wartości przestrzeni zmiennych.



Rys. IV-18. Definicja zbiorów rozmytych dla zmiennych systemu

1.5.3. Identyfikacja bazy wiedzy dla pełnej, statycznej charakterystyki węgla

Kryterium jakościowym identyfikacji jest otrzymanie jak najlepszej aproksymacji zależności wartości parametrów wyjściowych i wyjściowego tak, aby móc na podstawie zadanych wartości wejściowych określić właściwą wartość na wyjściu. Jako empiryczne dane

uczące zastosowano 400 rekordów wartości pomiarowych. Zdolność uogólniania otrzymanego systemu bazującego na *fuzzy graph* f_p^* została oceniona wyznaczając, dla kolejnych 95 pomiarów, wskaźnik błędu średniokwadratowego (RMSE). Oczywiście, im wartość wskaźnika błędu dla danych uczących jest mniejsza tym otrzymujemy lepszą aproksymację zależności parametrów, ponadto im wartość wskaźnika błędu dla danych testujących jest mniejszy, tym zdolność do uogólniania systemu jest większa.

Początkowo założono otrzymanie *fuzzy graph* f_p^* z pełnym rozkładem prawdopodobieństwa zdarzeń rozmytych. Badania przeprowadzono dla różnych sposobów interpretacji reguł (za pomocą interpretacji koniunkcyjnej oraz interpretacji logicznej z wykorzystaniem operatora implikacji rozmytej tj.: operatora implikacji Łukasiewicza, Goguena, Gödela, Kleene-Dienesa, Reichenbacha oraz Zadeha) oraz różnych operatorów t-normy, jako sposobu łączenia zdarzeń w poprzedniku reguły. W celu oceny aproksymacji, wartość liczbową wyjścia otrzymywano na podstawie metody środka ciężkości (por. III-5.5). W przypadku rozszerzonej metody środka ciężkości otrzymywano gorsze wyniki aproksymacji. Uzyskane wartości błędu RMSE dla zbioru danych uczących i testujących oraz różnych operatorów są przedstawione w tabeli IV-8.

Tab. IV-8. Błąd RMSE [%] aproksymacji zależności parametrów za pomocą *fuzzy graph* f_p^* w zależności od różnych operatorów interpretacji reguł oraz t-normy jako spójnika logicznego AND (dla modelu uwzględniającego pełny rozkład prawdopodobieństwa zdarzeń rozmytych)

Spójnik log. AND Interpretacja reguł	T-norma Zadeha (min)	T-norma Łukasiewicza	T-norma Fodora	T-norma iloczynu	T-norma Hamachera	T-norma Einstaina	T-norma drastyczna							
Iloczyn algebraiczny	2,148	2,220	15,568	14,387	1,948	2,729	1,913	2,056	2,079	2,162	1,880	2,034	29,678	29,760
Minimum	2,190	2,250	15,567	14,389	1,946	2,728	1,920	2,050	5,120	5,302	1,879	2,017	29,678	29,760
Implikacja Gödela	2,453	2,428	15,572	14,393	1,930	2,720	2,436	2,415	4,087	4,506	2,422	2,395	29,678	29,760
Implikacja Goguena	2,427	2,400	15,571	14,394	1,926	2,715	2,406	2,377	3,867	4,349	2,394	2,363	29,678	29,760
Implikacja Łukasiewicza	3,923	4,243	15,780	14,759	3,383	4,128	4,113	4,468	4,197	4,569	4,150	4,506	29,678	29,757
Implikacja Reichenbacha	4,023	4,357	15,799	14,785	3,563	4,293	4,171	4,526	4,115	4,463	4,195	4,551	29,678	29,757
Implikacja Kleene-Dienesa	4,137	4,488	15,821	14,816	3,784	4,496	4,233	4,590	4,690	4,983	4,243	4,599	29,678	29,756
Implikacja Zadeha	4,202	4,563	15,835	14,831	3,965	4,662	4,259	4,615	4,717	5,009	4,261	4,617	29,678	29,756
	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące		

Na podstawie tabeli, można sformułować następujące wnioski. Największy wpływ na wyniki aproksymacji ma prawidłowy dobór operatora jako spójnika logicznego AND. Rodzaj interpretacji reguł nie wpływa na wynik w tak znaczącym stopniu. Zdolność do aproksymacji dla danych uczących i jej uogólnienie idzie ze sobą w parze tzn. przy małej wartości błędu dla danych uczących, błąd dla danych testujących jest nieco większy ale również mały, i na odwrót, przy dużych wartościach błędów dla danych uczących, błąd dla danych testujących jest jeszcze nieco większy. Dla każdego operatora stanowiącego o interpretacji reguł, najlepszą aproksymację otrzymano przy użyciu t-normy Einstaina, t-normy Fodora, minimum oraz iloczynu algebraicznego. Kolejno gorsze wyniki otrzymano dla t-normy Hamachera, natomiast wykorzystanie t-normy Łukasiewicza oraz t-normy drastycznej powodowało zafałszowanie wyników. Niekwestionowanie od wybrania operatora dla spójnika logicznego AND najlepsze wyniki aproksymacji uzyskano dla interpretacji reguł z wykorzystaniem operatorów: minimum i iloczynu algebraicznego. Porównując wszystkie otrzymane wyniki, najlepszą aproksymację wykazało użycie następującego zestawu parametrów: t-normy Einstaina, interpretacji reguł jako minimum oraz wyostrzania metodą środka ciężkości.

Dla najlepszych parametrów (t-normy Einstaina oraz interpretacji reguł z operatorem minimum) wykreślono zależności błędów RMSE oraz liczby reguł elementarnych od wartości minimalnego wsparcia (*min w*). Rysunek IV-19 przedstawia omawiane zależności zarówno dla

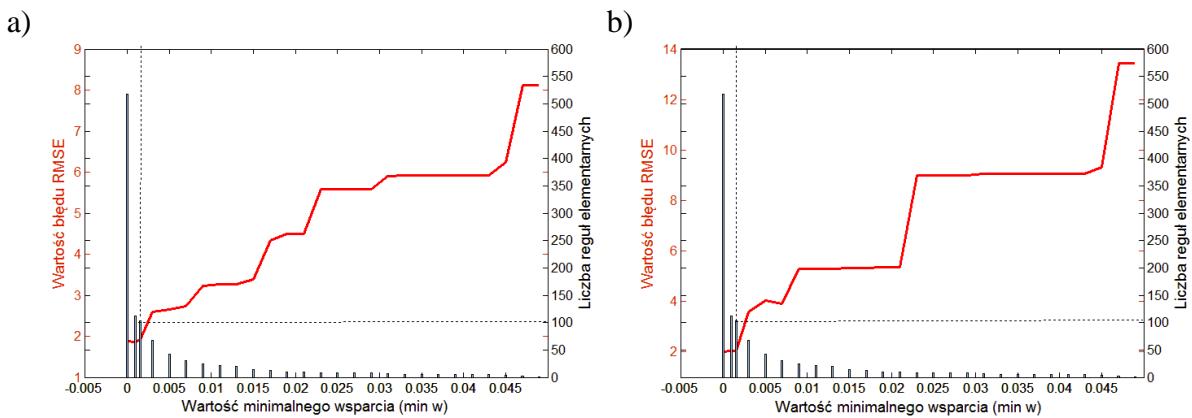
danych uczących, jak i danych testujących. Liczba reguł elementarnych, w przypadku *fuzzy graph* f_p^* zawierającego prawdopodobieństwa brzegowe i warunkowe wynikające z pełnego rozkładu prawdopodobieństw dla zdarzeń rozmytych w regułach, wynosi 518. Wraz z wzrostem wartości minimalnego wsparcia dla reguł, liczba reguł elementarnych maleje, początkowo bardzo szybko, później wolniej. Jednakże, również wraz z wzrostem wartości minimalnego wsparcia, wartości błędów zarówno dla danych uczących, jak i danych testujących ulegają wzrostowi. Biorąc pod uwagę najlepsze dopasowanie wartości wyjścia do danych testujących i uczących oraz jednocześnie możliwie najmniej złożone odwzorowanie zależności, wartość minimalnego wsparcia równa 0,0015 wyznacza optymalną strukturę bazy wiedzy (na rysunku IV-19 zaznaczona linią przerywaną). Wówczas błąd RMSE dla danych uczących wynosi 1,91%, błąd RMSE dla danych testujących - 2,05%, natomiast liczba reguł elementarnych jest równa 103, co stanowi 62 reguły plikowe. Całość *fuzzy graph* f_p^* dla odwzorowania zależności przedstawionych parametrów węgla, została zamieszczona w dodatku C. Poniżej przedstawiono tylko 4 najważniejsze reguły plikowe zależności parametrów:

- 1: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS duża) [0.1581]
 - THEN (Ac IS średnia) [0.6946]
 - ALSO (Ac IS duża) [0.2756]
 - ALSO (Ac IS mała) [0.0298]
- 2: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS duża) [0.0758]
 - THEN (Ac IS średnia) [0.5966]
 - ALSO (Ac IS mała) [0.4034]
- 3: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS duża) [0.0640]
 - THEN (Ac IS średnia) [0.7554]
 - ALSO (Ac IS duża) [0.1965]
 - ALSO (Ac IS mała) [0.0481]
- 4: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0584]
 - THEN (Ac IS średnia) [0.7821]
 - ALSO (Ac IS duża) [0.1103]
 - ALSO (Ac IS mała) [0.1076]

Porównanie ciągów pomiarowych danych uczących i testujących oraz ciągów uzyskanych przy zastosowaniu systemu o optymalnej strukturze probabilistyczno-rozmytej bazy wiedzy zamieszczono na rysunkach IV-27 ÷ IV-28. W celu weryfikacji struktury modelu wiedzy przeprowadzono również analizę błędów dopasowania modelu do danych rzeczywistych, czyli tzw. analizę reszt:

$$e = A_c - \hat{A}_c, \quad (180)$$

gdzie: A_c oznacza ciąg danych odpowiednio uczących lub testujących, \hat{A}_c – ciąg danych uzyskanych na wyjściu utworzonego modelu (odpowiednio dla danych uczących lub testujących).

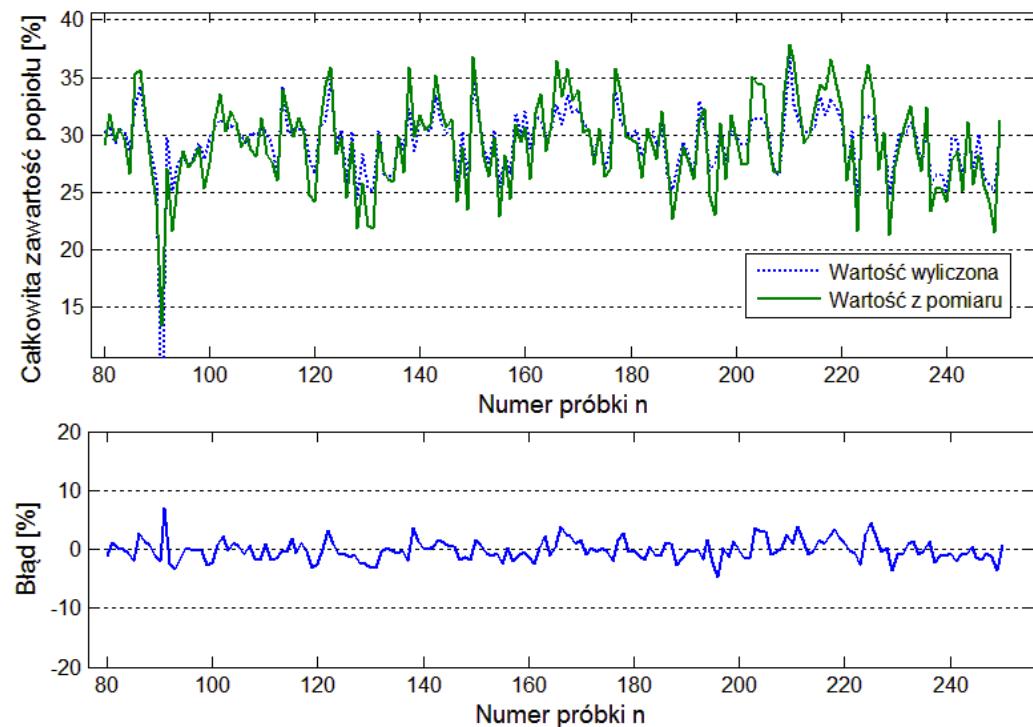


Rys. IV-19. Zależności błędu RMSE oraz liczby reguł elementarnych od wartości minimalnego wsparcia dla:
a) danych uczących, b) danych testujących

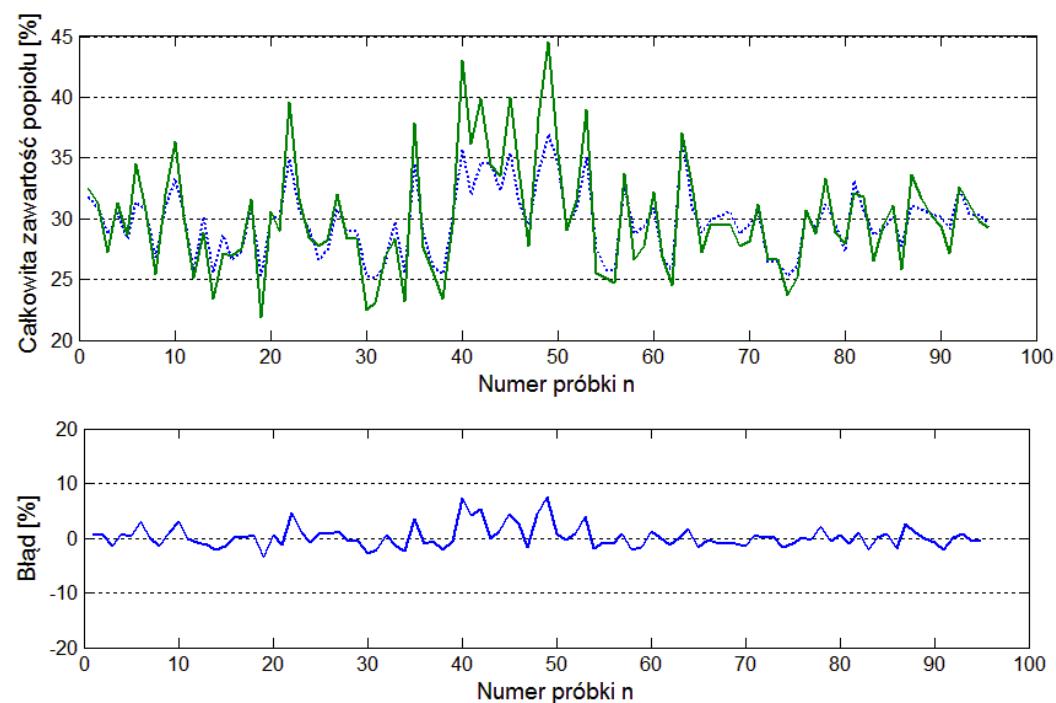
Rysunki IV-20 – IV-21 przedstawiają ciąg błędów dla danych uczących i testujących. Jak można zauważyć, w przypadku danych uczących, błędy rozkładają się równomiernie wzdłuż blisko zerowej wartości średniej zakłóceń (dokładnie wartość średnia wynosiła: -0.1081% , odchylenie standardowe: 1.9121%), ponadto można przypuszczać o wartości stałej wariancji. Przedstawiona na rysunku IV-22 unormowana funkcja autokorelacji błędów $R_e(s)$ (unormowana kowariancja) wyliczona ze wzoru:

$$R_e(s) = \frac{E[(e(n) - \bar{e})(e(n + s) - \bar{e})]}{\sqrt{\sigma_{e(n)}^2 \sigma_{e(n+s)}^2}} = \frac{\sum_{n=1}^{N-s} [(e(n) - \bar{e})(e(n + s) - \bar{e})]}{\sqrt{\sum_{n=1}^{N-s} (e(n) - \bar{e})^2 \sum_{n=1}^{N-s} (e(n + s) - \bar{e})^2}} \quad (181)$$

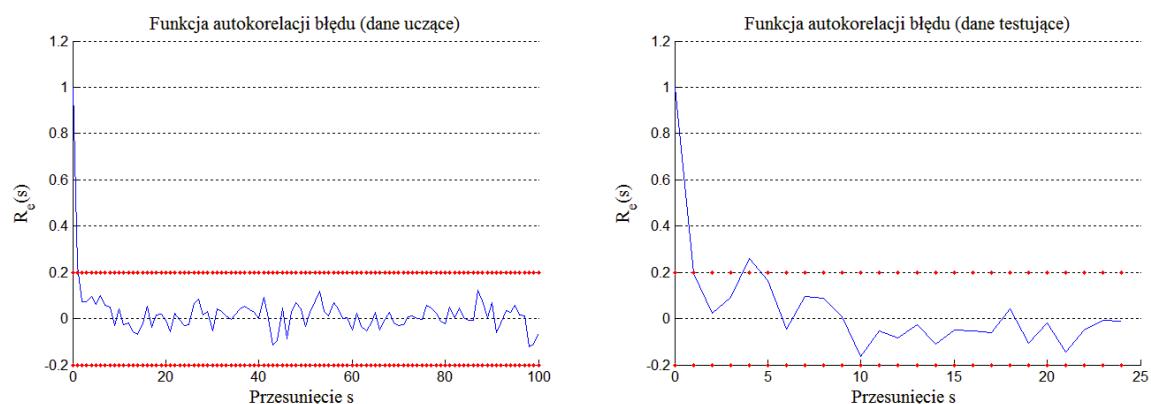
świerczy również o niezależności błędów - kształt funkcji zbliżony jest do funkcji autokorelacji białego szumu. W przypadku błędów dla danych testujących, rozkład ich jest również przeważająco równomierny wokół wartości zerowej, jednak wyjątkiem są próbki o numerach 40-55, gdzie średnia błędu jest większa od wartości zerowej. Wartość średnia błędu dla wszystkich pomiarów danych testujących wynosi $0,24\%$. Przedstawiona na rysunku IV-22 unormowana funkcja autokorelacji błędów $R_e(s)$ dla danych testujących świadczy o ich nieskorelowaniu dla przesunięć, za wyjątkiem przesunięcia $s=4$, dla którego autokorelacja błędu jest większa od wartości $0,2$. Jednakże wartość na poziomie $0,259$ oznacza niską dodatnią korelację błędów [ost99]. Rysunek IV-23 przedstawia histogram błędów dla danych uczących i testujących łącznie. Parametry statystyczne błędów wynoszą odpowiednio: moda $mo=-2.9236$, średnia $\bar{e}=-0.0413$, mediana $me=-0.2503$. Zachodzi wobec tego nierówność $\bar{e} > me > mo$, co świadczy o asymetrii prawostronnej rozkładu błędów. Asymetrię tą można również zaobserwować na wykresach zależności między wartością rzeczywistą wynikającą z pomiarów A_c a wartością wyliczoną na podstawie modelu \hat{A}_c (rys. IV-24). Wykresy te ponadto wyraźnie ukazują, że dla wartości wyjścia bliskiej średniej wartości danych wyjściowych ($\bar{A}_c = 29,37$ dla danych uczących, $\bar{A}_c = 30,04$ dla danych testujących) jest najmniejszy poziom błędów. Im dalej od wartości średniej tym błąd e jest większy – dodatni dla wartości mniejszej od średniej \bar{A}_c , ujemny dla wartości większej od średniej \bar{A}_c .



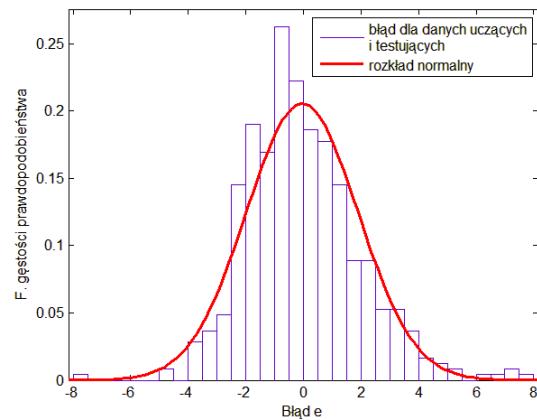
Rys. IV-20. a) Ciąg pomiarów danych uczących (linia kropkowana), oraz ciąg uzyskany przy zastosowaniu systemu z probabilistyczno-rozmytą bazą wiedzy (linia ciągła), b) Błąd jako różnica wartości wyliczonej od wartości rzeczywistej



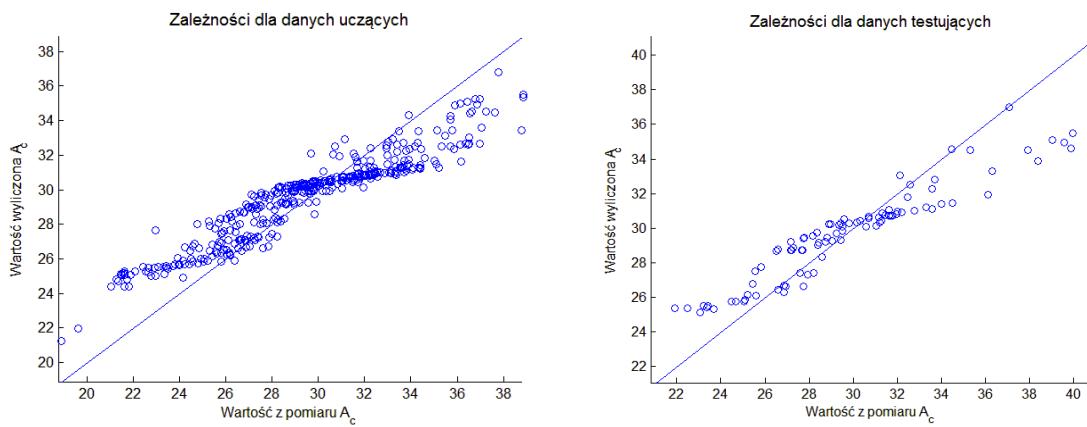
Rys. IV-21. a) Ciąg pomiarów danych testujących (linia kropkowana), oraz ciąg uzyskany przy zastosowaniu systemu z probabilistyczno-rozmytą bazą wiedzy (linia ciągła), b) Błąd jako różnica wartości wyliczonej i wartości rzeczywistej



Rys. IV-22. Unormalowana funkcja autokorelacji błędów dla danych uczących i testujących



Rys. IV-23. Histogram błędów dla danych uczących i testujących łącznie



Rys. IV-24. Zależność między wartością rzeczywistą wynikającą z pomiarów A_c a wartością wyliczoną na podstawie modelu \hat{A}_c , dla danych uczących i testujących

2. System decydujący o wyborze algorytmu do budowy probabilistyczno-rozmytej bazy wiedzy

Przykład zastosowania systemu z probabilistyczno-rozmytą bazą wiedzy jako systemu decyzyjnego, zostanie pokazany na podstawie wyboru efektywnego algorytmu do budowy probabilistyczno-rozmytej bazy reguł. Jak zaznaczono w rozdziale III-3.3, za efektywny algorytm uznaje się algorytm, który generuje bazę reguł w możliwie najkrótszym czasie.

Czas budowy ww. bazy reguł jest zdeterminowany strukturą danych analizowanego problemu badawczego. Dlatego też, istnieją trudności w utworzeniu uściślonego modelu matematycznego, który pozwoli jednoznacznie stwierdzić, jaki algorytm przy danych założeniach wejściowych będzie działać efektywniej dla każdego rodzaju danych uczących. Ta ograniczona wiedza powoduje, iż symulacje na podstawie jednego zbioru danych stają się nie w pełni wiarygodne dla ogółu możliwych danych doświadczalnych. Stąd też proponowane jest zbudowanie rozmytego systemu podejmowania decyzji na podstawie analizowanego w dysertacji systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy. System pozwoli na podjęcie decyzji w oparciu o prawdopodobieństwa zdarzeń rozmytych określonych m.in. na podstawie analiz wykonanych w rozdziale III-3.3. Takie rozwiązanie umożliwia wnioskowanie bazując na wiedzy niepewnej oraz pokaże zależności pomiędzy wartościami lingwistycznymi analizowanych parametrów.

2.1.1. Dobór zmiennych lingwistycznych i ich wartości

W rozdziale III omówiono algorytmy pozwalające na zbudowanie probabilistyczno-rozmytej bazy wiedzy. Określono też, które z nich należy wziąć pod uwagę przy budowie uniwersalnego narzędzia. Parametrami mającymi wpływ na wybór efektywnego algorytmu są:

- liczba zmiennych systemu X_1 [2-5],
- średnia liczba zdefiniowanych wartości rozmytych dla każdej zmiennej wejściowej i wyjściowej X_2 [2-10],
- liczba rozłącznych przedziałów dyskretyzacji wartości zmiennych X_3 [1-50],
- wartość minimalnego wsparcia reguł modelu X_4 [0-0,001],
- liczba rekordów danych uczących X_5 [1-10000].

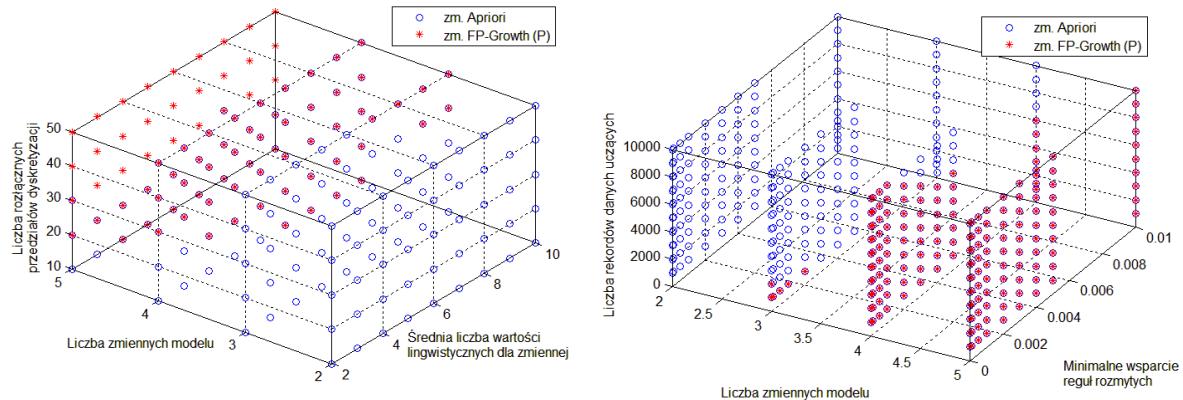
Wynikiem decyzji Y jest wybór jednego z algorytmów: zmodyfikowanego algorytmu Apriori (rozdz. III-3.2.2) (wartość 2) oraz zmodyfikowanego algorytmu FP-Growth (P) (rozdz. III-3.2.3) (wartość 4).

Macierz danych uczących otrzymano w wyniku 15 961 symulacji uczenia probabilistyczno-rozmytej bazy wiedzy na podstawie danych charakteryzujących się kombinacjami poniższych wartości parametrów wejściowych:

- liczba zmiennych systemu ze zbioru {2,3,4,5},
- średnia liczba zdefiniowanych wartości rozmytych dla każdej zmiennej wejściowej i wyjściowej ze zbioru {2,3,...,10},
- liczba rozłącznych przedziałów dyskretyzacji wartości zmiennych ze zbioru {10,20,30,40,50},
- liczba rekordów danych uczących ze zbioru {1000,2000,...,10000},
- wartość minimalnego wsparcia reguł modelu ze zbioru {0, 0.0001, 0.0005, 0.001, 0.002, 0.003, 0.004, 0.005, 0.01}.

Wyjście dla danych uczących stanowi algorytm dla którego czas trwania symulacji, przy jednakowych założeniach wejściowych, był krótszy.

Aby pokazać charakterystykę zmiennych, na podstawie danych uczących zostały przedstawione wykresy zależności pomiędzy danymi wejściowymi i wynikiem decyzji (rys. IV-25). Wynik decyzji na wykresach często jest dwuznaczny z uwagi, iż inny wynik można otrzymać przy założeniu różnych wartości pozostałych, nie ujętych na wykresie zmiennych wejściowych.



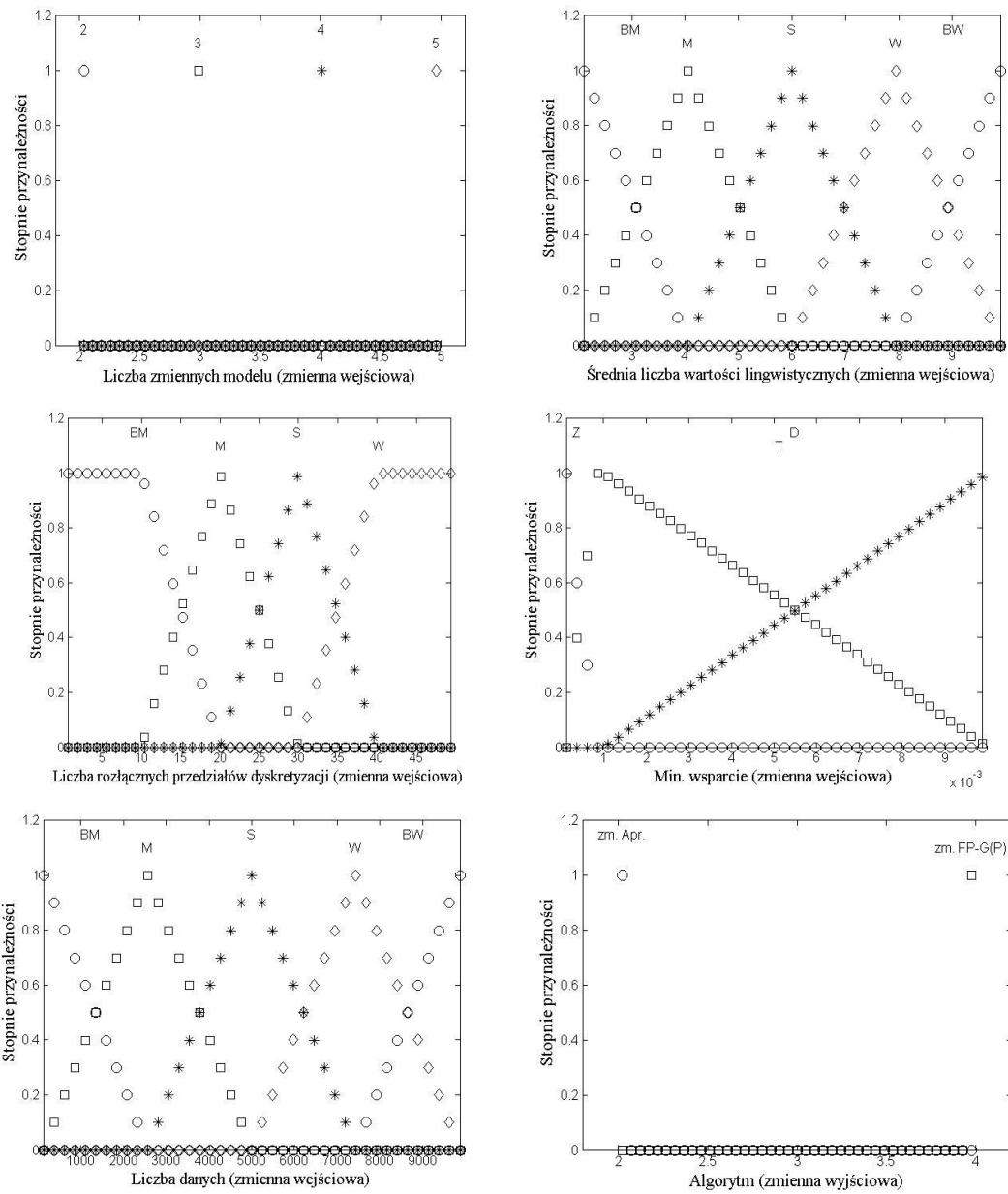
Rys. IV-25. Zależności między wybranymi zmiennymi wejściowymi a zmienną wyjściową modelu rozmytego (na podstawie danych uczących)

Macierz danych testujących jest wektorem wybranych 100 rekordów. Wybór algorytmu dla danych testujących był również określony poprzez mierzenie czasów symulacji obu algorytmów.

Bazę danych dla analizowanego systemu określono następująco:

- zmienna wejściowa: *liczba zmiennych modelu* – cztery wartości lingwistyczne dla zbiorów typu singleton: $L(X_1)=\{'2', '3', '4', '5'\}$,
- zmienna wejściowa: średnia liczba wartości lingwistycznych dla zmiennych w modelu – pięć wartości lingwistycznych: $L(X_2)=\{'Bardzo Mała', 'Mała', 'Średnia', 'Wysoka', 'Bardzo Wysoka'\}=\{'BM', 'M', 'S', 'W', 'BW'\}$,
- zmienna wejściowa: *liczba rozłącznych przedziałów dyskretyzacji* – cztery wartości lingwistyczne: $L(X_3)=\{'Bardzo Mała', 'Mała', 'Średnia', 'Wysoka'\}=\{'BM', 'M', 'S', 'W'\}$,
- zmienna wejściowa: *minimalne wsparcie reguł rozmytych* – trzy wartości lingwistyczne: $L(X_4)=\{'Zerowe', 'Typowe', 'Duże'\}=\{'Z', 'T', 'D'\}$,
- zmienna wejściowa: *liczba rekordów danych uczących* – pięć wartości lingwistycznych: $L(X_5)=\{'Bardzo Mała', 'Mała', 'Średnia', 'Wysoka', 'Bardzo Wysoka'\}=\{'BM', 'M', 'S', 'W', 'BW'\}$,
- zmienna wyjściowa: *wybór algorytmu* – dwie wartości lingwistyczne dla zbiorów typu singleton: $L(Y)=\{'zm. Apriori', 'zm. FP-Growth (P)\}=\{'zm. Apr.', 'zm. FP-G(P)'\}$.

Definicję zbiorów rozmytych dla ww. wartości lingwistycznych zostały zestawione na rysunku IV-26.



Rys. IV-26. Stopnie przynależności przedziałów poszczególnych wartości zmiennych wejściowych i zmiennej wyjściowej do zdefiniowanych zbiorów rozmytych

2.1.2. Identyfikacja probabilistyczno-rozmytej bazy wiedzy dla systemu decyzyjnego

Jako kryterium jakości identyfikacji zastosowano wskaźnik będący stosunkiem liczby błędnych decyzji dla zbioru danych uczących (testujących) do jego liczebności. Oczywiście, przy projektowaniu bazy wiedzy najbardziej interesuje nas procent błędnych decyzji dla danych testujących, który określa zdolność systemu do uogólniania.

Analizowano wpływ wyboru operatorów wnioskowania rozmytego na procent błędnych odpowiedzi systemu, zarówno dla danych uczących, jak i danych testujących. Podczas obliczeń, brano pod uwagę wszystkie możliwe probabilistyczno-rozmyte reguły modelu (wartość minimalnego wsparcia równa 0). Z analizy otrzymanych wyników (tab. IV-9) można wywnioskować, które założenia wnioskowania rozmytego należy wziąć pod uwagę przy

zadanym problemie decyzyjnym. Wartości ostre na wyjściu systemu zostały obliczone na podstawie wyostrzania metodą środka ciężkości COG (rozdział III-5.5).

Tab. IV-9. Procent błędnych decyzji w zależności od różnych operatorów interpretacji reguł oraz t-normy jako spójnika logicznego AND (dla modelu uwzględniającego pełny rozkład prawdopodobieństwa zdarzeń rozmytych)

Spójnik log. AND Interpretacja reguł	T-norma Zadeha (min)	T-norma Łukasiewicza		T-norma Fodora		T-norma iloczynu		T-norma Hamachera		T-norma Einstaina		T-norma drastyczna	
Iloczyn algebraiczny	2,52% 15,00%	6,19%	26,00%	2,75% 12,00%	2,66%	15,00%	2,53% 15,00%	2,66% 14,00%	21,70%	57,00%			
Minimum	2,52% 15,00%	6,19%	26,00%	2,75% 12,00%	2,66%	15,00%	2,53% 15,00%	2,66% 14,00%	21,70%	57,00%			
Implikacja Gödela	5,36% 16,00%	6,21%	27,00%	2,84% 12,00%	5,36%	16,00%	5,36% 16,00%	5,36% 16,00%	21,77%	57,00%			
Implikacja Goguena	5,36% 16,00%	6,21%	27,00%	2,84% 12,00%	5,36%	16,00%	5,36% 16,00%	5,36% 16,00%	21,77%	57,00%			
Implikacja Łukasiewicza	2,52% 15,00%	6,23%	26,00%	2,75% 12,00%	2,66%	15,00%	2,53% 15,00%	2,66% 14,00%	21,70%	57,00%			
Implikacja Reichenbacha	2,52% 15,00%	6,23%	26,00%	2,75% 12,00%	2,66%	15,00%	2,53% 15,00%	2,66% 14,00%	21,70%	57,00%			
Implikacja Kleene-Dienesa	2,52% 15,00%	6,23%	26,00%	2,75% 12,00%	2,66%	15,00%	2,53% 15,00%	2,66% 14,00%	21,70%	57,00%			
Implikacja Zadeha	25,51% 31,00%	30,60%	44,00%	22,92% 34,00%	32,13%	34,00%	30,62%	47,00%	33,51%	38,00%	21,93%	58,00%	
	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	

Podobnie jak w rozdziale IV-1.5.3, większy wpływ na procent poprawnych odpowiedzi ma wybór odpowiedniego operatora odzwierciedlającego spójnik logiczny AND, łączący przesłanki proste w regułach. Dla danych uczących najmniejszy procent błędnych odpowiedzi systemu (2,52%) uzyskano w przypadku użycia t-normy Zadeha (min), nieco większy procent błędu uzyskano przy użyciu t-normy Hamachera (2,53%), iloczynu algebraicznego (2,66%), t-normy Einstaina (2,66%) oraz t-normy Fodora (2,75%). Natomiast, dla danych testujących najniższy procent błędnych odpowiedzi (12%) uzyskano dla t-normy Fodora, większy o dwa punkty procentowe dla t-normy Einstaina oraz większy o trzy punkty procentowe dla t-normy Zadeha, iloczynu algebraicznego i t-normy Hamachera. Wykorzystanie operatora t-normy Łukasiewicza oraz t-normy drastycznej powoduje najmniejsze dopasowanie modelu wiedzy do analizowanego problemu decyzyjnego.

W przypadku określonego operatora t-normy jako spójnika logicznego AND, wyboru interpretacji reguł można dokonać spośród 4 grup, wydzielonych z uwagi na dopasowanie modelu wiedzy do problemu decyzyjnego:

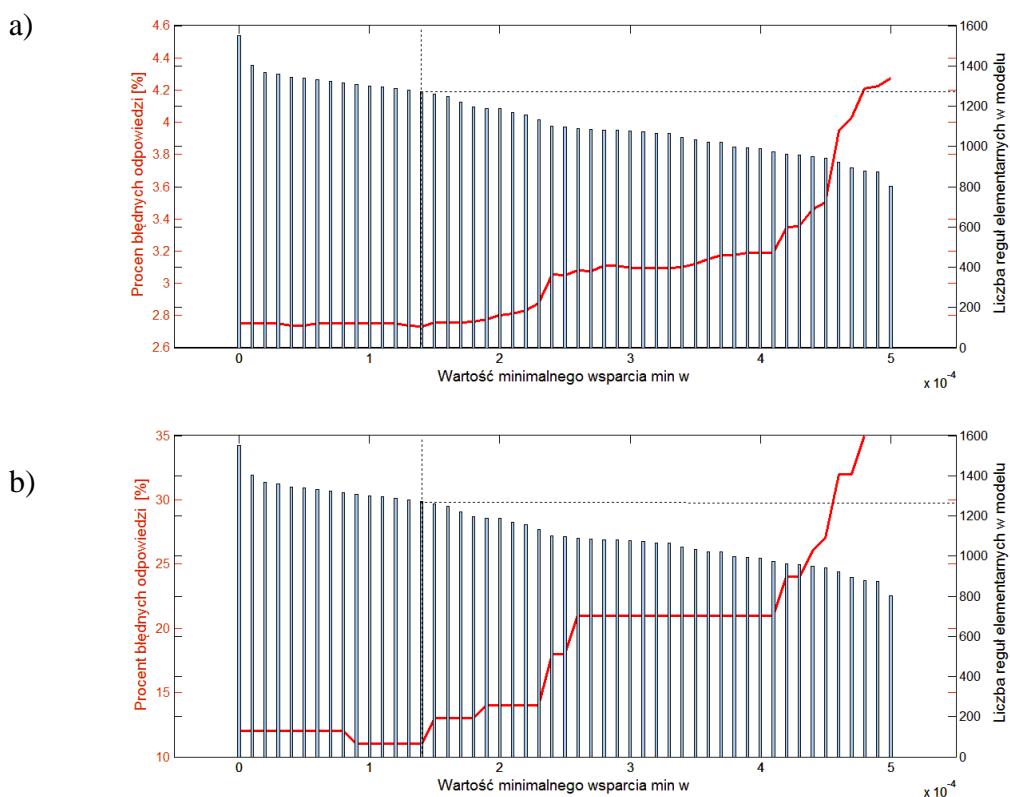
- grupa pierwsza (iloczyn algebraiczny, minimum), najlepsze dopasowanie,
- grupa druga (implikacja rozmyta Łukasiewicza, Reichenbacha, Kleene-Dienesa), niewiele gorsze dopasowanie (różnica dla t-normy Łukasiewicza),
- grupa trzecia (implikacja rozmyta Gödela, Goguena), słabe dopasowanie,
- grupa czwarta (implikacja rozmyta Zadeha), model niedopasowany.

Istnieje duża rozbieżność między błędami dla danych uczących i danych testujących. Powodem rozbieżności jest sposób doboru ww. danych. Rekordów uczących było aż 15 961. Składają się one z kombinacji wartości zmiennych wejściowych pochodzących z całych zakresów ich zmienności. Zawierają zatem statystyczną większość wartości zmiennych wejściowych dla których powierzchnia modelu jest gładka, a udział danych dla których powierzchnia modelu jest pofałdowana jest znacznie mniejszy, niż w przypadku danych testujących. Stąd też możliwości dokonania błędów są też znacznie mniejsze. Danych testujących jest dużo mniej (w ilości 100) i zostały tak wyselekcjonowane, aby wartości zmiennych dotyczyły możliwie jak najwięcej przypadków na granicy zmiany efektywności algorytmów.

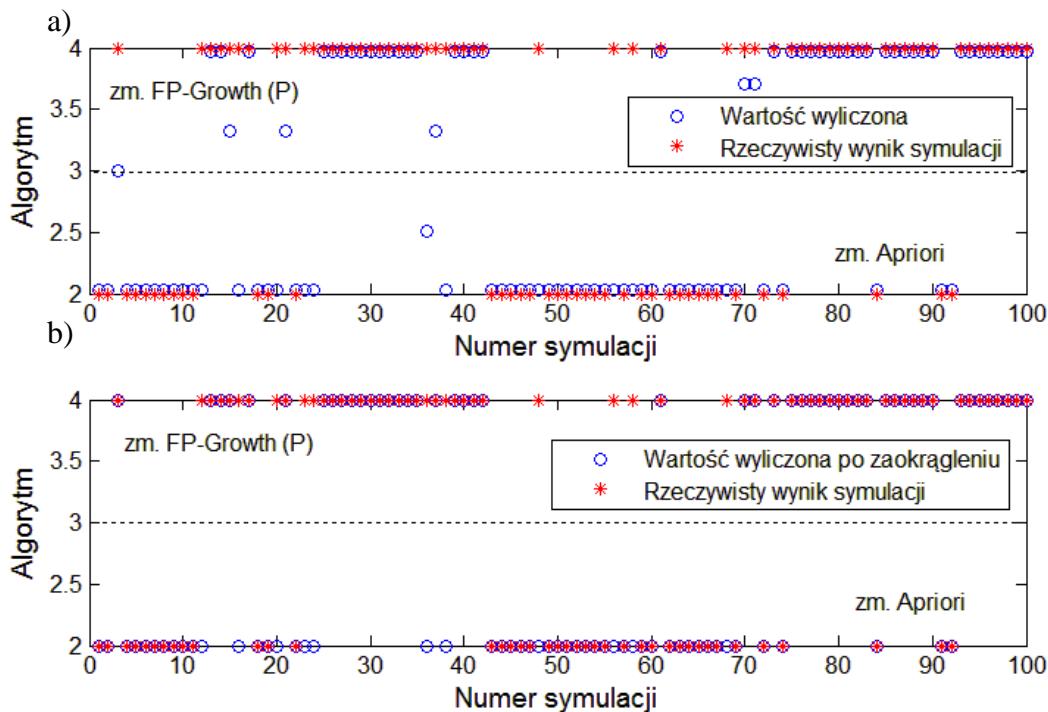
Ponieważ najważniejsza jest zdolność systemu do uogólniania, postanowiono wybrać operatory do wnioskowania rozmytego w oparciu o najwyższy uzyskany procent błędnych

decyzji dla danych testujących. Mianowicie, wybrano t-normę Fodora, jako odzwierciedlenie iloczynu przesłanek prostych w regule, oraz interpretację reguł Larsena (iloczyn algebraiczny).

Dla ww. założeń ograniczano liczbę reguł modelu wiedzy zwiększając wartość minimalnego ich wsparcia. Badano również wpływ zmiany parametrów modelu na procent błędnych decyzji, dla danych uczących i danych testujących. Wyniki przedstawiono na rysunku IV-27. Z wykresów można wyznaczyć optymalną wartość parametru minimalnego wsparcia, dla którego błąd odpowiedzi systemu jest najmniejszy oraz struktura modelu wiedzy jest możliwie najmniej skomplikowana. Dla wartości minimalnego wsparcia z przedziału $\langle 9 \cdot 10^{-5}, 1,4 \cdot 10^{-4} \rangle$ procent błędnych odpowiedzi, dla danych testujących, zmalał z 12% do 11%. Procent błędnych odpowiedzi z wykorzystaniem danych uczących, dla ww. wartości minimalnego wsparcia, ustalił się na tym samym poziomie, jak dla pełnego rozkładu prawdopodobieństwa zdarzeń rozmytych (2,73%). Stąd też optymalną wartością minimalnego wsparcia jest $1,4 \cdot 10^{-4}$, dla której otrzymujemy najniższe błędy przy mniejszej liczbie reguł elementarnych. Niestety model zmniejszył się jedynie o 282 reguły elementarnego w stosunku do pełnego rozkładu zdarzeń rozmytych. Cały model systemu zawiera więc 1270 reguł elementarnych, co stanowi 1186 reguł plikowych. Jednakże w przypadku zastosowania systemu do podejmowania decyzji o wyborze algorytmu generowania bazy wiedzy w implementowanym narzędziu, złożoność modelu nie jest tak istotna jak trafność wyboru.



Rys. IV-27. Zależności procentu błędnych odpowiedzi systemu oraz liczby reguł elementarnych od wartości minimalnego wsparcia dla: a) danych uczących, b) danych testujących



Rys. IV-28. Porównanie wyników symulacji z wartościami wyliczonymi na podstawie modelu z probabilistyczno-rozmytą bazą wiedzy (dla danych testujących)

Rysunek IV-28 przedstawia porównanie rzeczywistych wyników symulacji danych testujących y z wartościami wyliczonymi na podstawie utworzonego modelu wiedzy \hat{y} . Podpunkt a) rysunku pokazuje wartości \hat{y} otrzymane bezpośrednio po wnioskowaniu rozmytym i wyostrzaniu metodą środka ciężkości COG. Natomiast podpunkt b) rysunku przedstawia wartości po uwzględnieniu zależności:

$$\text{algorytm} = \begin{cases} 2 & (\text{zm.} \text{Apriori}) \text{ dla } \hat{y} < 3 \\ 4 & (\text{zm.} \text{FP} - \text{Growth}) \text{ dla } \hat{y} \geq 3 \end{cases}. \quad (182)$$

Poniżej znajdują się bardziej znaczące reguły plikowe utworzonego modelu wiedzy:

- 1: IF (X1 IS 2) AND (X2 IS W) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. Apr.) [1.0000]
- 2: IF (X1 IS 3) AND (X2 IS W) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. Apr.) [1.0000]
- 3: IF (X1 IS 5) AND (X2 IS W) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. FP-G(P)) [1.0000]
- 4: IF (X1 IS 2) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. Apr.) [1.0000]
- 5: IF (X1 IS 2) AND (X2 IS S) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. Apr.) [1.0000]

- 6: IF (X1 IS 3) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. Apr.) [1.0000]
- 7: IF (X1 IS 3) AND (X2 IS S) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. Apr.) [1.0000]
- 8: IF (X1 IS 5) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. FP-G(P)) [1.0000]
- 9: IF (X1 IS 5) AND (X2 IS S) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. FP-G(P)) [1.0000]
- 10: IF (X1 IS 4) AND (X2 IS W) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. FP-G(P)) [0.7657] ALSO (Y IS zm. Apr.) [0.2343]
- 11: IF (X1 IS 4) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. FP-G(P)) [0.5350] ALSO (Y IS zm. Apr.) [0.4650]
- 12: IF (X1 IS 4) AND (X2 IS S) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0029]
THEN (Y IS zm. FP-G(P)) [0.7038] ALSO (Y IS zm. Apr.) [0.2962]
- 13: IF (X1 IS 2) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS W) [0.0027]
THEN (Y IS zm. Apr.) [1.0000]
- 14: IF (X1 IS 5) AND (X2 IS W) AND (X3 IS W) AND (X4 IS T) AND (X5 IS W) [0.0027]
THEN (Y IS zm. FP-G(P)) [1.0000]
- 15: IF (X1 IS 4) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS M) [0.0027]
THEN (Y IS zm. FP-G(P)) [0.9360] ALSO (Y IS zm. Apr.) [0.0640]
- 16: IF (X1 IS 4) AND (X2 IS M) AND (X3 IS W) AND (X4 IS T) AND (X5 IS W) [0.0027]
THEN (Y IS zm. Apr.) [0.8550] ALSO (Y IS zm. FP-G(P)) [0.1450]
- 17: IF (X1 IS 3) AND (X2 IS S) AND (X3 IS W) AND (X4 IS T) AND (X5 IS M) [0.0027]
THEN (Y IS zm. Apr.) [1.0000]
- 18: IF (X1 IS 4) AND (X2 IS W) AND (X3 IS W) AND (X4 IS T) AND (X5 IS M) [0.0026]
THEN (Y IS zm. FP-G(P)) [1.0000]
- 19: IF (X1 IS 2) AND (X2 IS BM) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0022]
THEN (Y IS zm. Apr.) [1.0000]
- 20: IF (X1 IS 3) AND (X2 IS BM) AND (X3 IS W) AND (X4 IS T) AND (X5 IS S) [0.0022]
THEN (Y IS zm. Apr.) [1.0000]
- 21: IF (X1 IS 5) AND (X2 IS BM) AND (X3 IS W) AND (X4 IS T) AND (X5 IS BW) [0.0015]
THEN (Y IS zm. FP-G(P)) [1.0000]
- 22: IF (X1 IS 3) AND (X2 IS BW) AND (X3 IS W) AND (X4 IS T) AND (X5 IS BW) [0.0015]
THEN (Y IS zm. Apr.) [1.0000]
- 23: IF (X1 IS 4) AND (X2 IS BW) AND (X3 IS W) AND (X4 IS T) AND (X5 IS BW) [0.0015]
THEN (Y IS zm. Apr.) [0.7787] ALSO (Y IS zm. FP-G(P)) [0.2213]
- 24: IF (X1 IS 2) AND (X2 IS S) AND (X3 IS W) AND (X4 IS Z) AND (X5 IS S) [0.0015]
THEN (Y IS zm. Apr.) [1.0000]
- 25: IF (X1 IS 2) AND (X2 IS W) AND (X3 IS W) AND (X4 IS Z) AND (X5 IS S) [0.0015]
THEN (Y IS zm. Apr.) [1.0000]

2.1.3. Analiza probabilistyczno-rozmytej bazy wiedzy przy faktach wyrażonych za pomocą wartości lingwistycznych i umerycznych

Analizując strukturę modelu wiedzy utworzonego systemu wnioskującego możemy dowiedzieć się o charakterystyce badanego zadania. W rozważanym przykładzie, prawdopodobieństwo jednoczesnego zajścia zdarzeń rozmytych znajdujących się w regułach wskazuje, który algorytm jest bardziej efektywny w danej sytuacji – przy określonych wartościach lingwistycznych zmiennych wejściowych. Wybrane zależności przedstawione zostały w tabelach IV-10 ÷ IV-12.

Tabela IV-10 ukazuje wybór konkretnego algorytmu dla średnio bardzo wysokiej liczby wartości lingwistycznej, wysokiej liczby rozłącznych przedziałów dyskretyzacji przestrzeni wartości zmiennej oraz wartości minimalnego wsparcia bliskiej zeru. Wówczas, dla modelu typu SISO, dla dowolnej liczby danych uczących do budowy probabilistyczno-rozmytej bazy wiedzy lepiej wybrać zmodyfikowany algorytm Apriori. Przy dwóch wejściach do systemu, algorytm Apriori nie jest zalecany dla bardzo małej liczby danych uczących i w 22,7% przypadkach – dla małej liczby danych uczących. Natomiast dla trzech wejść do systemu bez względu na liczbę danych uczących zalecane jest użycie zmodyfikowanego algorytmu FP-Growth (P).

Tab. IV-10. Prawdopodobieństwo jednoczesnego zajścia wybranych zdarzeń rozmytych w regułach

Założenia wartości lingwistycznych pozostałych zmiennych wejściowych				Prawdopodobieństwo jednoczesnego zajścia zdarzeń rozmytych w regułach		
				Wynik		
Liczba zmiennych	Śred. liczba wartości ling.	Liczba rozłącznych przedziałów dyskretyz.	Min. wsparcie	Liczba rekordów danych uczących	'zm. Apr.'	'zm. FP-G(P)'
'2'	'BW'	'W'	'Z'	'BM'	0,00034	0
				'M'	0,00102	0
				'S'	0,00110	0
				'W'	0,00102	0
				'BW'	0,00076	0
Zmienne wejściowe				Wynik		
Liczba zmiennych	Śred. liczba wartości ling.	Liczba rozłącznych przedziałów dyskretyz.	Min. wsparcie	Liczba rekordów danych uczących	'zm. Apr.'	'zm. FP-G(P)'
'3'	'BW'	'W'	'Z'	'BM'	0	0,00026
				'M'	0,00078	0,00023
				'S'	0,00110	0
				'W'	0,00102	0
				'BW'	0,00076	0
Zmienne wejściowe				Wynik		
Liczba zmiennych	Śred. liczba wartości ling.	Liczba rozłącznych przedziałów dyskretyz.	Min. wsparcie	Liczba rekordów danych uczących	'zm. Apr.'	'zm. FP-G(P)'
'4'	'BW'	'W'	'Z'	'BM'	0	0,00034
				'M'	0	0,00102
				'S'	0	0,00110
				'W'	0	0,00101
				'BW'	0	0,00072

Dla dwóch zmiennych wejściowych, średnio bardzo wysokiej liczby wartości lingwistycznych każdej zmiennej systemu, wysokiej liczby rozłącznych przedziałów dyskretyzacji, ale za to bardzo niskiej liczby rekordów danych uczących, algorytm

FP-Growth (P) jest zalecany przy wartości minimalnego wsparcia bliskiej zeru oraz tylko w 20,9% przypadków przy typowej wartości wsparcia (tab. IV-11). W pozostałych przypadkach, także dla średniej liczby rozłącznych przedziałów i małej liczby rekordów danych uczących, uznaje się zmodyfikowany algorytm za bardziej efektywny.

Tabela IV-12 pokazuje analogiczne zależności wyboru algorytmu od różnej liczby zmiennych modelu, przy zadanych wartościach pozostałych zmiennych wejściowych.

Tab. IV-11. Prawdopodobieństwo jednoczesnego zajścia wybranych zdarzeń rozmytych w regułach

Założenia wartości lingwistycznych pozostałych zmiennych wejściowych				Prawdopodobieństwo jednoczesnego zajścia zdarzeń rozmytych w regułach		
Zmienne wejściowe				Wynik		
Liczba zmiennych	Śred. liczba wartości ling.	Liczba rozłącznych przedziałów dyskretyz.	Liczba rekordów danych ucz.		'zm. Apr.'	'zm. FP-G(P)'
'3'	'BW'	'W'	'BM'	Min. wsparcie	'Z'	0,00026
					'T'	0,00053
					'D'	0,00031
Zmienne wejściowe				Wynik		
'3'	'BW'	'S'	'M'	Min. wsparcie	'zm. Apr.'	'zm. FP-G(P)'
					'Z'	0
					'T'	0,00107
					'D'	0,00050

Tab. IV-12. Prawdopodobieństwo jednoczesnego zajścia wybranych zdarzeń rozmytych w regułach

Założenia wartości lingwistycznych pozostałych zmiennych wejściowych				Prawdopodobieństwo jednoczesnego zajścia zdarzeń rozmytych w regułach		
Zmienne wejściowe				Wynik		
Liczba rekordów danych ucz.	Śred. liczba wartości ling.	Liczba rozłącznych przedziałów dyskretyz.	Min. wsparcie		'zm. Apr.'	'zm. FP-G(P)'
'M'	'BW'	'W'	'Z'	Liczba zmiennych	'2'	0,00102
					'3'	0,00078
					'4'	0,00102
					'5'	0,00102
Zmienne wejściowe				Wynik		
'S'	'BW'	'W'	'Z'	Liczba zmiennych	'zm. Apr.'	'zm. FP-G(P)'
					'2'	0,00110
					'3'	0,00110
					'4'	0,00110
Zmienne wejściowe				Wynik		
'S'	'BW'	'W'	'T'	Liczba zmiennych	'zm. Apr.'	'zm. FP-G(P)'
					'2'	0,00220
					'3'	0,00220
					'4'	0,00054
					'5'	0,00165

Biorąc pod uwagę konkretne ilościowe wartości wejściowe systemu decyzyjnego, możemy również, na podstawie wag modelu wiedzy, otrzymać prawdopodobieństwo warunkowe osiągnięcia danego wyniku.

Niech wejścia systemu będą równe odpowiednio:

- $x_1=4$ (ilość zmiennych modelu),
- $x_2=7$ (średnia ilość zbiorów rozmytych dla każdej zmiennej),
- $x_3=35$ (liczba rozłącznych przedziałów dyskretyzacji wartości zmiennych),
- $x_4=0,0001$ (wartość minimalnego wsparcia),
- $x_5=5250$ (liczba rekordów danych uczących).

Wówczas, aktywowane są 2 reguły plikowe modelu wiedzy, dla których stopień aktywacji τ_o (przy zastosowaniu t-normy Fodora) kształtuje się następująco:

466: IF (L. zmiennych IS 3) AND (Śr. l. wartości ling. IS W) AND (L. rozł. przedz. dyskr. IS S) AND (Min. w IS Z) AND (L. danych IS S) [0.000777]

THEN (Algorytm IS zm. FP-G(P)) [0.7471]

ALSO (Algorytm IS zm. Apr.) [0.2529] $\Rightarrow \tau_{466} = 0,5$,

472: IF (L. zmiennych IS 3) AND (Śr. l. wartości ling. IS S) AND (L. rozł. przedz. dyskr. IS S) AND (Min. w IS Z) AND (L. danych IS S) [0.000777]

THEN (Algorytm IS zm. FP-G(P)) [0.5545]

ALSO (Algorytm IS zm. Apr.) [0.4455] $\Rightarrow \tau_{472} = 0,5$,

Wykorzystując wagi probabilistyczno-rozmytej bazy wiedzy, można stwierdzić, iż przy warunku zajścia powyżej opisanych wartościach numerycznych na wejściach systemu:

- efektywnym algorytmem wyszukiwania reguł modelu jest zmodyfikowany algorytm Apriori z prawdopodobieństwem 0,3492,
- efektywnym algorytmem wyszukiwania reguł modelu jest zmodyfikowany algorytm FP-Growth (P) z prawdopodobieństwem 0,6508.

Posługując się pełnym wnioskowaniem rozmytym opisanym w rozdziale III-5 pracy wraz z wyostrzaniem wynikowego zbioru rozmytego metodą środka ciężkości, otrzymujemy na wyjściu systemu wartość ostrą $\hat{y} = 3.2943$. Po zastosowaniu zależności (182), otrzymujemy podpowiedź, że przy danych założeniach wejściowych, do budowy bazy wiedzy, powinniśmy wykorzystać zmodyfikowany algorytm FP-Growth (P). Jest to zgodnie z wcześniejszym rozważaniem.

W dodatku D pracy zostały zamieszczone wyniki identyfikacji modelu wiedzy analogicznego systemu wnioskującego, przy założeniu większej liczby rozłącznych przedziałów dyskretyzacji w przestrzeniach zmiennych oraz innych definicji zbiorów rozmytych dla wielkości liczby rekordów danych uczących. Otrzymano wówczas, dla bazy wiedzy z pełnym rozkładem prawdopodobieństwa zdarzeń rozmytych, mniejszy procent błędów dla danych uczących (1,48%) (przy założeniu t-normy Zadeha (min) jako spójnik logiczny AND) i taki sam procent błędów dla danych testujących (12%). Jednakże liczba reguł takiego modelu jest o wiele większa, a zwiększenie wartości minimalnego wsparcia powoduje także zwiększenie błędów decyzji (np. dla $\min w = 2 \cdot 10^{-4}$ procent błędnych odpowiedzi wynosi 1,58% dla danych uczących, 13% – dla danych testujących, liczba reguł elementarnych wynosi 1793).

V. Podsumowanie i wnioski końcowe

W niniejszej dysertacji przedstawiono rozwiązanie dla problemu identyfikacji probabilistyczno-rozmytej bazy wiedzy oraz utworzenia na jej podstawie systemu wnioskującego.

Wkładem własnym autorki jest:

- opracowanie struktury systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy,
- wykorzystanie idei metod wyszukiwania rozmytych reguł asocjacji jako możliwość automatycznego pozyskiwania probabilistyczno-rozmytej bazy reguł,
- zmodyfikowanie algorytmów, które pozwalają bezpośrednio znaleźć wiarygodne reguły rozmyte wraz z wagami, stanowiące podstawę do wnioskowania w oparciu o budowany model wiedzy,
- opracowanie systemu decyzyjnego, który pozwoli na wybór efektywnego algorytmu generowania probabilistyczno-rozmytej bazy reguł, przy określonych parametrach modelu wiedzy,
- utworzenie w środowisku Matlab, nowego modułu narzędziowego PFIS wraz z interfejsem graficznym PFISEDIT, który może być rozszerzeniem pakietu *Fuzzy Logic Toolbox*.

Opracowanie koncepcji systemu otwiera nowe możliwości w modelowaniu zagadnień, które wymagają uwzględnienia niepewności w kategoriach probabilistycznych i rozmytych jednocześnie. Zastosowanie logiki rozmytej z regułową bazą wiedzy daje możliwość wyrażania informacji niepełnej i niepewnej w języku naturalnym, w sposób charakterystyczny dla człowieka. Zastosowanie dodatkowo prawdopodobieństw zdarzeń ujętych w kategoriach lingwistycznych, pozwala na utworzenie modelu w oparciu o wiedzę eksperta i dane empiryczne. Ograniczenie reguł elementarnych, do tych o odpowiednim poziomie wsparcia, umożliwia zmniejszenie struktury modelu wiedzy. Zwiększając wartość minimalnego wsparcia reguł elementarnych, jakość modelu wiedzy (mierzona pierwiastkiem błędu średniokwadratowego pomiędzy odpowiedzią systemu a danymi uczącymi) maleje. Jednakże, można określić optymalną strukturę modelu, dla której liczba reguł elementarnych jest najmniejsza, przy dopuszczalnym poziomie błędu dopasowania. W niniejszej pracy, za optymalną strukturę modelu wiedzy uznawano strukturę, dla zadanej wartości minimalnego wsparcia reguł, przy której system był w stanie odwzorować proces rzeczywisty na tym samym poziomie, co model wiedzy o pełnym rozkładzie prawdopodobieństwa zdarzeń rozmytych wejścia i wyjścia.

Dodatkowym atutem opracowanej koncepcji systemu jest możliwość korekcji wyników wnioskowania w oparciu o właściwy dobór operatorów, które wraz z wagami reguł decydują o wartościach stopni przynależności wyprowadzonych zbiorów rozmytych wyjścia. System umożliwia wnioskowanie w oparciu o logiczną, jak i koniunkcyjną interpretację reguł. Jednakże o jakości modelu wiedzy decyduje głównie właściwy dobór operatora t-normy, który interpretuje spójnik logiczny AND w przesłankach reguł.

Specyfika modelu wiedzy utworzonego systemu wnioskującego pozwala na zastosowanie systemu także do określenia probabilistycznych własności zmiennych rozmytych wejścia i wyjścia. Wykorzystanie w bazie wiedzy wag, stanowiących odpowiednio brzegowy oraz warunkowy rozkład prawdopodobieństwa zbioru wartości lingwistycznych zmiennych modelu, umożliwia badanie:

-
- brzegowych rozkładów prawdopodobieństwa zmiennych lingwistycznych wejść modelu,
 - warunkowych rozkładów prawdopodobieństwa zmiennej lingwistycznej wyjścia modelu,
 - łącznych rozkładów prawdopodobieństwa zmiennych lingwistycznych wejść i wyjścia modelu.

Wyznaczenie i ukazanie powyższych miar statystycznych pozwala na dogłębną interpretację charakterystyki analizowanego obiektu (procesu), w sposób przystępny dla człowieka (eksperta) oraz wpływa na łatwość w zrozumieniu otrzymanych wyników. Ma to głównie znaczenie przy projektowaniu systemów diagnostycznych oraz systemów podejmowania decyzji, gdzie możemy sugerować się nie tylko ilościowym wynikiem na wyjściu systemu, ale również informacją o prawdopodobieństwie zajścia innych możliwych zdarzeń, przy faktach określonych z wykorzystaniem języka naturalnego (za pomocą wartości lingwistycznych) lub wartości liczbowych.

Właściwe podejmowanie decyzji jest niezmiernie ważnym czynnikiem zarówno podczas pracy każdego człowieka, stabilnej pracy systemów komputerowych, jak i mechanizmów sterujących. Wybór prawidłowej opcji czy to w postaci konkretnego elementu ze zbioru możliwych elementów, czy też w postaci konkretnej wartości danego elementu ze zbioru dostępnych wartości, jest istotą różnorodnych zagadnień o charakterze decyzyjnym. W systemie sterującym podstawowym zadaniem jest dobór właściwej wielkości sterującej, która utrzyma obiekt w zadanym stanie. W systemach diagnostycznych wymaga się prawidłowego określenia wielkości stanu danego obiektu. Natomiast w systemach klasyfikacji niezbędny staje się wybór określonej klasy (kategorii), do której należy analizowane zagadnienie. W pracy zaprezentowano system decyzyjny, który wpływa na wybór efektywnego algorytmu do budowy probabilistyczno-rozmytej bazy reguł. Zatem jeżeli system z probabilistyczno-rozmytą bazą wiedzy jest w stanie zamodelować proces decyzyjny to może być także użyty do szeregu ww. zagadnień automatyki. Określenie prawdopodobieństw warunkowych zajścia zdarzeń na wyjściu systemu wpływa szczególnie na użyteczność zaproponowanego systemu w kontekście problemów biomedycznych np. do zautomatyzowanej diagnostyki medycznej w odniesieniu do różnych chorób. Jest to możliwe, gdyż proces wskazywania poprawnej diagnozy oparty jest zwykle na zasadach decyzyjnych, mimo że w przypadku różnych chorób odwołuje się do odmiennych zestawów symptomów.

Zaprezentowane w niniejszej rozprawie możliwości modelowania charakterystyk węgla wykazują możliwości aplikacyjne systemu w zadaniach klasyfikacji i diagnostyki sterowania procesami wzbogacania węgla.

Utworzenie modułu narzędziowego PFIS, w środowisku Matlab, pozwoli na dalsze testowanie systemu, z wykorzystaniem różnych danych empirycznych. Utworzona narzędzie posiada pewną przewagę w porównaniu z istniejącymi rozwiązaniami w pakiecie Fuzzy Logic Toolbox. Zbudowanie modelu wiedzy, za pomocą dostępnego w Matlabie narzędzia zwanego FIS (ang. *Fuzzy Inference System*) [fuz08], który charakteryzuje się logiczną interpretacją reguł, wymaga samodzielnego określenia przez eksperta podstawowych parametrów struktury, tj.: definicji zbiorów rozmytych, metod użytych przy wnioskowaniu oraz ilość i dokładnej budowy bazy reguł. Brak możliwości automatycznego dostrojenia struktury modelu wiedzy stanowi ograniczenie tego narzędzia. Rozważany w dysertacji system wnioskujący z probabilistyczno-rozmytą bazą wiedzy posiada "rozszerzoną", niż model lingwistyczny, strukturę, która umożliwia logiczną oraz koniunkcyjną interpretację reguł. Ponadto analizowany model wiedzy daje możliwość dostrojenia wybranych jego parametrów do danych empirycznych.

Przewagą utworzonego systemu nad narzędziem ANFIS [fuz08], jest łatwość interpretacji modelu wiedzy. W systemie ANFIS, odpowiednie metody uczenia powodują modyfikację funkcji przynależności co wpływa na to, iż użytkownik nie może wprost interpretować reguł z użyciem zrozumiałych dla niego wartości lingwistycznych. Ponadto nauczona baza reguł narzędzia ANFIS, zawiera różne wartości lingwistyczne dla każdej reguły, co również nie wpływa pozytywnie na zrozumienie i analizę badanego zagadnienia.

Wadą zaproponowanej koncepcji systemu wnioskującego jest konieczność określenia wielu parametrów wstępnych modelu (m.in. liczby rozłącznych przedziałów dyskretyzacji przestrzeni zmiennych, liczby oraz definicji zbiorów rozmytych itp.). Ponadto, otrzymane wyniki aproksymacji analizowanych zależności są ściśle zależne od ustalonych przez eksperta wartości lingwistycznych a brak dopasowania wartości lingwistycznych do danych empirycznych powoduje obniżenie jakości dopasowania modelu wiedzy do danych uczących.

Utworzony system wnioskujący jest zatem alternatywą dla systemów opartych na tradycyjnych modelach matematycznych oraz innych systemów z bazami wiedzy, w sytuacjach, gdy dane są nieścisłe i nieprecyzyjne, a zależy nam na wyrażeniu zależności z prawdopodobieństwem zajścia określonych zdarzeń w ujęciu lingwistycznym.

Mając powyższe na uwadze, wydaje się, że postawione w rozprawie cele szczegółowe zostały osiągnięte, tj.:

1. Zaproponowano strukturę systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy.
2. Zaproponowano metodę identyfikacji probabilistyczno-rozmytego modelu wiedzy, zapewniającą ograniczenie liczby reguł przy zadanej wartości błędu identyfikacji.
3. Opracowano algorytm generowania reguł z wykorzystaniem idei wyszukiwania rozmytych reguł asocjacji.
4. Dokonano implementacji zaprojektowanego systemu wnioskującego w środowisku Matlab, z uwzględnieniem plików funkcyjnych oraz interfejsu graficznego.
5. Zweryfikowano działanie systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy dla wybranych zastosowań.

Mając na uwadze spełnienie powyższych celów szczegółowych, nie stwierdzono podstaw do odrzucenia tezy, postawionej we wstępie niniejszej rozprawy doktorskiej:

Istnieje możliwość opracowania koncepcji i implementacji systemu wnioskującego z bazą wiedzy, który uwzględnia niepewność informacji, jednocześnie w kategoriach probabilistycznych i rozmytych, dla zadań modelowania i podejmowania decyzji.

Prezentowana praca przedstawia jedną koncepcję systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy i nie wyczerpuje z pewnością wszystkich możliwych reprezentacji wiedzy w ujęciu probabilistycznym i rozmytym jednocześnie. Dlatego też rozważana tematyka pozostawia miejsce do dalszych badań. Wydaje się również, że możliwymi kierunkami badań są:

- wykorzystanie pełnej idei rozmytych reguł asocjacji z uwzględnieniem eliminacji zmiennych mniej ważnych dla danego zagadnienia,
- dopracowanie metody wnioskowania w oparciu o reguły z różną ilością zmiennych w poprzedniku.

Literatura

- [agr93] Agrawal R., Imielinski T., Swami A.: Mining association rules between sets of items in large databases. In Proc. ACM SIGMOD, Washington, D.C., 1993, pp. 207-216,
- [agr94] Agrawal R., Srikant R.: Fast algorithms for mining association rules. In Proc. VLDB, 1994, pp. 487-499.
- [ahn10] Ahn C.K.: Delay-dependent state estimation for T-S fuzzy delayed Hopfield neural networks, Springer Science+Business Media B.V. 2010, published online: 06 Feb 2010.
- [alc09] Alcalá-Fdez J., Alcalá R., Gacto M.J., Herrera F.: Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms. In Fuzzy Sets and System, Vol. 160, 2009, pp. 905–921.
- [all87] Allard J.L., Kaemmerer W. F.: The goal/subgoal knowledge representation for real-time process monitoring, Proceedings IJCAI 1987, Milano, Italy, 1987, pp. 394-398.
- [au99] Au W.-H., Chan K.C.C.: FARM: A data mining system for discovering fuzzy association rules, Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE '99. 1999 IEEE International, vol.3, pp. 1217-1222.
- [bac99] Buckley J.J., Feuring T.: Fuzzy and neural: interactions and applications, Physica-Verlag, Springer-Verlag Comp., Heidelberg 1999.
- [ban09] Banaszak Z.: Modele i algorytmy sztucznej inteligencji, Politechnika Koszalińska, Wydawnictwo Uczelniane Politechniki Koszalińskiej, Koszalin, 2009.
- [bez93] Bezdek J.: Editorial, fuzzy models – What are they, and why? IEEE Trans. Fuzzy Syst., Vol. 1, No. 1, pp. 1-6.
- [bez05] Bezdek J.C., Keller J., Krisnapuram R., Pal N. R.: Fuzzy Models and Algorithms for Pattern Recognition and Image Processing, Springer, 2005.
- [bis06] Bishop C. M.: Pattern Recognition and Machine Learning, Springer, 2006.
- [bla02] Blaschke W.: Propozycja nowej formuły sprzedażowej węgla energetycznego przeznaczonego dla energetyki zawodowej. Studia, rozprawy, monografie, Nr 112. Wydawnictwo Instytutu Gospodarki Surowcami Mineralnymi i Energią PAN, Kraków 2002.
- [bla07a] Błaszczyk K.: Asocjacyjne reguły rozmyte dla modelowania szeregów czasowych, Zeszyt Naukowy, Elektryka, z. 59, I Środowiskowe Warsztaty Doktorantów PO, Opole-Jarnołtówek 2007, str. 9-10.
- [bla07b] Błaszczyk K.: Reguły asocjacji dla rozmytego modelowania szeregów czasowych, IX Międzynarodowe Warsztaty Doktoranckie, OWD, Warszawa, 2007, str. 305-310.
- [bla07c] Błaszczyk K., Ruszczak B.: Wielowymiarowe reguły asocjacji w modelowaniu tendencji rozwojowych MSP, Programowanie rozwoju regionu, Instrumentarium rozwoju, Ład społeczny, red. K Malik, Wydawnictwo Instytut Śląski, Opole 2007, str. 159-169.
- [bla08] Błaszczyk K.: Implementation of a probabilistic-fuzzy modelling system in Matlab, X Międzynarodowe Warsztaty Doktoranckie, OWD, Warszawa, 2008, str.74-78.
- [bla09] Błaszczyk K.: A probabilistic-fuzzy system applied to modeling of time series. III Środowiskowe Warsztaty Doktorantów PO, Zeszyt Naukowy Politechnika Opolska, z.62, Elektryka, Nr 329/2009, OW, Opole - Głucholaży 2009, str. 15-16.
- [bla10] Błaszczyk K.: Notes on Defining fuzzy sets in the created inference system with probabilistic-fuzzy knowledge base, IV Środowiskowe Warsztaty Doktorantów PO, Zeszyt Naukowy Politechnika Opolska, z.63, Elektryka, Nr 335/2010, OW, Opole- Pokrzywna 2010, str. 9-10.
- [bob86] Bobrowski D.: Probabilistyka w zastosowaniach technicznych, WNT, Warszawa 1986.
- [bor05] Borgelt C.: An Implementation of the FP-growth Algorithm, OSDM'05, August 21, 2005, Chicago, Illinois, USA, pp. 1-5.

- [bro06] Broekhoven E. V., Baets B. D.: Fast and accurate center of gravity defuzzification of fuzzy system outputs defined on trapezoidal fuzzy partitions, *Fuzzy Sets and Systems*, 157, 2006, pp. 904 – 918.
- [bro08] Broel-Plater B.: Wykorzystanie logiki rozmytej do sterowania ruchem ramienia mobilnego przenośnika, *Sterowanie i Automatyzacja: aktualne problemy i ich rozwiązania* (Red.: Malinowski K., Rutkowski L.), AOW EXIT, 2008, str. 563-572.
- [bub05] Bubnicki Z.: *Teoria i algorytmy sterowania*, PWN, Warszawa, 2005.
- [cha10] Chai C., Li B.: A novel association rules method based on genetic algorithm and fuzzy set strategy for Web Mining, *Journal Of Computers*, Vol. 5, No. 9, sept. 2010, pp. 1448-1455.
- [che98] Chen C.-L., Wang S.-N., Hsieh C.-T., Chang F.-Y.: Theoretical analysis of crisp-type fuzzy logic controllers using various t-norm sum-gravity inference methods, *IEEE Trans. Fuzzy Systems*, 6 (1) (1998), pp. 122–136.
- [cho96] Cho K.B., Wang B. H.: Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction, *Fuzzy Sets and Systems*, 83, 1996, pp. 325-339.
- [cic00] Cichosz P.: *Systemy uczące się*, WNT, Warszawa, 2000.
- [cie05] Cierpisz Stanisław: Parametry jakości węgla – pomiary i sterowanie. Wydawnictwo Politechniki Śląskiej, Gliwice 2005.
- [cio92] Cios J.K., Sztandera L.M.: Continuous ID3 algorithm with fuzzy entropy measures, Proc. IEEE Internat. Conf. on Fuzzy Systems, San Diego, CA, March 1992, pp. 469-476.
- [cpa09] Cpałka K.: Zagadnienie interpretowalności wiedzy i dokładności działania systemów rozmytych, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2009.
- [cor97] Cordon O., Herrera F., Peregrin A.: Applicability of the fuzzy operators in the design of fuzzy logic controllers, Vol. 86, 1, 1997, pp. 15-41.
- [cor00] Cordon O., Herrera F., Peregrin A.: Searching for basic properties obtaining robust implication operators in fuzzy control, *Fuzzy Sets and Systems*, 111, 2000, pp. 237-251.
- [cza78] Czaja-Pośpiech D., Czogała E., Pedrycz W.: Sterowanie rozmyte jako matematyczna formalizacja heurystycznego sposobu sterowania złożonymi procesami, podstawy Sterowania, Vol. 3, 1978, str. 73-84.
- [cza09] Chang W.-J., Ku C.-C., Chang W.: Fuzzy control with passivity synthesis for continuous affine Takagi-Sugeno fuzzy systems, *International Journal of Intelligent Computing and Cybernetics*, Vol. 2 No. 2, 2009, pp. 386-408.
- [cze04] Czekalski P.: Systemy ewolucyjno-rozmyte z parametrycznymi konkluzjami w regułach IF-THEN, Praca doktorska, Politechnika Śląska, Wydział Automatyki, Elektrotechniki i Informatyki, Gliwice, 2004.
- [cze08] Czemplik A.: Modele dynamiki układów fizycznych dla inżynierów, Wydawnictwo Naukowo-Techniczne WNT, Warszawa 2008.
- [che03] Chen J.C., Susanto V.: Fuzzy logic based in-Process tool-wear monitoring system in face milling operations, *The International Journal of Advanced Manufacturing Technology*, Vol. 21, No. 3, Springer-Verlag London, 2003, pp. 186-192.
- [czo85] Czogała E., Pedrycz W.: Elementy i metody teorii zbiorów rozmytych, PWN, Warszawa 1985.
- [del00] Delgado M., Sánchez D., Vila M. A.: Fuzzy cardinality based evaluation of quantified sentences, *Inter. J. of Approximate Reasoning*, 23:2000, pp. 23–66.
- [det01] Detyniecki M.: Praca doktorska: Mathematical aggregation operators and their application to video querying, promotorzy: B. Bouchon-Meunier, R. Yager, [<http://www.cs.berkeley.edu/~marcin/agop.pdf>], data odczytu: kwiecień 2010].
- [din03] Ding Y.-S., Ying H., Shao S.-H.: Typical Takagi–Sugeno PI and PD fuzzy controllers: analytical structures and stability analysis, Elsevier, *Information Sciences* 151 (2003), pp. 245–262.
- [don08] Donchenko V.: Fuzzy Sets: Abstraction Axiom, Statistical Interpretation, Observations Of Fuzzy Sets, *International Journal Information Theories & Applications*, Vol.13, 2008, pp. 233-238.

- [dri93] Driankov D., Hellendorf H., Reinfrank M.: An introduction to fuzzy control. Springer-Verlag, Berlin, Heidelberg 1993.
- [duc00] Rutkowska D., Rutkowski L.: Systemy rozmyte i rozmyto-neuronowe, (w:) Biocybernetyka i Inżynieria Biomedyczna 2000, Sieci Neuronowe T6, (red.) Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R., Akademicka Oficyna Wydawnicza EXIT, Warszawa 2000.
- [dub84] Dubois D., Prade H.: A theorem on implication functions defined from triangular norms, Stochastica, Vol. VIII, No. 3, 1984, pp. 267-279.
- [dub98] Dubois D., Prade H.: An introduction to fuzzy systems, Elsevier, Clinica Chimica Acta 270 (1998), pp. 3-29.
- [dub03] Dubois D., Prade H., Sudkamp T.: A Discussion of Indices for the Evaluation of Fuzzy Associations in Relational Databases, Springer Berlin / Heidelberg, 2003, pp.111-118.
- [dub06] Dubois D.: Hüllermeier E., Prade H., A systematic approach to the assessment of fuzzy association rules, Data Mining and Knowledge Discovery, Vol. 13 , Issue 2, 2006.
- [fay96] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. In AI Magazine, 1996, pp. 37-54.
- [fil93] Filev D.P., Yager R.R.: An adaptive approach to defuzzification based on level sets, Fuzzy sets and systems, 53, 1993, pp. 355-360.
- [fod91] Fodor J.C.: On fuzzy implication operators, Fuzzy Sets and Systems, Vol. 42, Issue 3, 15 August 1991, pp. 293-300.
- [fod94] Fodor J.C., Roubens M.: Fuzzy preference modelling and multicriteria decision support, Kluwer Academic Publishers, Dordrecht, 1994.
- [fra92] Frawley, W., Piatetsky-Shapiro, G., Matheus, C.: Knowledge discovery in databases: an overview. In AI Magazine, Vol 13 (1992), No. 3, pp. 57-70.
- [fuk00] Fukunaga K.: Introduction to statistical pattern recognition, Academic Press, 1990.
- [fuz08] Fuzzy Logic Toolbox 2, User's Guide, The MathWorks, Release 2008a.
- [gen97] Genther H., Runkler T., Glesner M.: Defuzzification based on fuzzy clustering, Third IEEE Conference on fuzzy systems, 1994, pp. 1646-1648.
- [gon96] Gonzalez A., Herrera F.: Multi-stage genetic fuzzy systems based on the iterative rule learning approach, DECSAI, 1996.
- [gor02] Gorzałczany M.B.: Computational intelligence systems and applications: neuro-fuzzy and fuzzy neural synergisms, Physica-Verlag, Springer-Verlag Comp., Heidelberg , 2002.
- [gor05] Gornig A.: Zastosowanie metod klasyfikacji danych w analizie jakości węgla. Praca doktorska, Wydział Górnictwa i Geologii, Politechnika Śląska, Gliwice, 2005.
- [got03] Gottwald S.: Fuzzy Relation Equations and Approximate Solutions. Some Recent Results, Proceedings of the 3rd Conference of the European Society for Fuzzy Logic and Technology, (Eds.) Wagenknecht M. and Hampel R, Zittau, Germany, September 10-12, 2003.
- [gye00] Gyenesi A.: Mining weighted association rules for fuzzy quantitative items, Zighed D.A., Komorowski J., Żytkow J. (Eds.), PKDD 2000, LNAI 1910, Springer-Verlag Berlin Heidelberg 2000, pp. 416-423.
- [háj06] Hájek P.: What is mathematical fuzzy logic, Fuzzy Sets and Systems 157, 2006, pp. 597 – 603.
- [han95] Han J., Fu Y.: Discovery of multiple level association rules from large databases. In: Proceedings of the 21st International Conference on Very Large Data Bases, Sept. 95, pp. 420-431.
- [han00a] Han J., Pei J., Mortazavi-Asl B., Chen Q., Dayal U., Hsu M.: FreeSpan: frequent pattern-projected sequential pattern mining. Conf. on Knowledge Discovery in Data. Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining. Boston, Massachusetts, USA, 2000, pp. 355 – 359.

- [han00b] J. Han, H. Pei, and Y. Yin.: Mining Frequent Patterns without Candidate Generation. In: Proc. Conf. on the Management of Data (SIGMOD'00, Dallas, TX), ACM Press, New York, NY, USA 2000.
- [han01] Han J., Kamber M.: Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, 2001.
- [he06] He Y., Tang Y., Zhang Y.-G., Sunderraman R.: Adaptive Fuzzy Association Rule mining for effective decision support in biomedical applications, Int. J. Data Mining and Bioinformatics, Vol. 1, No. 1, 2006.
- [hel97] Hellendorf H., Driankov D.: Fuzzy model identification. Selected approaches, Springer-Verlag, Berlin, Heidelberg, 1997.
- [hir77] Hirota K.: Concepts of probabilistic sets, Proc. IEEE Conf. Decision and Control 1977, pp. 1361-1366.
- [hol93] Holmes G., Cunningham S.J.: Using data mining to support the construction and maintenance of expert systems. In Proc Artificial Neural Networks and Expert Systems, pages 156-159, Dunedin, New Zealand, 1993.
- [ham00] Hammouda K. M., Karray F.: A Comparative Study of Data Clustering Techniques, SYDE 625: Tools of Intelligent Systems Design. Course Project.
- [hon01] Hong T.-P., Kuo C.-S., Chi S.-C.: Trade-off between computation time and number of rules for fuzzy mining from quantitative data, Inter. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 9, No. 5(2001), pp. 587-604.
- [hor92] Horikawa S., Furuhashi T., Uchikawa Y.: On fuzzy modeling using fuzzy neural networks with the backpropagation algorithm, IEEE Trans. on Neural Networks, Sep 1992, Volume 3, Issue 5, pp. 801-806.
- [hoy93] Hoyo T., Terano T., Masui S.: Design of quasi-optimal fuzzy controller by fuzzy dynamic programming, Proceedings of Second IEEE Intern. Conf. on Fuzzy Systems, San Francisco, CA, USA, Vol. 2, pp. 1253-1258.
- [hül05] Hüllermeier E.: Fuzzy methods in machine learning and data mining: status and prospects. In Fuzzy Sets and System, Vol. 156, Issue 3, 2005, pp. 387-406.
- [ian98] Iancu I.: A Method For Constructing T-Norms, Korean J. Comput. & Appl. Math. Vol. 5 (1998), No. 2, pp. 407 - 414.
- [ich96] Ichihashi H., Shirai T., Nagasaka K., Miyoshi T.: Neuro-fuzzy ID3: a method of inducing fuzzy decision trees with linear programming for maximizing entropy and an algebraic method for incremental learning, Fuzzy Sets and Systems, Vol 81, Issue 1, 8 July 1996, pp. 157-167.
- [int09] Intan R., Yuliana O. Y., Handoko A.: Mining Fuzzy Multidimensional Association Rules Using Fuzzy Decision Tree Induction Approach, (IJCNS) International Journal of Computer and Network Security, Vol. 1, No. 2, November 2009, pp. 60-69.
- [ish93] Ishibuchi H., Fujioka R., Tanaka H.: Neural networks that learn from fuzzy if-then rules, Fuzzy Systems, IEEE, May 1993, Volume 1, Issue 2, pp. 85 - 97.
- [jan93] Jang J.-S. R.: ANFIS: Adaptive-Network-Based Fuzzy Inference System, IEEE Transactions On Systems, Man, And Cybernetics, Vol. 23, No. 3, May/june 1993, pp. 665-685.
- [jan95] Jang J.-S. R., Sun C.-T., J.-S. R.: Neuro-fuzzy modeling and control, Proceedings IEEE, 83 (3), 1995, pp. 378-406.
- [jan02] Janiszowski K., Identyfikacja modeli parametrycznych w przykładach, EXIT, 2002.
- [jan07] Januszewski E.: Roczniki filozoficzne, Logiczne i filozoficzne problemy związane z logiką rozmytą, Towarzystwo Wydawnicze KUL, T. LV, RF 55, nr 1, 2007.
- [jay09] Jayaram B., Mesiar R.: On special fuzzy implications, Fuzzy Sets and Systems Volume: 160, Issue: 14, July 16, 2009, pp. 2063-2085.
- [jin03] Jin Y.: Advanced fuzzy systems design and applications, Studies in fuzziness and soft computing, Physica-Verlag Heidelberg 2003.

- [jon90] Jones R.D. et al.: Function approximation and time series prediction with neural networks. 1990 IJCNN International Joint Conference on Neural Networks, 1990, pp. 649-665.
- [kac86] Kacprzyk J.: Zbiory rozmyte w analizie systemowej, PWN, Warszawa 1986.
- [kac01] Kacprzyk J.: Wieloetapowe sterowanie rozmyte, Wydawnictwo Naukowo-Techniczne, Warszawa 2001.
- [kac81] Kaczorek T.: Teoria sterowania, PWN, Warszawa, 1981.
- [kac09] Kaczorek T., Dzieliński A., Dąbrowski W., Łopata R.: Podstawy teorii sterowania, Wydawnictwo WNT, Warszawa, 2009.
- [kair] Strategiczny Program Badawczy, Grupa Robocza KAiR PAN, [http://www.kair.pan.pl/images/stories/pliki/pdf/propozycjagrroboczej9.pdf], data odczytu: maj 2011]
- [kan08] Kandyba A., Kalus M., Piasecki A., Skoczkowski T.: Regulacja w logice rozmytej temperatury stalowej rury w procesie nagrzewania i wyżarzania oporowego, Sterowanie i Automatyzacja: aktualne problemy i ich rozwiązania (Red.: Malinowski K., Rutkowski L.), AOW EXIT, 2008, str.406-416.
- [kay00] Kaymak U., Setnes M.: Extended Fuzzy Clustering Algorithms, [http://publishing.eur.nl/ir/repub/asset/57/erimrs20001123094510.pdf, 13.08.2010].
- [kay02] Kaya M., Alhajj R., Polat F., Arslan A.: Efficient Automated Mining of Fuzzy Association Rules, R. Cicchetti et al. (Eds.), DEXA 2002, LNCS 2453, 2002, pp. 133-142.
- [kay05] Kaya M., Alhajj R.: Genetic algorithm based framework for mining fuzzy association rules, Fuzzy Sets and Systems 152, 2005, pp. 587–601.
- [kaz09] Kazienko P.: Mining indirect association rules for web recommendation, Int. J. Appl. Math. Comput. Sci., 2009, Vol. 19, No. 1, pp. 165–186.
- [kic78] Kickert W.J.M., Mamdani E. H.: Analysis of a fuzzy controller, Fuzzy Sets an Systems, Vol. 1, pp. 29-44.
- [kle04] Klement E. P., Mesiar R., Pap E.: Triangular norms. Position paper I: basic analytical and algebraic properties, Fuzzy Sets and Systems 143, 2004, pp. 5–26.
- [kli95] Klir G. J., Yuan B.: Fuzzy Sets and Fuzzy Logic. NJ: Prentice-Hall, 1995.
- [kon07] Kononenko I., Kukar M.: Machine Learning and Data Mining, Horwood Publ., 2007.
- [kor94] Korbicz J., Obuchowicz A.: Sztuczne sieci neuronowe. Podstawy i zastosowanie, Akademicka Oficyna Wydawnicza, Warszawa 1994.
- [kor02] Korbicz J., Kościelny J. M., Kowalcuk Z., Cholewa W., Diagnostyka procesów: modele, metody sztucznej inteligencji, zastosowania, Lubuskie Towarzystwo Naukowe w Zielonej Górze, Warszawa, Wydaw-a Naukowo-Techniczne, 2002.
- [kor08] Koronacki J., Cwik J.: Statystyczne systemy uczące się. Warszawa, 2008.
- [kos92a] Kosko B.: Fuzzy systems as universal approximators, Proceedings IEEE International Conference on Fuzzy Systems, San Diego, 1992, pp. 1153-1162.
- [kos92b] Kosko B.: Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence, Prentice-Hall, Englewood Cliffs, NJ, 1992.
- [kow08] Kowalczyk A.: Rachunek prawdopodobieństwa i statystyka matematyczna. Przykłady rachunkowe do wykładu, Politechnika Rzeszowska 2008.
- [kra04] Krawiec K., Stefanowski J.: Uczenie maszynowe i sieci neuronowe, Wydawnictwo Politechniki Poznańskiej, Poznań, 2 wydanie, 2004.
- [kry97] Krysicki W., Bartos J., i inni.: Rachunek prawdopodobieństwa i statystyka matematyczna część I, Wydawnictwo Naukowe PWN, Warszawa 1997.
- [kuo98] Kuok C. M., Fu A. W., Wong M. H.: Mining fuzzy association rules in databases, SIGMOD Record, 17(1), 1998, pp. 41-46.
- [kus83] Kushner H.: Wprowadzenie do teorii sterowania stochastycznego, PWN, Warszawa, 1983.

- [kwi07] Kwiatkowska A., M.: Systemy wspomagania decyzji. Jak korzystać z WIEDZY i informacji, Wydawnictwo Naukowe PWN/MIKOM, 2007.
- [lar06] Larose D.T.: Odkrywanie wiedzy z danych, PWN, 2006.
- [lee99] Leekwijck W.V., Kerre E.E.: Deffuzzification: criteria and classification, Fuzzy Sets and Systems, 108, 1999, pp. 159-178.
- [lee03] Lee C.W., Shin Y.C.: Construction of fuzzy systems using least-squares method and genetic algorithm, Fuzzy Sets and Systems, 137, pp. 297 - 323.
- [lig06] Ligęza A.: Logical foundations for rule-based systems, Springer-Verlag Berlin Heidelberg 2005.
- [liu04] Liu P., Li H.: Fuzzy neural network theory and application, World Scientific Publishing, Singapore, 2004.
- [liu07] Liu X.: Parameterized defuzzification with maximum entropy weighting function - another view of weighting function expectation method, Math. and Comp. Modelling, 45, 2007, pp. 177-188.
- [lu05] Lu Y.-H., Yeh F.-H., Li C.-L., Wu M.-T.: Study of using ANFIS to the prediction in the bore-expanding process, Int. J. Adv. Manuf. Technol. (2005) 26, pp. 544–551.
- [łes99] Łęski J., Czogała E.: A new fuzzy inference system based on artificial neural network and its application, in: Zadeh L., Kacprzyk A. (Eds.), Computing with words in information/intelligent systems. Volume 2, Physica-Verlag, Springer-Verlag Com., Heidelberg, New York, 1999, pp.75–94.
- [łes08] Łęski, J.: Systemy neuronowo-rozmyte. WNT, Warszawa 2008.
- [luk20] Łukasiewicz J.: O logice trójwartościowej (On three-valued logic), Ruch Filozoficzny 5, 1920, str. 169-171.
- [mab97] Mabuchi S.: An interpretation of membership functions and the properties of general probabilistic operators as fuzzy set operators – Part I, Fuzzy Sets and Systems 89, 1997, pp. 69-76.
- [mam74] Mamdani E.H.: Application of fuzzy algorithms for simple dynamic plant, Proc. IEEE 121, 2, 1974, pp. 1585–1588.
- [mam75] Mamdani E.H., Assilian S.: An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies. vol. 7, 1975, pp. 1-13.
- [mań71] Mańczak K.: Metody identyfikacji wielowymiarowych obiektów sterowania, Wydawnictwo Naukowo-Techniczne, Warszawa 1971.
- [mań83] Mańczak K., Nahorski Z.: Komputerowa identyfikacja obiektów dynamicznych, Państwowe Wydawnictwo Naukowe, Warszawa 1983.
- [mal02] Malinowski G.: Many-valued logics, Oxford University Press, New York 2002.
- [mas04] Mastorakis N.E.: General fuzzy systems as extensions of the Takagi-Sugeno methodology, Wseas Transactions on Systems, Issue 2, Volume 3, April 2004, p. 795-801.
- [mat] MATLAB uniwersalne środowisko obliczeń naukowo technicznych, [\[http://www.ont.com.pl/matlab.php?disp=kat&id=117\]](http://www.ont.com.pl/matlab.php?disp=kat&id=117), data odczytu: marzec 2010].
- [mat02] Mata J., Alvarez J.L., Riquelme J.C.: An evolutionary algorithm to discover numeric association rules, SAC '2002, Madrid, Spain, pp. 590-594.
- [men05] Mennis J., Liu J.W.: Mining Association Rules in Spatio-Temporal Data: An Analysis of Urban Socioeconomic and Land Cover Change, Transactions in GIS, 2005, 9(1), pp. 5–17.
- [mil97] Miller R.J., Yang Y.: Association Rules over Interval Data, ACM, Vol. 26 , Iss. 2 , June 1997, pp. 452 - 461.
- [mit97] Mitchell T.M.: Machine Learning, McGraw Hill, 1997.
- [miy08] Miyamoto S., Ichihashi H., Honda K.: Algorithms for Fuzzy Clustering, Springer Verlag 2008.
- [mor] Morzy T., Morzy M., Leśniewska A.: Wykłady z przedmiotu eksploracja danych, Studia Informatyczne, [\[http://wazniak.mimuw.edu.pl/index.php?title=Eksploracja_danych\]](http://wazniak.mimuw.edu.pl/index.php?title=Eksploracja_danych), data odczytu: październik 2010].

- [mor05] Morzy M., Oracle Data Mining – odkrywanie wiedzy w dużych wolumenach danych, XI Konferencja PLOUG Kościelisko, Październik 2005.
- [nał00] Nałęcz T., Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R.: Biocybernetyka i inżynieria biomedyczna 2000. Sieci neuronowe. T 6. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2000.
- [nie00] Niederliński A.: Regułowe systemy ekspertowe, Wydawnictwo Pracownia Komputerowa Jacka Skalmierskiego PKJS, 2000.
- [nie06] Niederliński A.: Regułowo-modelowe systemy ekspertowe rmse, Wydawnictwo Pracownia Komputerowa Jacka Skalmierskiego PKJS, 2006.
- [now09] Nowicki R.K.: Rozmyte systemy decyzyjne w zadaniach z ograniczoną wiedzą, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2009.
- [noz97] Nozaki K., Ishibuchi H., Tanaka H.: A Simple but Powerful Heuristic Method for Generating Fuzzy Rules From Numerical Data, *Fuzzy Sets and Systems*, Vol. 86, No. 3, 1997, pp. 251-270.
- [nur09] Nurmi H.: Probability and fuzziness - echoes from 30 years back, in: Views on Fuzzy Sets and Systems from Different Perspectives, (Ed.) Seising R., Springer-Verlag Berlin Heidelberg, 2009, pp. 161-174.
- [oh07] Oh S.-K., Pedrycz W., Park K.-J.: Identification of fuzzy systems by means of genetic optimization and data granulation, *Journal of Intelligent & Fuzzy Systems* 18, 2007, pp. 31-41.
- [ols07] Olson D. L., Li Y.: Mining fuzzy weighted association rules, Proceedings of the 40th Hawaii International Conference on System Sciences, IEEE 2007, pp. 53-62.
- [osb86] Osborne R. L., Gonzalez A.J., Weeks C.A.: First years experience with on-line generator diagnostics, Proceddings American Control Conference, Chicago, IL, 1986.
- [oso96] Osowski S.: Sieci neuronowe w ujęciu algorytmicznym, WNT, Warszawa 1996.
- [ost99] Ostasiewicz S., Rusnak Z., Siedlecka U.: Statystyka. Elementy teorii i zadania. Wydawnictwo Akademii Ekonomicznej im. O. Langego we Wrocławiu, Wrocław 1999.
- [paj07] Paja W.: Rozprawa doktorska pt. Budowa optymalnych modeli uczenia na podstawie wtórnego źródła wiedzy, Promotor Z.S. Hippe, Wyższa Szkoła Informatyki i Zarządzania w Rzeszowie, Kraków 2007.
- [pal01] Palit A.K., Babuška R.: Efficient training algorithm for Takagi-Sugeno type neuro-fuzzy network, In Proceedings of FUZZ-IEEE'2001, pp.1367-1371.
- [par91] Park J. Sandberg I.W.: Universal approximation using radial-basis-function Networks, *Neural Computation*, 3, 1991, pp. 246-257.
- [pat02] Patel A.V., Mohan B. M.: Some numerical aspects of center of area defuzzification method, *Fuzzy sets and systems*, 132, 2002, pp. 401-409.
- [pie04] Pieczyński A.: Rozprawa habilitacyjna pt. Reprezentacja wiedzy w diagnostycznym systemie ekspertowym, Lubuskie Towarzystwo Naukowe w Zielonej Górze, 2004.
- [ped84] Pedrycz W.: An identification algorithm in fuzzy relational systems, *Fuzzy Sets and Systems*, Vol. 13, 1984, pp. 153-167.
- [ped93] Pedrycz W.: Fuzzy control and fuzzy systems, John Wiley and Sons, New York 1993.
- [per06] Perfilieva I., Lehmke S.: Correct models of fuzzy IF-THEN rules are continuous, *Fuzzy Sets and Systems*, Vol. 157, Issue 24, 16 December 2006, pp. 3188-3197.
- [pie03] Piegał A.: Modelowanie i sterowanie rozmyte, EXIT, Warszawa 2003.
- [pis09] Pisz I., Błaszczyk K.: Estimation of project realization costs with the use of fuzzy sets, Monografia drukowana w ramach II International Interdisciplinary Technical Conference of Young Scientists "InterTech 2009", 20-22 May 2009, Poznań.
- [plu00] Plucińska A., Pluciński E.: Probabilistyka, WNT, Warszawa 2000.

- [pra07] Pradera A., Trillas E., Guadarrama S., Renedo E.: On Fuzzy Set Theories, Studies in Fuzziness and Soft Computing, Springer Berlin Heidelberg, Vol. 215, 2007, pp. 15-47.
- [pri05] Priyono A., Ridwan M., Alias A. J., Rahmat R. A. O. K., Hassan A. ,Ali M. A. M.: Generation of fuzzy rules with subtractive clustering, Journal Technology, 43(D), 2005, pp. 143-153.
- [qui86] Quinlan J.R.: Induction of decision tree, Mach. Learning, 1 (1986), pp. 81-106.
- [qui90] Quinlan J.R.: Decision trees and decision making, IEEE Trans. Systems Man Cybernet, 20 (1990), pp. 339-346.
- [roj00] Rojek R., Bartęcki K., Korniak J.: Zastosowanie sztucznych sieci neuronowych i logiki rozmytej w automatyce, Praca zbiorowa pod redakcją R. Rojka, Skrypt Politechniki Opolskiej nr 234, Opole 2000.
- [rud10] Rudnik K., Walaszek-Babiszewska A.: Rozmyty system wnioskujący o modelu bazującym na regułach asocjacji, Zarządzanie przedsiębiorstwem, Nr 2 (2010), str. 50-60.
- [rud11] Rudnik K.: Conception and implementation of the inference system with probabilistic-fuzzy knowledge base, V Środowiskowe Warsztaty Doktorantów PO, Zeszyt Naukowy Politechnika Opolska, 2011.
- [rus69] Ruspini E.H.: A new approach to clustering, Inform. Control, 15(1), 1969, pp. 22-32.
- [rut97a] Rutkowska D.: Inteligentne systemy obliczeniowe, Algorytmy genetyczne i sieci neuronowe w systemach rozmytych, Akad. Oficyna Wyd., Warszawa, 1997.
- [rut97b] Rutkowska D., Piliński M., Rutkowski L.: Sieci neuronowe, algorytmy genetyczne i systemy rozmyte, PWN, Warszawa-Łódź 1997.
- [rut06] Rutkowski L.: Metody i techniki sztucznej inteligencji. Wydawnictwo naukowe PWN, Warszawa 2006.
- [ryk] Rykaczewski K.: Systemy rozmyte i ich zastosowania, [<http://gold.uwb.edu.pl/~akuzel/pdfy/duszek.pdf>, data odczytu: grudzień 2009].
- [sav95] Savasere A., Omiecinski E., Navathe S.: An Efficient Algorithm for Mining Association Rules in large database, Proceedings of the 21st VLDB Conference Zurich, Switzerland, 1995, pp. 432-444.
- [shu08] Shu H., Zhu X., Dai S.: Mining association rules in geographical spatio-temporal data, ISPRS Congress Beijing 2008, Vol. XXXVII, pp. 225-228.
- [smi80] Smith S.F.: A learning system based on genetic adaptive algorithms, Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA, 1980.
- [sik07] Sikora M.: Konstruktywna indukcja i rozmywanie dla poprawy jakość reguł z liniowymi konkluzjami, Mat. konf. Krajowej Konferencji Naukowej Technologie Przetwarzania danych, KKNTPD 2007, Poznań, 24-26 wrzesień 2007.
- [sim] Simiński R.: Generowanie reguł minimalnych, [http://prac.us.edu.pl/~siminski/se/se_04.pdf, data odczytu: sierpień 2010].
- [sri95] Srikant R., Agrawal R.: Mining Generalized Association Rules, In Proc. of the 21st Int. Conference on Very Large Databases, Zurich, Switzerland, 1995.
- [sri96] Srikant R., Agrawal R.: Mining Quantitative Association Rules in Large Relational Tables, In Proc. of the ACM SIGMOD Conference on Management of Data, Montreal, Canada, 1996.
- [su10] Su C.-L., Li P.: Adaptive predictive functional control based on Takagi-Sugeno model and its application to pH process, Journal of Central South University of Technology, Vol. 17, Nr 2, April, 2010, pp. 363-371.
- [sug88] Sugeno M., Kang G.T.: Structure identification of a fuzzy model, Fuzzy Sets and Systems, 28, 1988, pp. 15-33.
- [sug93] Sugeno M., Yasukawa T.: A Fuzzy-Logic Based Approach to Qualitative Modeling. IEEE Trans. Fuzzy Systems, 1 (1), 1993, pp. 7-31.

- [sza03] Szabatin J.: Przetwarzanie sygnałów, 2003, [<http://www.ise.pw.edu.pl/~szabatin/>, data odczytu: kwiecień 2011].
- [świ09] Świątek J.: Wybrane zagadnienia identyfikacji statycznych systemów złożonych, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2009.
- [tak85] Takagi T., Sugeno M.: Fuzzy identification of systems and its application to modeling and control, IEEE Trans. Systems, Man and Cybernetics, 15(1), 1985, pp. 116-132.
- [tan05] Tang Y., Sun F., Sun Z.: Neuro-fuzzy system modeling based on automatic fuzzy, Journal of Control Theory and Applications, 2 (2005), pp. 121 - 130.
- [tan06] Tan P.-N., Steinbach M., Kumar V.: Introduction to data mining, Addison-Wesley Companion Book Site, 2006.
- [tat06] Tatjewski, P. ; Ławryńczuk, M.: Soft computing in model-based predictive control, International Journal of Applied Mathematics and Computer Science, 2006, Vol. 16, No 1, pp. 7-26.
- [tso01] Tsoukatos I., Gunopoulos D.: Efficient Mining of Spatiotemporal Patterns. In Proc. of SSTD, 2001, pp. 425–442.
- [uma93] Umano M.: Generation of fuzzy decision trees by extended algorithm and its application to diagnosis by analyzing gas in oil, Proc. 4th Intelligent FA Symposium, 1993, pp. 201-204.
- [vis09] Visalakshi N.K., Thangavel K.: Distributed Data Clustering: A Comparative Analysis, in A. Abraham et al. (Eds.), Foundations of Comput. Intel. Vol. 6, SCI 206, pp. 371–397, Springer-Verlag Berlin Heidelberg 2009.
- [wb97] Walaszek-Babiszewska A. (Kierownik projektu): Projekt badawczy, Reprezentacja wiedzy i jej modele w zakresie opróbkowania materiału dla systemów kontroli procesów przeróbki kopalin, Nr proj. 9T12A 006 10, Politechnika Śląska, Wydział Górnictwa i Geologii, Termin realizacji: 1996-1997.
- [wb05] Walaszek-Babiszewska A.: Measurements and expert knowledge for creating a fuzzy representation of stochastic systems, in: Methods of artificial intelligence, (Eds Burczyński T., Cholewa W., Moczulski W.), AI-METH Series, Gliwice, 2005.
- [wb07] Walaszek-Babiszewska A.: Construction of Fuzzy Models Using Probability Measures of Fuzzy Events, in Proc. 13th IEEE Internat. Conf. on Methods and Models in Automation and Robotics, MMAR 2007, Szczecin, Poland, pp. 661-666.
- [wb08a] Walaszek-Babiszewska A.: Fuzzy Knowledge Representation Using Probability Measures of Fuzzy Events, Automation and Robotics, Ed. Juan Manuel Ramos Arreguin, May 2008.
- [wb08b] Walaszek-Babiszewska A., Czabak A., Błaszczyk K.: Rozmyte modele dyskretnych procesów stochastycznych w opisie procesów technologicznych, Automation 2008, PAR, luty (2) 2008, str. 642-651.
- [wb08c] Walaszek-Babiszewska A., Błaszczyk K., Czabak A.: Budowa rozmytych modeli procesów stochastycznych przy użyciu reguł asocjacji, Sterowanie i automatyzacja: aktualne problemy i ich rozwiązania, red. Malinowski K., Rutkowski L., EXIT, Warszawa 2008, str. 246-256.
- [wb09] Walaszek-Babiszewska A., Błaszczyk K.: A modified Apriori algorithm to generate rules for inference system with probabilistic-fuzzy knowledge base, 7th Workshop on Advanced Control and Diagnosis 19-20 November 2009, Zielona Góra, CD-ROOM.
- [wb10] Walaszek-Babiszewska A.: Modelowanie rozmyte systemów stochastycznych. Teoria, modele, bazy wiedzy, Oficyna Wydawnicza Politechnika Opolska, 2010.
- [wan92] Wang L.-X., Mendel J.M.: Generating fuzzy rules by learning from examples, IEEE Transactions on Systems, Man, and Cybernetics 22, 6 (1992), pp. 1414-1427.
- [wan98] Wang L.-X.: A course in fuzzy systems and control. Prentice-Hall, New York 1998.
- [wan00] Wang X., Chen B., Qian G., Ye F.: On the optimization of fuzzy decision trees, Fuzzy Sets and Systems 112 (2000), pp. 117-125.
- [wan09] Wang C.-H., Pang C.-T.: Finding Fuzzy Association Rules Using FWFP-Growth with Linguistic Supports and Confidences, World Academy of Science, Engineering and Technology, 53, 2009, pp. 1139-1147.

- [web92] Weber R.: Fuzzy ID3: a class of methods for automatic knowledge acquisition, Proc. 2nd Internat. Conf. on Fuzzy Logic and Neural Networks (1992), pp. 265-268.
- [wic08] Wiciak M.: Elementy probabilistyczne w zadaniach, Podręcznik dla studentów wyższych szkół technicznych, Politechnika Krakowska, Kraków 2008.
- [wie09] Wierzchoń S.T.: Elementy teorii zbiorów rozmytych, <http://www.ipipan.waw.pl/~stw/esi/sterowniki.pdf>, data odczytu: kwiecień 2009].
- [yag80] Yager R.R.: An approach to inference in approximate reasoning, International Journal on Man-Machine Studies, vol.13, 1980, pp. 323-338.
- [yag95] Yager R.R., Filev D. P.: Podstawy modelowania i sterowania rozmytego, WNT, Warszawa 1995.
- [yua95] Yuan Y. Shaw M. J.: Induction of fuzzy decision trees, Fuzzy Sets and Systems, 69(2), 1995, pp. 125-139.
- [zim96] Zimmermann H.J.: Fuzzy set theory, Boston: Kluwer, 1996.
- [zim05] Zimmer A., Englot A.: Identyfikacja obiektów i sygnałów. Politechnika Krakowska , Kraków 2005.
- [zad65] Zadeh L.A.: Fuzzy sets. Inform. Contr., 1965 vol. 8, pp. 338 – 353.
- [zad68] Zadeh L.A.: Probability measures of fuzzy events, Journal of Mathematical Analysis and Applications, vol. 23, 2, 1968, pp. 421-427.
- [zad71] Zadeh L.A.: Toward a theory of fuzzy systems. in: Aspects of Network and System Theory, R.E. Kalman, N. DeClaris (Eds.), Rinehart & Winston, New York, 1971, pp. 469-490.
- [zad73] Zadeh L.A.: Outline of a new approach to the analysis of complex system and decision processes, IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-3, pp. 28-44.
- [zad75a] Zadeh L.A.: The concept of a linguistic variable and its application to approximate reasoning Part 1, Information Sciences 8(1975), pp. 199 – 249.
- [zad75b] Zadeh, L.A.: The concept of a linguistic variable and its application to approximate Part 2, Information Sciences 8(1975), pp. 301-357.
- [zad75c] Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning, Part 3, Information Sciences 9(1975), pp. 43-80.
- [zad79] Zadeh L.A.: A theory of approximate reasoning. In: Machine Intelligence, Hayes J.E., Michie D. and Mikulich L.I. (Eds.), Vol. 9, New York, 1979, pp.149-194.
- [zak97] Zaki M.J., Parthasarathy S., Ogihsara M., Li W.: New Algorithms for Fast Discovery of Association Rules, in Proc. of 3rd International Conference on Knowledge Discovery and Data Mining (KDD-97), Newport Beach, California, USA, 1997.
- [zha99] Zhang W.: Mining fuzzy quantitative association rules, IEEE 1999, pp. 99-102.
- [zhi05] Zhi-Chao L., Pi-Lian H., Ming L.: A high efficient AprioriTid algorithm for mining association rule, Machine Learning and Cybernetics, 2005, Proceedings of 2005 International Conference, Vol. 3, Issue, 18-21 Aug. 2005, pp. 1812 – 1815.

Spis rysunków

Rys. II-1. Identyfikacja w rozmytych systemach z bazami wiedzy (na podstawie [hel97]).....	11
Rys. II-2. Interpretacja graficzna wybranych parametrów zbioru rozmytego	18
Rys. II-3. Funkcja przynależności zbioru rozmytego A oraz jego dopełnienia $\neg A$	19
Rys. II-4. Funkcje przynależności przecięcia zbiorów rozmytych A i B z uwzględnieniem różnych operatorów t -normy	22
Rys. II-5. Funkcje przynależności sumy zbiorów rozmytych A i B z uwzględnieniem różnych operatorów s -normy	22
Rys. II-6. Schemat rozmytego systemu wnioskującego (na podstawie [rut06], [lęs08], [kac01])	28
Rys. II-7. Struktura systemu neuronowo-rozmytego (na podstawie [duc00]).....	34
Rys. II-8. Typy anomalii regułowej reprezentacji wiedzy (na podstawie [lig08],[paj07])	35
Rys. II-9. Przykład generowania reguły metodą Wanga-Mendela.....	44
Rys. II-10. Rozkład danych wejście-wyjście z uwzględnieniem szablonów: a) siatka rozmyta, b) siatka nieroźmyta (na podstawie [yag95]).....	46
Rys. III-1. Schemat struktury systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy oraz jego powiązanie z otoczeniem (na podstawie [rud10]).....	60
Rys. III-2. Przykład definicji zbiorów rozmytych dla 10 rozłącznych przedziałów wartości zmiennej (na podstawie [rud10])	62
Rys. III-3. Definicje zbiorów rozmytych z uwzględnieniem znormalizowanych funkcji przynależności oraz wartości stopni przynależności.....	63
Rys. III-4. Liczba elementarnych reguł w zależności od wartości minimalnego wsparcia (min w).....	64
Rys. III-5. Wartości błędów RMSE dla danych uczących i testujących w zależności od parametru minimalnego wsparcia (min w)	64
Rys. III-6. Czas generowania bazy reguł w zależności od wartości minimalnego wsparcia (min w).	65
Rys. III-7. Czas wnioskowania na podstawie danych uczących w zależności od wartości minimalnego wsparcia (min w)	65
Rys. III-8. Graficzne przedstawienie rozkładu prawdopodobieństwa zajścia jednaczesnego zdarzeń dla zmiennych wejściowych x_1, x_2 i wyjściowych y (macierz E^p)	68
Rys. III-9. Schematyczne przedstawienie obliczeń (120).....	71
Rys. III-10. Liczba reguł elementarnych oraz czas tworzenia reprezentacji wiedzy w zależności od ilości zdefiniowanych zbiorów rozmytych dla każdej zmiennej w modelach: a) AR(2) b) AR(3)	73
Rys. III-11. Czas tworzenia reprezentacji wiedzy w zależności od liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych dla 3,5,7 zdefiniowanych zbiorów rozmytych każdej zmiennej w modelach: a) AR(2) b) AR(3)	74

<i>Rys. III-12. Liczba reguł elementarnych reprezentacji wiedzy w zależności od liczby rozłącznych przedziałów dyskretyzacji wartości zmiennych dla 3,5,7 zdefiniowanych zbiorów rozmytych każdej zmiennej w modelach: a) AR(2) b) AR(3)</i>	74
<i>Rys. III-13. Zależność czasu tworzenia reprezentacji wiedzy od liczby zmiennych modelu dla różnej liczby zdefiniowanych zbiorów rozmytych każdej zmiennej</i>	75
<i>Rys. III-14. Zależność liczby reguł elementarnych reprezentacji wiedzy od liczby zmiennych modelu dla różnej liczby zdefiniowanych zbiorów rozmytych każdej zmiennej</i>	75
<i>Rys. III-15. Etapy modelowania z wykorzystaniem algorytmu wykrywającego ilościowe reguły asocjacji.....</i>	77
<i>Rys. III-16. Algorytm generowania reguł modelu probabilistyczno-rozmytego bazujący na założeniach algorytmu Apriori (na podstawie [rud10]).....</i>	87
<i>Rys. III-17. Algorytm generowania reguł modelu probabilistyczno-rozmytego bazujący na założeniach algorytmu FP-Growth (P)</i>	93
<i>Rys. III-18. Zależność czasu tworzenia reprezentacji wiedzy od liczby zmiennych modelu dla porównywanych algorytmów.....</i>	95
<i>Rys. III-19. Zależność liczby reguł elementarnych modelu wiedzy od liczby zmiennych modelu dla porównywanych algorytmów.....</i>	95
<i>Rys. III-20. Zależność czasu tworzenia reprezentacji wiedzy od liczby zbiorów rozmytych definiowanych dla każdej zmiennej w przypadku: a) procesu AR(1) udziału frakcji lekkiej węgla, b) procesu AR(3) udziału frakcji lekkiej węgla.....</i>	96
<i>Rys. III-21. Zależność liczby reguł elementarnych modelu wiedzy od wartości minimalnego wsparcia</i>	97
<i>Rys. III-22. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia</i>	97
<i>Rys. III-23. Zależności czasu tworzenia reprezentacji wiedzy od liczby rozłącznych przedziałów dyskretyzacji przestrzeni wartości zmiennych dla porównywanych algorytmów oraz dla minimalnego wsparcia o wartościach równych odpowiednio: a) min w=0.0001, b) min w=0.0005, c) min w=0.0016, d) min w=0.005</i>	98
<i>Rys. III-24. Zależność czasu tworzenia reprezentacji wiedzy od liczby danych uczących dla różnych algorytmów, przy założeniach: a) min w=0.0001, b) min w=0.01</i>	99
<i>Rys. III-25. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia dla różnych algorytmów, przy wykorzystaniu: a) 1 tys. rekordów uczących, b) 4 tys. rekordów uczących</i>	99
<i>Rys. III-26. Identyfikacja modelu wiedzy w systemie z probabilistyczno-rozmytą bazą wiedzy</i>	102
<i>Rys. III-27. Przykład wyrowadzenia stopnia aktywacji o-tej reguły plikowej dla wybranych operatorów t-normy.....</i>	105
<i>Rys. III-28. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla operatora: a) Larsena b) Mamdaniego.....</i>	106
<i>Rys. III-29. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla implikacji rozmytej: a) Lukasiewicza b) Zadeha.....</i>	106

Rys. III-30. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla implikacji rozmytej: a) Reichenbacha b) Kleene-Dienes'a	106
Rys. III-31. Rozmyta konkluzja reguły elementarnej (deseń w kropki) na tle zbioru odniesienia następnika (deseń w paski) przy stopniu aktywacji 0.5 dla implikacji rozmytej: a) Gödela b) Goguena	107
Rys. III-32. Ilustracja wyprowadzenia rozmytego wniosku o-tej reguły plikowej	108
Rys. III-33. Wyznaczanie wartości y^* za pomocą: a) metody środka ciężkości (rozszerzonej metody środka ciężkości) b)-c) rozszerzonej metody środka ciężkości d)-e) metody środka ciężkości.....	110
Rys. III-34. Graficzna interpretacja rozmytego wnioskowania w oparciu o probabilistyczno-rozmytą bazę wiedzy (na podstawie [rud10])	110
Rys. III-35. Przykład obiektu systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy jako struktury w Matlabie (na podstawie [bla08]).....	112
Rys. III-36. Diagram sekwencji wywołania funkcji z modułu narzędziowego PFIS, przy tworzeniu nowego systemu z probabilistyczno-rozmytą bazą wiedzy	113
Rys. III-37. Funkcje przynależności dostępne w Toolbox Fuzzy Logic (na podstawie [fuz08])	113
Rys. III-38. Diagram sekwencji wywołania funkcji z modułu narzędziowego PFIS, przy generowaniu probabilistyczno-rozmytej bazy wiedzy	114
Rys. III-39. Interfejs PFISEDIT – okno główne interfejsu z zaznaczonymi zakresami funkcjonalnymi	116
Rys. III-40. Interfejs PFISEDIT – generowanie probabilistyczno-rozmytej bazy reguł	117
Rys. III-41. Przegląd reguł plikowych probabilistyczno-rozmytej bazy wiedzy.....	117
Rys. III-42. Interfejs PFISEDIT – rozkład prawdopodobieństwa wartości rozmytych zmiennych wejścia-wyjścia.....	118
Rys. III-43. Interfejs PFISEDIT – rozkład prawdopodobieństwa warunkowego i brzegowego zdarzeń rozmytych znajdujących się w regułach probabilistyczno-rozmytej bazy wiedzy.....	118
Rys. III-44. Interfejs PFISEDIT – warunkowy rozkład prawdopodobieństwa zajścia zdarzeń rozmytych na wyjściu systemu	119
Rys. III-45. Interfejs PFISEDIT – definiowanie zbiorów rozmytych dla zmiennych systemu	120
Rys. IV-1. Zarys możliwych zastosowań systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy (na podstawie [wb97])	121
Rys. IV-2. Zależność zawartości popiołu w frakcji lekkiej [%] od masowego udziału tej frakcji	123
Rys. IV-3. Definicja zbiorów rozmytych dla wartości lingwistycznych zmiennych: udziału frakcji lekkiej i zawartości popiołu w tej frakcji.....	123
Rys. IV-4. Rozkład prawdopodobieństwa warunkowego wartości lingwistycznych w konkluzji (zawartości popiołu we frakcji lekkiej) dla wybranych reguł plikowych modelu.....	125
Rys. IV-5. Zależność zawartości popiołu we frakcji lekkiej od masowego udziału węgla tej frakcji dla: a) danych empirycznych, b) danych uśrednionych (średnia z 3 wartości), c) danych uśrednionych (średnia z 6 wartości)	126

Rys. IV-6. Rozkład prawdopodobieństwa warunkowego wartości lingwistycznych w konkluzji (zawartości popiołu we frakcji lekkiej) dla wybranych reguł plikowych modelu.....	127
Rys. IV-7. Rozkład prawdopodobieństwa warunkowego wartości lingwistycznych w konkluzji (zawartości popiołu w frakcji lekkiej) dla wybranych reguł plikowych modelu	128
Rys. IV-8. Schemat blokowy filtru jako średniej ruchomej z 7 sygnałów.....	130
Rys. IV-9. Stopnie przynależności dla wartości lingwistycznych modelu wiedzy oraz sposób ich pozyskiwania na podstawie grupowania danych wejście-wyjście.....	131
Rys. IV-10. Zależności błędów dopasowania modelu wiedzy i ilości reguł elementarnych w modelu od wartości minimalnego wsparcia dla: a) danych uczących, b) danych testujących	132
Rys. IV-11. Sygnał wejściowy uczący, sygnał będący średnią ruchomą z 7 próbek danych uczących oraz sygnał wyjściowy systemu rozmytego ($\min w=0,04$)	133
Rys. IV-12. Sygnał wejściowy testujący, sygnał będący średnią ruchomą z 7 próbek danych testujących oraz sygnał wyjściowy systemu rozmytego ($\min w=0,04$).....	133
Rys. IV-13 Funkcja autokorelacji rzeczywistego sygnału wejściowego, sygnału uśrednionego oraz wyjścia systemu rozmytego.....	134
Rys. IV-14. Widmo częstotliwościowe Fouriera dla sygnału wejściowego, uśrednionego po 7 próbkach i wyjścia modelu rozmytego	134
Rys. IV-15. Aproksymacja funkcji f za pomocą fuzzy graph f^* (na podstawie [zad97]).....	135
Rys. IV-16. Aproksymacja funkcji f za pomocą fuzzy graph f_p^* z rozkładem prawdopodobieństwa brzegowego i warunkowego zdarzeń rozmytych	136
Rys. IV-17. Schemat układu charakterystyki wzbogacania węgla.....	136
Rys. IV-18. Definicja zbiorów rozmytych dla zmiennych systemu	137
Rys. IV-19. Zależności błędu RMSE oraz liczby reguł elementarnych od wartości minimalnego wsparcia dla: a) danych uczących, b) danych testujących.....	140
Rys. IV-20. a) Ciąg pomiarów danych uczących (linia kropkowana), oraz ciąg uzyskany przy zastosowaniu systemu z probabilistyczno-rozmytą bazą wiedzy (linia ciągła), b) Błąd jako różnica wartości wyliczonej od wartości rzeczywistej	141
Rys. IV-21. a) Ciąg pomiarów danych testujących (linia kropkowana), oraz ciąg uzyskany przy zastosowaniu systemu z probabilistyczno-rozmytą bazą wiedzy (linia ciągła), b) Błąd jako różnica wartości wyliczonej i wartości rzeczywistej	141
Rys. IV-22. Unormowana funkcja autokorelacji błędów dla danych uczących i testujących	142
Rys. IV-23. Histogram błędów dla danych uczących i testujących łącznie	142
Rys. IV-24. Zależność między wartością rzeczywistą wynikającą z pomiarów A_c a wartością wyliczoną na podstawie modelu \hat{A}_c , dla danych uczących i testujących.....	142
Rys. IV-25. Zależności między wybranymi zmiennymi wejściowymi a zmienną wyjściową modelu rozmytego (na podstawie danych uczących).....	144
Rys. IV-26. Stopnie przynależności przedziałów poszczególnych wartości zmiennych wejściowych i zmiennej wyjściowej do zdefiniowanych zbiorów rozmytych.....	145

Rys. IV-27. Zależności procentu błędnych odpowiedzi systemu oraz liczby reguł elementarnych od wartości minimalnego wsparcia dla: a) danych uczących, b) danych testujących 147

Rys. IV-28. Porównanie wyników symulacji z wartościami wyliczonymi na podstawie modelu z probabilistyczno-rozmytą bazą wiedzy (dla danych testujących) 148

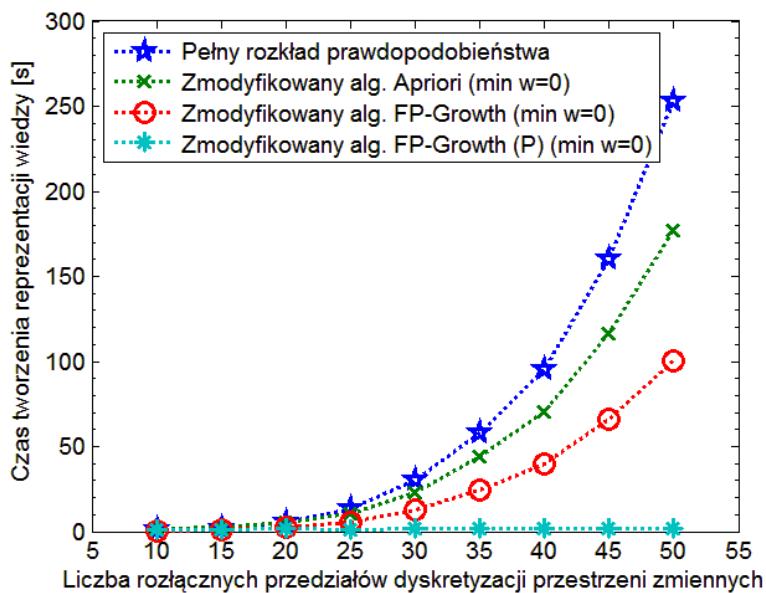
Spis tabel

<i>Tab. II-1. Przykłady odcinkowo-liniowych funkcji przynależności [rut06], [łęs08], [pie03], [dri96]</i>	14
<i>Tab. II-2. Funkcja przynależności – rozmyty singleton [rut06], [pie03]</i>	15
<i>Tab. II-3. Przykłady nieliniiowych funkcji przynależności [pie03], [łęs08]</i>	15
<i>Tab. II-4. Wybrane normy trójkątne [pie03], [łęs08], [cpa09]</i>	21
<i>Tab. II-5. Wybrane implikacje rozmyte [kac01], [łęs08], [now09]</i>	30
<i>Tab. II-6. Przykłady metod wystrzania [lee99], [pie03], [bro06], [łęs08], [now09], [duc00]</i>	41
<i>Tab. III-1. Operatory t-normy, które mogą być użyte do generowania reguł modelu probabilistyczno-rozmytego [pie03], [łęs08], [cpa09]</i>	69
<i>Tab. III-2. Wartości stopni przynależności $\mu_{A_j}(a_i)$ ($i=1, \dots, 40$) zbiorów rozmytych dla udziału frakcji lekkiej węgla gdzie: a) $j=1, \dots, 3$, b) $j=1, \dots, 5$, c) $j=1, \dots, 7$</i>	72
<i>Tab. III-3. Przykład ilościowych reguł asocjacji zmiennych $x(t), x(t-1)$</i>	79
<i>Tab. III-4. Wartości stopni przynależności $\mu_{A_j}(a_i)$ ($j=1, \dots, 5$; $i=1, \dots, 10$) zbiorów rozmytych dla udziału frakcji lekkiej węgla</i>	80
<i>Tab. III-5. Łączny rozkład prawdopodobieństwa dla zmiennych lingwistycznych $x(t), x(t-1)$</i>	80
<i>Tab. III-6. Warunkowy rozkład prawdopodobieństwa dla zmiennych lingwistycznych $x(t), x(t-1)$</i>	81
<i>Tab. III-7. Przykład rozmytych reguł asocjacji dla zmiennych lingwistycznych $x(t), x(t-1)$</i>	81
<i>Tab. III-8. Wartości stopni przynależności $\mu_{B_j}(b_i)$ zbiorów rozmytych dla zawartości popiołu we frakcji lekkiej</i>	82
<i>Tab. III-9. Przykład ilościowych reguł asocjacji modelu $y(t)=f(x(t-2), x(t-1), y(t-2), y(t-1))$</i>	82
<i>Tab. III-10. Łączny rozkład prawdopodobieństwa dla poszukiwanego probabilistyczno-rozmytego modelu wiedzy</i>	82
<i>Tab. IV-1. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych wejścia-wyjścia</i>	124
<i>Tab. IV-2. Rozkład prawdopodobieństwa brzegowego wartości lingwistycznych przesłanki (udziału frakcji lekkiej)</i>	124
<i>Tab. IV-3. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych wejścia-wyjścia (udziału frakcji lekkiej i zawartości popiołu w tej frakcji)</i>	126
<i>Tab. IV-4. Rozkład prawdopodobieństwa brzegowego wartości lingwistycznych przesłanki (udziału frakcji lekkiej)</i>	126
<i>Tab. IV-5. Rozkład prawdopodobieństwa wartości lingwistycznych zmiennych wejścia-wyjścia (udziału frakcji lekkiej i zawartości popiołu w tej frakcji)</i>	127
<i>Tab. IV-6. Rozkład prawdopodobieństwa brzegowego wartości lingwistycznych przesłanki (udziału frakcji lekkiej)</i>	128
<i>Tab. IV-7. Dokładność działania systemu jako filtru - średniej ruchomej z 7 sygnałów</i>	131

<i>Tab. IV-8. Błąd RMSE [%] aproksymacji zależności parametrów za pomocą fuzzy graph f_p* w zależności od różnych operatorów interpretacji reguł oraz t-normy jako spójnika logicznego AND (dla modelu uwzględniającego pełny rozkład prawdopodobieństwa zdarzeń rozmytych)</i>	138
<i>Tab. IV-9. Procent błędnych decyzji w zależności od różnych operatorów interpretacji reguł oraz t-normy jako spójnika logicznego AND (dla modelu uwzględniającego pełny rozkład prawdopodobieństwa zdarzeń rozmytych)</i>	146
<i>Tab. IV-10. Prawdopodobieństwo jednoczesnego zajścia wybranych zdarzeń rozmytych w regułach</i>	150
<i>Tab. IV-11. Prawdopodobieństwo jednoczesnego zajścia wybranych zdarzeń rozmytych w regułach</i>	151
<i>Tab. IV-12. Prawdopodobieństwo jednoczesnego zajścia wybranych zdarzeń rozmytych w regułach</i>	151

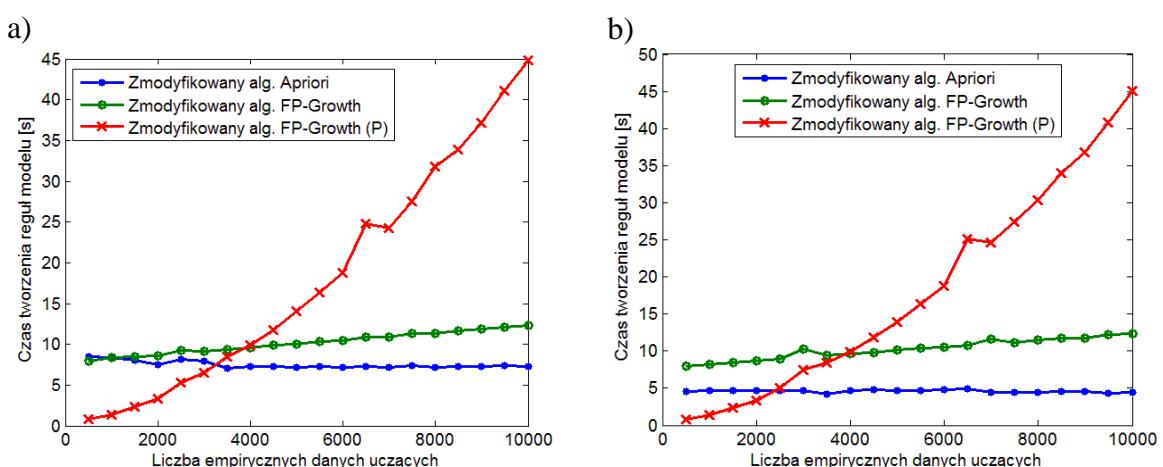
Dodatek A – Wyniki porównania algorytmów przy różnych założeniach wejściowych

Model wiedzy procesu AR(3) udziału frakcji lekkiej węgla (blisko 500 danych pomiarowych) z pełnym rozkładem prawdopodobieństwa zdarzeń rozmytych w regułach, w których zdefiniowano po 5 zbiorów rozmytych dla każdej zmiennej oraz zastosowano 35 rozłącznych przedziałów dyskretyzacji wartości zmiennych.

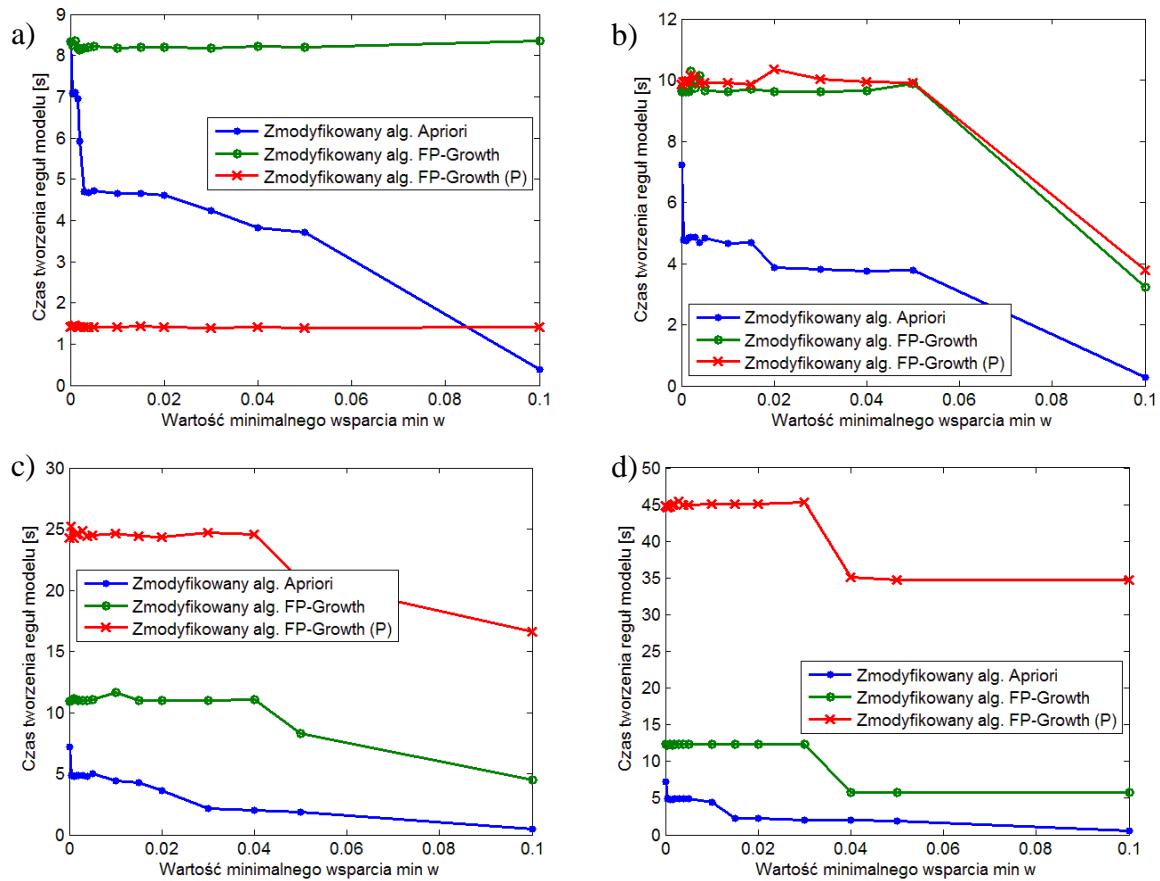


Rys. A-1. Zależności czasu tworzenia reprezentacji wiedzy od liczby rozłącznych przedziałów dyskretyzacji przestrzeni wartości zmiennych dla procesu AR(3) udziału frakcji lekkiej węgla

Model wiedzy dla procesu AR(3) prędkości wiatru z okresem próbkowania co 1 minutę. W modelu uwzględniono po 3 zbiory rozmyte typu trójkątnego dla każdej zmiennej oraz użyto 35 rozłącznych przedziałów dyskretyzacji wartości zmiennych.

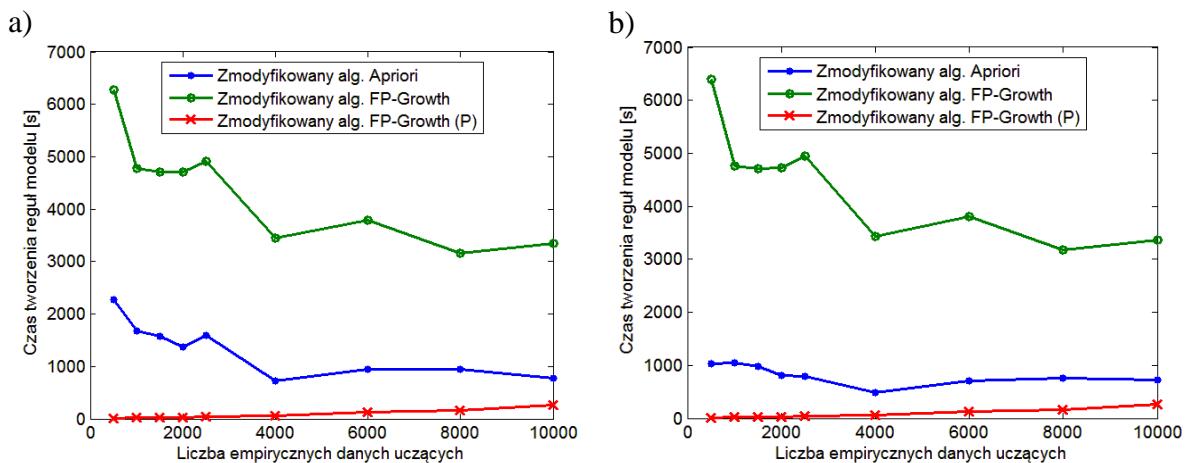


Rys. A-2. Zależność czasu tworzenia reprezentacji wiedzy od liczby danych uczących dla różnych algorytmów, przy założeniach: a) min w=0.0001, b) min w=0.01

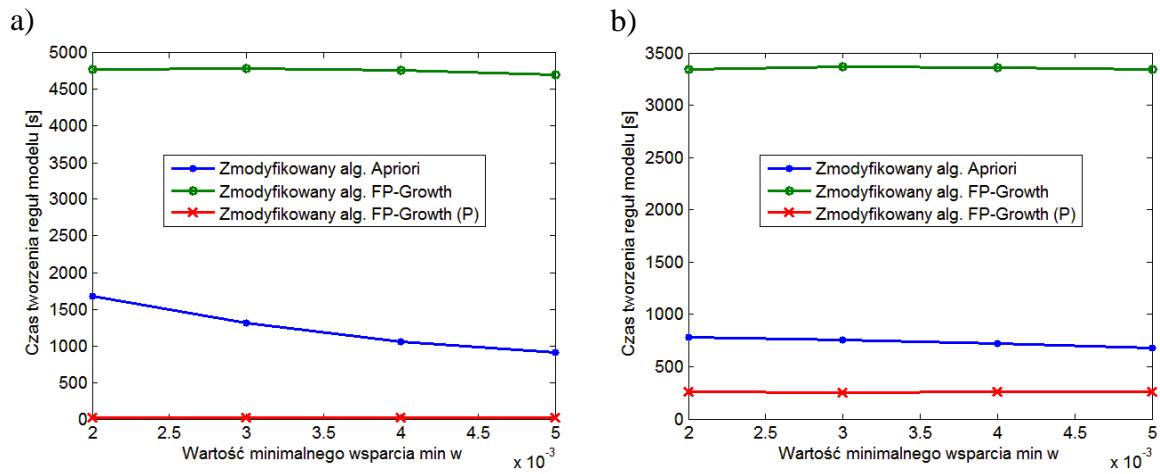


Rys. A-3. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia dla różnych algorytmów przy wykorzystaniu: a) 1 tys. rekordów uczących, b) 4 tys. rekordów uczących, c) 7 tys. rekordów uczących, d) 10 tys. rekordów uczących

Model wiedzy dla procesu AR(4) dla prędkości wiatru z okresem próbkowania co 1 minutę. W modelu uwzględniono po 5 zbiorów rozmytych typu trójkątnego dla każdej zmiennej oraz użyto 35 rozłącznych przedziałów dyskretyzacji wartości zmiennych.



Rys. A-4. Zależność czasu tworzenia reprezentacji wiedzy od liczby danych uczących dla różnych algorytmów, przy założeniach: a) min w=0.002, b) min w=0.004



Rys. A-5. Zależność czasu tworzenia reprezentacji wiedzy od wartości minimalnego wsparcia dla różnych algorytmów przy wykorzystaniu: a) 4 tys. rekordów uczących, b) 10 tys. rekordów uczących,

Dodatek B – Opis funkcji z modułu narzędziowego PFIS

Opis głównych funkcji z modułu narzędziowego PFIS, wraz z ich parametrami wywołania.

Tab. B-1. Opis wywołania funkcji *newpfis* - tworzącej system z probabilistyczno-rozmytą bazą wiedzy

Funkcja	
pfis=newpfis(pfisName, inX, outX, numGmfs, optionMf, optionMet, optionX)	
Nazwa argumentu	Opis
pfisName	Nazwa tworzonego systemu.
inX	Macierz z wartościami empirycznymi dla zmiennych wejściowych.
outX	Macierz z wartościami empirycznymi dla zmiennej wyjściowej.
numGmfs	Liczba rozłącznych przedziałów w przestrzeniach zmiennych (domyślnie: 27).
optionMf	Opcje definicji stopni przynależności zbiorów rozmytych (domyślnie {'trimf' 7}).
optionMet	Wektor z parametrami wnioskowania rozmytego: {'operAnd' 'operImp' 'metDefuzz'} (domyślnie {'prod' 'prod' 'cog'}).
optionX	Macierz dwu-wierszowa zawierająca granice dziedzin zmiennych wejścia/wyjścia (opcjonalnie).
pfis	Utworzona struktura dla systemu rozmytego typu MISO z parametrami określonymi poprzez argumenty wywołania funkcji.

Tab. B-2. Opis wywołania funkcji *genrulespfis* - generującej probabilistyczno-rozmytą bazę wiedzy w oparciu o idee rozmytych reguł asocjacji

Funkcja	
pfis=genrulespfis(pfisName, algName, inX, outX, tNorm, minSup)	
Nazwa argumentu	Opis
pfisName	Struktura systemu rozmytego typu MISO dla której ma być wygenerowana probabilistyczno-rozmyta baza wiedzy.
algName	Nazwa algorytmu użytego do utworzenia bazy wiedzy ('apriori' -zmodyfikowany algorytm Apriori , 'fpg' - zmodyfikowany algorytm FP-Growth lub 'fpg21' - zmodyfikowany algorytm FP-Growth (P)).
inX	Macierz z wartościami empirycznymi dla zmiennych wejściowych (jeżeli brak to wykorzystana jest macierz z struktury systemu).
outX	Macierz z wartościami empirycznymi dla zmiennej wyjściowej (jeżeli brak to wykorzystana jest macierz z struktury systemu).
tNorm	Operator t-normy użyty do tworzenia reguł (domyślnie: 'prod'). (Omawiana wersja modułu narzędziowego umożliwia wykorzystanie jedynie operatora: 'prod').
minSup	Wartość minimalnego wsparcia (minimalne łączne prawdopodobieństwo dla zdarzeń rozmytych w przesłance i konkluzji) (domyślnie: wartość podana w strukturze systemu).
pfis	Struktura systemu rozmytego typu MISO z wygenerowaną probabilistyczno-rozmytą bazą wiedzy.

Tab. B-3. Opis wywołania funkcji *genrules* - generującej probabilistyczno-rozmytą bazę wiedzy z pełnym rozkładem prawdopodobieństwa zdarzeń rozmytych w regułach

Funkcja	
pfis=genrules(pfisName, algName, inX, outX, tNorm, minSup)	
Nazwa argumentu	Opis
pfisName	Struktura systemu rozmytego typu MISO dla której ma być wygenerowana probabilistyczno-rozmyta baza wiedzy.
inX	Macierz z wartościami empirycznymi dla zmiennych wejściowych (jeżeli brak to wykorzystana jest macierz z struktury systemu).
outX	Macierz z wartościami empirycznymi dla zmiennej wyjściowej (jeżeli brak to wykorzystana jest macierz z struktury systemu).
tNorm	Operator t-normy użyty do tworzenia reguł (domyślnie: 'prod').
pfis	Struktura systemu rozmytego typu MISO z wygenerowaną probabilistyczno-rozmytą bazą wiedzy.

Tab. B-4. Opis wywołania funkcji *infermodpfis* - wnioskującej w oparciu o system z probabilistyczno-rozmytą bazą wiedzy

Funkcja	
outY=infermodpfis(pfisName,inX)	
Nazwa argumentu	Opis
pfisName	Struktura systemu rozmytego typu MISO z probabilistyczno-rozmytą bazą wiedzy.
inX	Macierz z wartościami empirycznymi dla zmiennej wyjściowej.
outY	Macierz z numerycznymi wartościami zmiennej wyjściowej.

Tab. B-5. Opis wywołania funkcji *getpfis* - zwracającą cechy opisujące system. Typ cechy jest uzależniony od argumentów funkcji.

Funkcja	
out=getpfis(pfisName,attr1,attr2)	
Nazwa argumentu	Opis
pfisName	Struktura systemu rozmytego typu MISO z probabilistyczno-rozmytą bazą wiedzy.
attr1	Nazwa parametru modelu wiedzy, który ma zwracać funkcja, dopuszczone są: - 'all' - zwraca wszystkie informacje o systemie, - 'rules' - zwraca reguły modelu, - 'gmfs' - zwraca wartości stopni przynależności dla rozłącznych przedziałów wartości zmiennych w systemie, - 'plotgmfs' - zwraca wykres wartości stopni przynależności dla zmiennych w systemie, kolejny atrybut (attr2) wówczas zawiera numer wejścia (wyjścia - liczba wejść+1), - 'minD' - zwraca minimalne wartości dla przedziałów wartości zmiennych, - 'argD' - zwraca środki przedziałów wartości zmiennych.
attr2	Nazwa atrybutu, który oznacza sposób wyświetlenia danych: - 'vec' - opis wektorowy, - 'disc' - opis jako string (domyślnie).

Opis pozostałych funkcji z modułu narzędziowego PFIS

- addinmf** - zdefiniowanie wartości rozmytej dla zmiennej wejściowej w strukturze systemu.
- addinp** - dodanie zmiennej wejściowej do struktury systemu.
- addoutmf** - zdefiniowanie wartości rozmytej dla zmiennej wyjściowej w strukturze systemu.
- addoutp** - dodanie zmiennej wyjściowej do struktury systemu.
- addpfis** - utworzenie struktury nowego systemu.
- addrules** - dodanie probabilistyczno-rozmytych reguł do struktury systemu.
- calcfrules** - generowanie probabilistyczno-rozmytych reguł według algorytmu tworzącego reguły z wagami stanowiącymi pełny rozkład prawdopodobieństwa zdarzeń rozmytych znajdujących się w regułach. Funkcja wywoływana obligatoryjnie przy okazji tworzenia bazy wiedzy za pomocą opcji *genrules*.
- calcfprob** - generowanie bazy reguł lingwistycznych wraz z pełnym rozkładem prawdopodobieństwa jednoczesnego zajścia zdarzeń rozmytych w regułach, z wykorzystaniem danych uczących i odpowiedniego operatora t-normy. Funkcja wywoywana obligatoryjnie przy okazji tworzenia bazy wiedzy za pomocą opcji *genrules*.
- calcrulespfis_apr** - generowanie bazy reguł lingwistycznych wraz z prawdopodobieństwem jednoczesnego zajścia zdarzeń rozmytych w regułach. Funkcja wykorzystuje zmodyfikowany algorytm Apriori uwzględniając do tego dane uczące i wskazany operator t-normy (obecna wersja modułu narzędziowego umożliwia wykorzystanie jedynie operatora 'prod').
- calcrulespfis_fpg** - generowanie bazy reguł lingwistycznych wraz z prawdopodobieństwem jednoczesnego zajścia zdarzeń rozmytych w regułach. Funkcja wykorzystuje zmodyfikowany algorytm FP-Growth uwzględniając do tego dane uczące i wskazany operator t-normy (obecna wersja modułu narzędziowego umożliwia wykorzystanie jedynie operatora 'prod').
- calcrulespfis_fpg21** - generowanie bazy reguł lingwistycznych wraz z prawdopodobieństwem jednoczesnego zajścia zdarzeń rozmytych w regułach. Funkcja wykorzystuje zmodyfikowany algorytm FP-Growth (P) (obliczanie prawdopodobieństwa zbiorów rozmytych z uwzględnieniem mocy zbiorów rozmytych), uwzględniając do tego dane uczące i wskazany operator t-normy (obecna wersja modułu narzędziowego umożliwia wykorzystanie jedynie operatora 'prod').
- calcw1** - obliczanie wag dla przesłanek reguł rozmytych w probabilistyczno-rozmytej bazie wiedzy.
- calcw2** - obliczanie wag dla konkluzji reguł rozmytych w probabilistyczno-rozmytej bazie wiedzy.

- change_minsup_pfis** - modyfikacja probabilistyczno-rozmytej bazy wiedzy przez zwiększenie wartości jednoczesnego prawdopodobieństwa zajścia zdarzeń rozmytych w przesłance i konkluzji reguł (wartości minimalnego wsparcia).
- delpfis** - usunięcie struktury systemu z probabilistyczno-rozmytą bazą wiedzy lub jego wybranych elementów składowych.
- gen_candcomp** - generowanie kandydatów zbiorów częstych k-elementowych (funkcja wywoływana podczas generowania probabilistyczno-rozmytej bazy wiedzy według zmodyfikowanego algorytmu Apriori).
- gen_candidatefar** - generowanie kandydatów zbiorów częstych k-elementowych (funkcja wywoływana podczas generowania probabilistyczno-rozmytej bazy wiedzy według zmodyfikowanego algorytmu Apriori).
- genMf_kmeans** - wygenerowanie parametrów zbiorów rozmytych dla zmiennych wejściowych i wyjściowych systemu na podstawie grupowania metodą k-means.
- genweight** - obliczanie wag reguł rozmytych dla probabilistyczno-rozmytej bazy wiedzy.
- getpfis** - pobranie wskazanych informacji na temat struktury systemu z probabilistyczno-rozmytą bazą wiedzy.
- getsupport** - pobranie wartości jednoczesnego prawdopodobieństwa zajścia zdarzeń rozmytych w przesłance i konkluzji reguł probabilistyczno-rozmytej bazy wiedzy.
- inferprobpfis** - wyliczenie prawdopodobieństw zajścia zdarzeń rozmytych w konkluzji przy faktach określonych w postaci numerycznej.
- modpfis** - modyfikacja wybranych pól struktury systemu.
- mult2one** - zamiana na jednowymiarowe indeksowanie macierzy.
- showrules** - wydrukowanie reguł probabilistyczno-rozmytej bazy wiedzy.
- showrulesf** - zapis reguł probabilistyczno-rozmytej bazy wiedzy do pliku.

Szczegóły dotyczące wyżej wymienionych funkcji oraz argumenty ich wywołania zostały zapisane w komentarzach plików funkcyjnych.

Dodatek C – Probabilistyczno-rozmyta baza wiedzy dla modelu wielowymiarowej charakterystyki statycznej

Fuzzy graph f^*_c , przedstawiony w postaci probabilistyczno-rozmytej bazy wiedzy, jako odzwierciedlenie zależności następujących parametrów węgla:

- Q_1 - procentowy masowy udział lżejszej frakcji gęstościowej węgla (w skrócie nazywany udziałem frakcji lżejszej węgla) o gęstości węgla mniejszej od $1,5 \cdot 10^3 \text{ kg/m}^3$,
- Q_2 - procentowy masowy udział cięższej frakcji gęstościowej węgla (w skrócie nazywany udziałem frakcji cięższej węgla) o gęstości węgla większej od $1,5 \cdot 10^3 \text{ kg/m}^3$,
- A_1 - zawartość popiołu w frakcji lżejszej węgla,
- A_2 - zawartość popiołu w frakcji cięższej węgla,
- A_c - całkowita zawartość popiołu w próbce węgla.

- 1: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS duża) [0.1581]
THEN (Ac IS średnia) [0.6946]
ALSO (Ac IS duża) [0.2756]
ALSO (Ac IS mała) [0.0298]
- 2: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS duża) [0.0758]
THEN (Ac IS średnia) [0.5966]
ALSO (Ac IS mała) [0.4034]
- 3: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS duża) [0.0640]
THEN (Ac IS średnia) [0.7554]
ALSO (Ac IS duża) [0.1965]
ALSO (Ac IS mała) [0.0481]
- 4: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0584]
THEN (Ac IS średnia) [0.7821]
ALSO (Ac IS duża) [0.1103]
ALSO (Ac IS mała) [0.1076]
- 5: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS duża) [0.0402]
THEN (Ac IS średnia) [0.7979]
ALSO (Ac IS mała) [0.1631]
ALSO (Ac IS duża) [0.0389]
- 6: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS duża) [0.0402]
THEN (Ac IS średnia) [0.7979]
ALSO (Ac IS mała) [0.1631]
ALSO (Ac IS duża) [0.0389]
- 7: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS duża) [0.0350]
THEN (Ac IS średnia) [0.5409]
ALSO (Ac IS mała) [0.4591]
- 8: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS średnia) [0.0314]

- THEN (Ac IS mała) [0.5734]
ALSO (Ac IS średnia) [0.4266]
- 9: IF (Q1 IS średni) AND (Q2 IS średni) AND (A1 IS średnia) AND (A2 IS duża) [0.0298]
THEN (Ac IS średnia) [0.5352]
ALSO (Ac IS duża) [0.4648]
- 10: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS duża) [0.0297]
THEN (Ac IS duża) [0.7631]
ALSO (Ac IS średnia) [0.1709]
ALSO (Ac IS b. duża) [0.0659]
- 11: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0262]
THEN (Ac IS średnia) [0.7612]
ALSO (Ac IS mała) [0.1750]
ALSO (Ac IS duża) [0.0638]
- 12: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS duża) [0.0226]
THEN (Ac IS duża) [0.6275]
ALSO (Ac IS średnia) [0.3725]
- 13: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS duża) [0.0225]
THEN (Ac IS duża) [0.6304]
ALSO (Ac IS średnia) [0.3696]
- 14: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS duża) [0.0180]
THEN (Ac IS średnia) [0.7775]
ALSO (Ac IS mała) [0.2225]
- 15: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS duża) [0.0180]
THEN (Ac IS średnia) [0.7775]
ALSO (Ac IS mała) [0.2225]
- 16: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS średnia) [0.0158]
THEN (Ac IS mała) [0.6686]
ALSO (Ac IS średnia) [0.3314]
- 17: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS średnia) [0.0152]
THEN (Ac IS średnia) [0.7003]
ALSO (Ac IS mała) [0.2997]
- 18: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0152]
THEN (Ac IS średnia) [0.7003]
ALSO (Ac IS mała) [0.2997]
- 19: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS b. duża) [0.0150]
THEN (Ac IS duża) [0.5023]
ALSO (Ac IS średnia) [0.4977]
- 20: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS średnia) [0.0148]
THEN (Ac IS duża) [0.5804]
ALSO (Ac IS średnia) [0.4196]

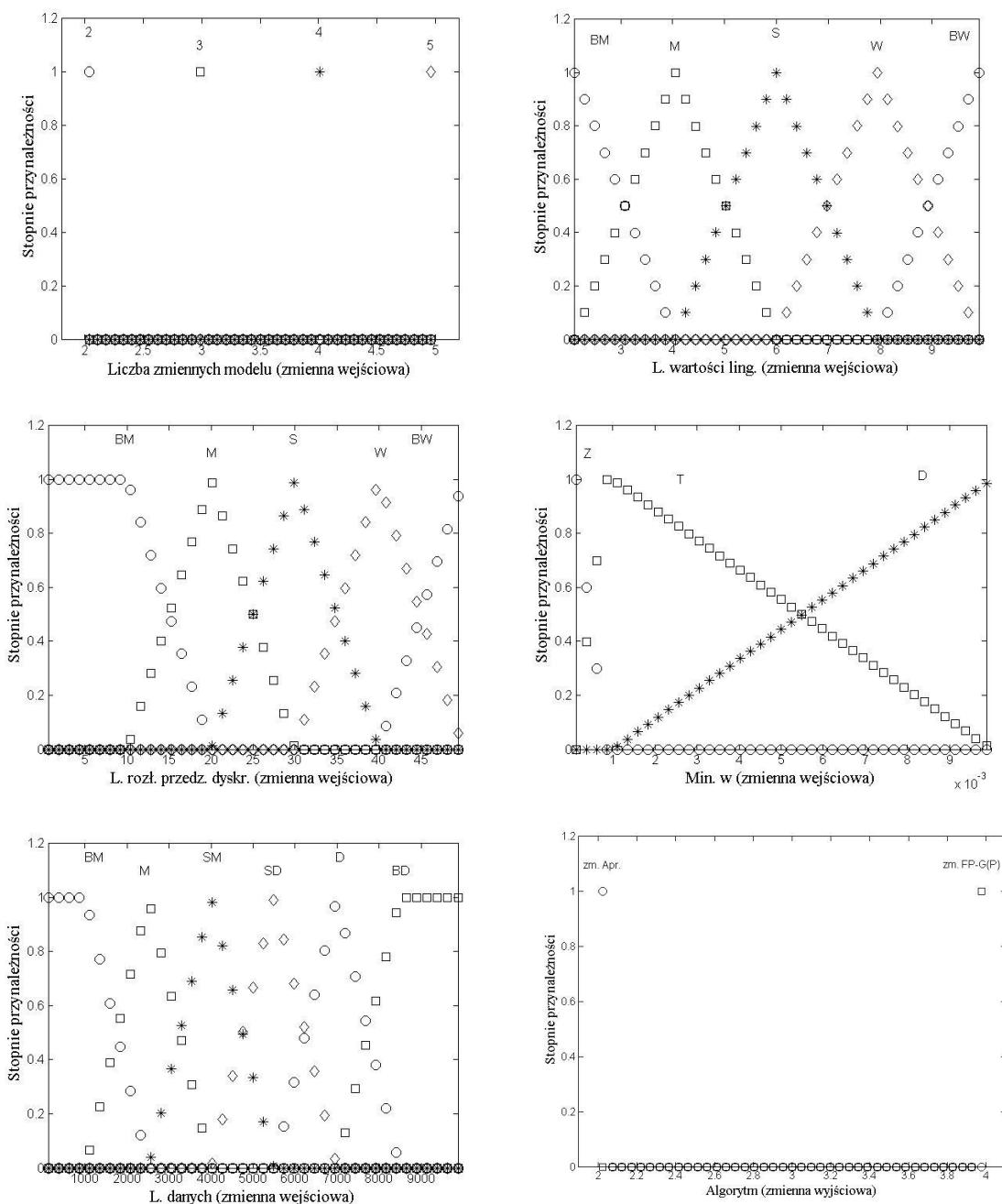
- 21: IF (Q1 IS mały) AND (Q2 IS b. mały) AND (A1 IS duża) AND (A2 IS duża) [0.0120]
THEN (Ac IS duża) [0.7625]
ALSO (Ac IS średnia) [0.2375]
- 22: IF (Q1 IS średni) AND (Q2 IS średni) AND (A1 IS średnia) AND (A2 IS średnia) [0.0102]
THEN (Ac IS średnia) [0.7065]
ALSO (Ac IS duża) [0.2935]
- 23: IF (Q1 IS mały) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS duża) [0.0091]
THEN (Ac IS duża) [0.5671]
ALSO (Ac IS średnia) [0.4329]
- 24: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS duża) AND (A2 IS duża) [0.0091]
THEN (Ac IS duża) [0.5673]
ALSO (Ac IS średnia) [0.4327]
- 25: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0079]
THEN (Ac IS średnia) [0.6489]
ALSO (Ac IS duża) [0.3511]
- 26: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS średnia) [0.0077]
THEN (Ac IS średnia) [0.6443]
ALSO (Ac IS duża) [0.3557]
- 27: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS średnia) [0.0075]
THEN (Ac IS średnia) [0.5510]
ALSO (Ac IS mała) [0.4490]
- 28: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0075]
THEN (Ac IS średnia) [0.5510]
ALSO (Ac IS mała) [0.4490]
- 29: IF (Q1 IS mały) AND (Q2 IS b. mały) AND (A1 IS duża) AND (A2 IS średnia) [0.0070]
THEN (Ac IS średnia) [0.5190]
ALSO (Ac IS duża) [0.4810]
- 30: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS mała) [0.0068]
THEN (Ac IS średnia) [0.5752]
ALSO (Ac IS mała) [0.4248]
- 31: IF (Q1 IS średni) AND (Q2 IS średni) AND (A1 IS mała) AND (A2 IS duża) [0.0064]
THEN (Ac IS średnia) [0.7135]
ALSO (Ac IS duża) [0.2865]
- 32: IF (Q1 IS duży) AND (Q2 IS średni) AND (A1 IS średnia) AND (A2 IS duża) [0.0064]
THEN (Ac IS średnia) [0.7135]
ALSO (Ac IS duża) [0.2865]
- 33: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS b. duża) [0.0055]
THEN (Ac IS średnia) [0.6410]
ALSO (Ac IS duża) [0.3590]
- 34: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS b. mała) AND (A2 IS duża) [0.0054]

- THEN (Ac IS mała) [1.0000]
- 35: IF (Q1 IS b. duży) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS duża) [0.0052]
 THEN (Ac IS mała) [1.0000]
- 36: IF (Q1 IS b. duży) AND (Q2 IS mały) AND (A1 IS b. mała) AND (A2 IS duża) [0.0046]
 THEN (Ac IS mała) [1.0000]
- 37: IF (Q1 IS duży) AND (Q2 IS średni) AND (A1 IS mała) AND (A2 IS duża) [0.0046]
 THEN (Ac IS średnia) [1.0000]
- 38: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS b. duża) [0.0042]
 THEN (Ac IS średnia) [1.0000]
- 39: IF (Q1 IS średni) AND (Q2 IS duży) AND (A1 IS średnia) AND (A2 IS duża) [0.0041]
 THEN (Ac IS duża) [1.0000]
- 40: IF (Q1 IS duży) AND (Q2 IS średni) AND (A1 IS mała) AND (A2 IS średnia) [0.0037]
 THEN (Ac IS średnia) [0.5923]
 ALSO (Ac IS mała) [0.4077]
- 41: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS b. duża) [0.0030]
 THEN (Ac IS duża) [1.0000]
- 42: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS b. duża) [0.0027]
 THEN (Ac IS średnia) [1.0000]
- 43: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS b. duża) [0.0027]
 THEN (Ac IS średnia) [1.0000]
- 44: IF (Q1 IS średni) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS b. duża) [0.0025]
 THEN (Ac IS duża) [1.0000]
- 45: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS średnia) AND (A2 IS b. duża) [0.0025]
 THEN (Ac IS duża) [1.0000]
- 46: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS b. duża) [0.0025]
 THEN (Ac IS średnia) [1.0000]
- 47: IF (Q1 IS mały) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS średnia) [0.0022]
 THEN (Ac IS średnia) [1.0000]
- 48: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS duża) AND (A2 IS średnia) [0.0022]
 THEN (Ac IS średnia) [1.0000]
- 49: IF (Q1 IS duży) AND (Q2 IS mały) AND (A1 IS b. mała) AND (A2 IS średnia) [0.0021]
 THEN (Ac IS mała) [1.0000]
- 50: IF (Q1 IS b. duży) AND (Q2 IS mały) AND (A1 IS mała) AND (A2 IS średnia) [0.0021]
 THEN (Ac IS mała) [1.0000]
- 51: IF (Q1 IS b. duży) AND (Q2 IS b. mały) AND (A1 IS b. mała) AND (A2 IS duża) [0.0020]
 THEN (Ac IS mała) [1.0000]
- 52: IF (Q1 IS b. duży) AND (Q2 IS mały) AND (A1 IS b. mała) AND (A2 IS średnia) [0.0020]
 THEN (Ac IS mała) [1.0000]
- 53: IF (Q1 IS średni) AND (Q2 IS średni) AND (A1 IS mała) AND (A2 IS średnia) [0.0020]
 THEN (Ac IS średnia) [1.0000]

- 54: IF (Q1 IS duży) AND (Q2 IS średni) AND (A1 IS średnia) AND (A2 IS średnia) [0.0020]
THEN (Ac IS średnia) [1.0000]
- 55: IF (Q1 IS mały) AND (Q2 IS średni) AND (A1 IS duża) AND (A2 IS duża) [0.0019]
THEN (Ac IS duża) [1.0000]
- 56: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS b. mała) AND (A2 IS duża) [0.0019]
THEN (Ac IS mała) [1.0000]
- 57: IF (Q1 IS b. duży) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS duża) [0.0018]
THEN (Ac IS mała) [1.0000]
- 58: IF (Q1 IS średni) AND (Q2 IS b. mały) AND (A1 IS mała) AND (A2 IS b. duża) [0.0016]
THEN (Ac IS średnia) [1.0000]
- 59: IF (Q1 IS duży) AND (Q2 IS b. mały) AND (A1 IS średnia) AND (A2 IS b. duża) [0.0016]
THEN (Ac IS średnia) [1.0000]
- 60: IF (Q1 IS mały) AND (Q2 IS mały) AND (A1 IS duża) AND (A2 IS mała) [0.0016]
THEN (Ac IS średnia) [1.0000]
- 61: IF (Q1 IS średni) AND (Q2 IS średni) AND (A1 IS duża) AND (A2 IS duża) [0.0016]
THEN (Ac IS duża) [1.0000]
- 62: IF (Q1 IS mały) AND (Q2 IS średni) AND (A1 IS średnia) AND (A2 IS duża) [0.0016]
THEN (Ac IS duża) [1.0000]

Dodatek E – Wyniki analizy systemu decyzyjnego z większą ilością zbiorów rozmytych

Wyniki analizy przykładu systemu wnioskującego z probabilistyczno-rozmytą bazą wiedzy służącego do wyboru efektywnego algorytmu generującego probabilistyczno-rozmytą bazę wiedzy w oparciu o założenia wejściowe modelu wiedzy tj.: liczba zmiennych modelu (4 zbiorów rozmyte w postaci singletonów), średnia liczba wartości lingwistycznych dla zmiennych w modelu (5 zbiorów rozmytych), liczba rozłącznych przedziałów dyskretyzacji w przestrzeniach zmiennych (5 zbiorów rozmytych), minimalna wartość wsparcia (3 zbiorów rozmyte) oraz liczba rekordów danych uczących (2 zbiorów rozmyte w postaci singletonów).

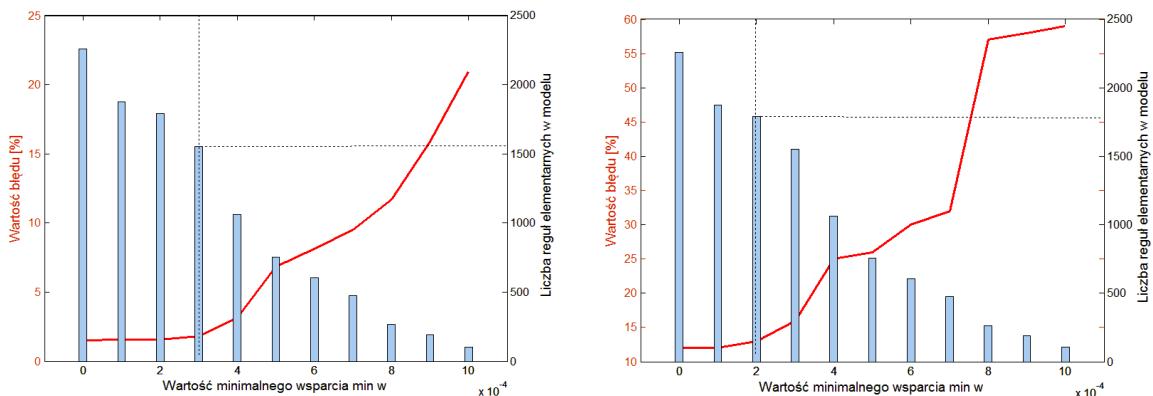


Rys. E-1. Stopnie przynależności wartości poszczególnych zmiennych wejściowych i zmiennej wyjściowej do zdefiniowanych zbiorów rozmytych

Tab. E-1. Procent błędnych decyzji w zależności od różnych operatorów interpretacji reguł oraz t-normy jako spójnika logicznego AND (dla modelu wiedzy uwzględniającego pełny rozkład prawdopodobieństwa zdarzeń rozmytych)

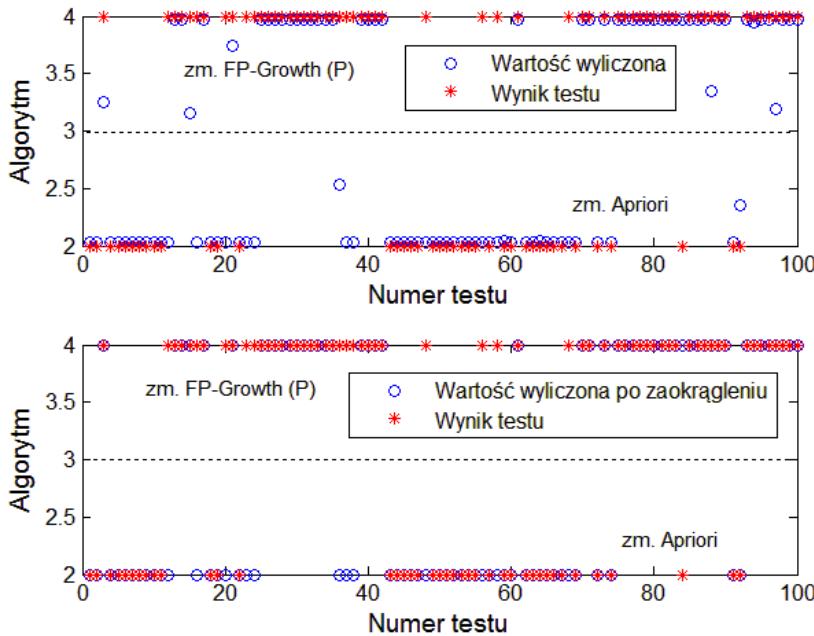
Spójnik log. AND Interpretacja reguł	T-norma Zadeha (min)		T-norma Łukasiewicza		T-norma Fodora		T-norma iloczynu		T-norma Hamachera		T-norma Einsteina		T-norma drastyczna	
Iloczyn algebraiczny	1,48%	12,00%	4,71%	20,00%	1,90%	13,00%	1,53%	12,00%	1,49%	12,00%	1,53%	12,00%	23,47%	57,00%
Minimum	1,48%	12,00%	4,71%	20,00%	1,90%	13,00%	1,53%	12,00%	1,49%	12,00%	1,53%	12,00%	23,47%	57,00%
Implikacja Gödela	5,59%	15,00%	4,72%	21,00%	2,00%	14,00%	5,59%	15,00%	5,59%	15,00%	5,59%	15,00%	23,49%	57,00%
Implikacja Goguena	5,59%	15,00%	4,72%	21,00%	2,00%	14,00%	5,59%	15,00%	5,59%	15,00%	5,59%	15,00%	23,49%	57,00%
Implikacja Łukasiewicza	1,48%	12,00%	4,71%	20,00%	1,90%	13,00%	1,53%	12,00%	1,49%	12,00%	1,53%	12,00%	23,47%	57,00%
Implikacja Reichenbacha	1,48%	12,00%	4,71%	20,00%	1,90%	13,00%	1,53%	12,00%	1,49%	12,00%	1,53%	12,00%	23,47%	57,00%
Implikacja Kleene-Dienesa	1,48%	12,00%	4,71%	20,00%	1,90%	13,00%	1,53%	12,00%	1,49%	12,00%	1,53%	12,00%	23,47%	57,00%
Implikacja Zadeha	25,74%	29,00%	28,58%	40,00%	26,53%	28,00%	30,14%	34,00%	28,66%	29,00%	31,53%	34,00%	23,49%	57,00%
Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące	Dane testujące	Dane uczące

Do dalszej analizy wybrano następujące założenia mechanizmu wnioskowania: interpretacja reguł Larsena (iloczyn algebraiczny), t-norma Zadeha (min) jako spójnik logiczny AND. Dla tych parametrów otrzymano najlepsze wyniki dopasowania modelu wiedzy (tab. E-1).



Rys. E-2. Zależności procentu błędnych decyzji oraz liczby reguł elementarnych od wartości minimalnego wsparcia dla: a) danych uczących, b) danych testujących

Na podstawie zależności procentu błędnych decyzji i wartości minimalnego wsparcia (rys. E-2) można uznać wartość wsparcia równą $2 \cdot 10^{-4}$ jako wartość optymalną dla dopasowania modelu, przy danych warunkach założeniowych modelu. Wówczas wartości błędów dla danych uczących i testujących utrzymują się na zbliżonym poziomie (1,58% dla danych uczących, 13% dla danych testujących), jak dla pełnego rozkładu prawdopodobieństwa zdarzeń rozmytych w regułach a liczba reguł ulega zmniejszeniu się (z 2259 reguł elementarnych do 1793 reguł). Porównanie wyników symulacji oraz wartości wyliczonych z modelu dla wartości wsparcia równej $2 \cdot 10^{-4}$, przedstawia rysunek E-3.



Rys. E-3. Porównanie wyników symulacji z wartościami wyliczonymi na podstawie modelu z probabilistyczno-rozmytą bazą wiedzy (dla danych testujących)

Przykładowe reguły plikowe probabilistyczno-rozmytej bazy wiedzy (założenia: baza danych zdefiniowana wg rys. E-1, interpretacja reguł Larsena (iloczyn algebraiczny), t-norma Zadeha (min) jako spójnik logiczny AND, minimalne wsparcie $2 \cdot 10^{-4}$):

- 1: IF (L. zmiennych IS 2) AND (Śr. l. wartości ling. IS W) AND
(L. rozł. przedz. dyskr. IS M) AND (Min. w IS T) AND (L. danych IS BD) [0.0016]
THEN (Algorytm IS zm. Apr.) [1.0000]
- 2: IF (L. zmiennych IS 3) AND (Śr. l. wartości ling. IS W) AND
(L. rozł. przedz. dyskr. IS M) AND (Min. w IS T) AND (L. danych IS BD) [0.0016]
THEN (Algorytm IS zm. Apr.) [1.0000]
- 3: IF (L. zmiennych IS 4) AND (Śr. l. wartości ling. IS W) AND
(L. rozł. przedz. dyskr. IS M) AND (Min. w IS T) AND (L. danych IS BD) [0.0016]
THEN (Algorytm IS zm. Apr.) [1.0000]
- 4: IF (L. zmiennych IS 2) AND (Śr. l. wartości ling. IS M) AND
(L. rozł. przedz. dyskr. IS M) AND (Min. w IS T) AND (L. danych IS BD) [0.0016]
THEN (Algorytm IS zm. Apr.) [1.0000]
- 5: IF (L. zmiennych IS 2) AND (Śr. l. wartości ling. IS S) AND
(L. rozł. przedz. dyskr. IS M) AND (Min. w IS T) AND (L. danych IS BD) [0.0016]
THEN (Algorytm IS zm. Apr.) [1.0000]
- 6: IF (L. zmiennych IS 5) AND (Śr. l. wartości ling. IS W) AND
(L. rozł. przedz. dyskr. IS W) AND (Min. w IS T) AND (L. danych IS BD) [0.0015]
THEN (Algorytm IS zm. FP-G(P)) [1.0000]

- 7: IF (L. zmiennych IS 5) AND (Śr. l. wartości ling. IS M) AND
(L. rozł. przedz. dyskr. IS W) AND (Min. w IS T) AND (L. danych IS BD) [0.0015]
THEN (Algorytm IS zm. FP-G(P)) [1.0000]
- 8: IF (L. zmiennych IS 4) AND (Śr. l. wartości ling. IS W) AND
(L. rozł. przedz. dyskr. IS BW) AND (Min. w IS T) AND (L. danych IS BD) [0.0014]
THEN (Algorytm IS zm. Apr.) [0.5203]
ALSO (Algorytm IS zm. FP-G(P)) [0.4797]
- 9: IF (L. zmiennych IS 4) AND (Śr. l. wartości ling. IS S) AND
(L. rozł. przedz. dyskr. IS BW) AND (Min. w IS T) AND (L. danych IS BD) [0.0014]
THEN (Algorytm IS zm. Apr.) [0.6726]
ALSO (Algorytm IS zm. FP-G(P)) [0.3274]
- 10: IF (L. zmiennych IS 5) AND (Śr. l. wartości ling. IS M) AND
(L. rozł. przedz. dyskr. IS M) AND (Min. w IS Z) AND (L. danych IS BD) [0.0008]
THEN (Algorytm IS zm. Apr.) [0.6775]
ALSO (Algorytm IS zm. FP-G(P)) [0.3225]