

**Zastosowanie zbiorów rozmytych w
zadaniach sterowania ruchem obiektów i
prognozowaniu.**

Sławomir T. Wierzchoń
Instytut Podstaw Informatyki PAN
Uniwersytet Gdański, Instytut Informatyki

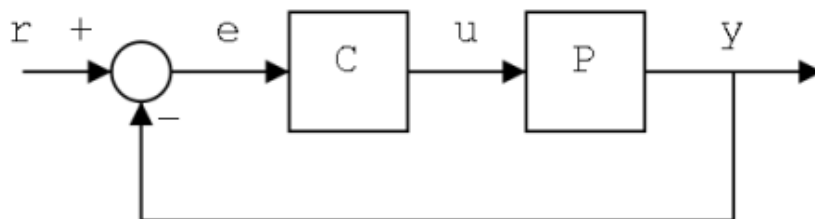
30 marca 2009

Spis treści

1	Wprowadzenie. Sterowanie i sterowniki	2
2	Architektura rozmytego sterownika	4
2.1	Elementy teorii zbiorów rozmytych	4
2.1.1	Podstawowe definicje	4
2.1.2	Funkcja przynależności	6
2.2	Baza reguł	7
2.3	Blok rozmywania	8
2.4	Blok wnioskowania	10
2.5	Blok wyostrzania	15
2.5.1	Sterownik Mamdaniego	15
2.5.2	Sterownik Takagi-Sugeno	17
2.6	Przykłady	17
2.6.1	Odwrócone wahadło	17
2.6.2	Automatyczny pilot	19
3	Podstawowy algorytm generowania reguł rozmytych	21
3.1	Podział wejściowej i wyjściowej przestrzeni na obszary	21
3.2	Tworzenie reguł rozmytych na podstawie przykładów	22
3.3	Przyporządkowanie stopni prawdziwości regułom	22
3.4	Utworzenie bazy rozmytych reguł	23
3.5	Wyostrzanie	23
4	Metody alternatywne	24
4.1	Generowanie uproszczonych reguł	24
4.1.1	Wnioskowanie z użyciem prostych reguł rozmytych	26
4.1.2	Identyfikacja prostych reguł rozmytych	26
4.1.3	Konwersja prostych reguł rozmytych	27
4.2	Zastosowanie algorytmów genetycznych	28
4.2.1	Genetyczna regresja	28
4.2.2	Uczenie reguł	29
5	Przykładowe problemy	32
5.1	Zadanie parkowania ciężarówki	32
5.2	Równanie Mackey-Glass’a	33
6	Związki między rozmytością a prawdopodobieństwem	36
6.1	Zbiory losowe	36
6.2	Konsonantne zbiory losowe	38
6.3	Próbkowanie ze zbioru rozmytego (losowego)	39
6.4	Losowe eksperymenty z nieprecyzyjnymi obserwacjami	40
7	Systemy neuro-rozmyte	41

1 Wprowadzenie. Sterowanie i sterowniki

Regulatory to urządzenia generujące sygnał sterujący w oparciu o tzw. *uchyb regulacji*, czyli różnicę między wartością zadaną sygnału regulowanego a wartością aktualnie zmierzoną – por. rysunek 13. Za przykłady mogą służyć regulatory ciśnienia, temperatury, czy regulatory natężenia przepływu.



Rysunek 1: Podstawowy schemat sterowania. Wyjście układu u jest porównywane z wartością zadaną r , a różnica tych wielkości, czyli uchyb regulacji e jest podana na wejście regulatora C . Dzięki temu regulator posiada informację o aktualnym błędzie sterowania i może tak zmieniać wartość wejściową u obiektu P , aby zmniejszyć uchyb do zera, gwarantując utrzymywanie wielkości wyjściowej układu możliwie blisko wartości zadanej.

W przypadku np. regulatorów temperatury można wyróżnić dwa rodzaje rozwiązań¹:

- Czujnik temperatury jest umieszczony wewnątrz zaworu i dzięki zmianie swych wymiarów powoduje przesuwanie grzybka względem gniazda.
- Czujnikiem jest termometr manometryczny dostarczający ciśnienia dla sprężystego mieszka, którego denko wprawia w ruch trzpień zaworu regulacyjnego.

Ogólnie, rolą regulatora jest doprowadzenia obiektu dożądanego stanu, ewentualnie poprawa niekorzystnych cech obiektu regulowanego lub też utrzymywanie obiektu w pożądanym stanie (jak to ma miejsce w przypadku regulatorów temperatury).

W automatyce przemysłowej² pod pojęciem regulatora (nazywanego też sterownikiem lub kontrolerem³) rozumie się samodzielny układ (komputerowy, elektryczny, elektroniczny, elektromechaniczny) zajmujący się nadzorowaniem pewnego urządzenia czy procesu technologicznego.

¹Por. np. http://www.sciaga.pl/tekst/34497-35-regulatory_bezposredniego_dzialania.

²Rozumianej jako dziedzina wiedzy, która zajmuje się możliwościami wyeliminowania lub ograniczenia udziału człowieka w czynnościach związanych ze sterowaniem różnorodnymi procesami, głównie technologicznymi i przemysłowymi.

³Nazwę tę wprowadza się, aby odróżnić je od programowych sterowników urządzeń w systemie operacyjnym.

Podstawowe typy regulatorów to: regulator proporcjonalny P , proporcjonalno-różniczkujący, PD oraz proporcjonalno-całkujący, PI , a ich uogólnieniem jest regulator proporcjonalno-całkująco-różniczkujący, PID ; składa się on z członu proporcjonalnego P o wzmacnieniu K_p , całkującego I o czasie zdwojenia T_i oraz różniczkującego D o czasie wyprzedzenia T_d . Realizuje on algorytm

$$u(t) = K_p \left[e(t) + \frac{1}{T_i} \int e(t) dt + T_d \frac{de(t)}{dt} \right] \quad (1)$$

gdzie $u(t)$ oznacza sygnał wyjściowy, a $e(t)$ – uchyb regulacji (sygnał wejściowy).

W dalszym ciągu będziemy używać nazwy *sterownik* dla odróżnienia regulatora implementowanego jako program komputerowy w odróżnieniu od wspomnianych wcześniej regulatorów realizowanych w postaci urządzeń mechanicznych.

Projektowanie sterownika wymaga znajomości modelu procesu P poddawanego sterowaniu. Jednakże znalezienie odpowiedniego modelu jest problemem czasochłonnym i trudnym, wymuszającym zazwyczaj przyjęcie szeregu założeń upraszczających.

Zastosowanie teorii zbiorów rozmytych do sterowania procesami technologicznymi (ang. *Fuzzy Logic Control*) nie wymaga szczegółowej znajomości modeli tych procesów⁴. Formuluje się jedynie reguły postępowania w postaci rozmytych zdań warunkowych typu **if** A **then** B , gdzie przesłanka A określa warunki jakie muszą być spełnione aby zastosować akcję przewidzianą w konkluzji B . Jeżeli np. sygnałami wejściowymi sterownika układu klimatyzacji są wartość temperatury i wilgotności w pomieszczeniu, a sygnałem wyjściowym jest intensywność chłodzenia to przykładowa reguła może mieć postać:

if [(temperatura jest *wysoka*) **and** (wilgotność jest *duża*)]
then (intensywność chłodzenia jest **duża**)

Sterowniki rozmyte są przykładami urządzeń działających w oparciu o wiedzę pozyskiwaną na jeden z czterech sposobów:

- odwołanie się do wiedzy eksperta z wykorzystaniem pomocy inżyniera wiedzy,
- modelowanie zachowania operatora procesu,
- wykorzystanie rozmytego modelu procesu,
- zastosowanie algorytmów uczących się i samoorganizujących

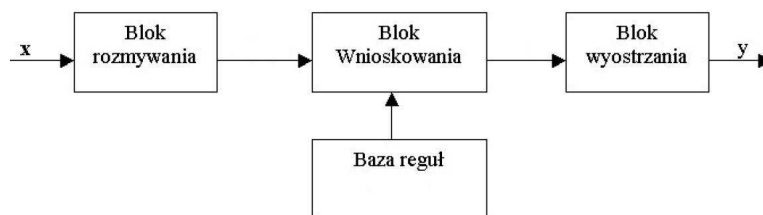
Z tego powodu sterowniki rozmyte traktowane są jako reprezentanci sterowania eksperckiego (ang. *Expert Control*).

⁴Wspomnijmy tu o pionierskiej pracy Mamdani, E.H., Assilian, S. An experiment in linguistic synthesis with fuzzy controller. *Internat. J. of Man-Machine Studies*. **7**, 1975, 1-13.

Sterowniki konstruowane z wykorzystaniem teorii zbiorów rozmytych znajdują obecnie szerokie zastosowanie ze względu na prostotę i taniość implementacji. Są wykorzystywane zarówno w prostych urządzeniach gospodarstwa domowego (pralki, lodówki, itp.) jak również w sterowaniu złożonymi procesami technologicznymi.

2 Architektura rozmytego sterownika

Schemat blokowy sterownika rozmytego przedstawiono na rysunku 2.



Rysunek 2: Schemat blokowy sterownika rozmytego.

W zastosowaniach wyróżnia się dwa typy sterowników rozmytych:

- (a) sterowniki Mamdaniego, w których zarówno przesłanki jak i konkluzje reguł mają charakter rozmyty,
- (b) sterowniki Takagi-Sugeno, w których przesłanki mają charakter rozmyty, a konkluzje – charakter funkcyjny, np.

$$\text{if } (x \text{ is } A) \text{ then } (u = f(x))$$

Poniżej omawiamy dokładniej poszczególne bloki tworzące sterownik rozmyty.

2.1 Elementy teorii zbiorów rozmytych

Zanim jednak przejdziemy do zasadniczego tematu, poświęćmy trochę uwagi sposobowi określania zbioru rozmytego i metodom definiowania służących temu celowi funkcji przynależności.

2.1.1 Podstawowe definicje

Zakładamy, że wyróżniono pewne uniwersum X ; reprezentuje ono np. zakres mierzonych wartości, czy zakres sygnałów wyjściowych. W tym wypadku $X \subset \mathbb{R}$. W zbiorze X możemy teraz wyróżniać pewne podzbiory A_1, \dots, A_n grupujące elementy o określonych własnościach. Typowym sposobem definiowania zbioru jest odwołanie się do pojęcia funkcji charakterystycznej $\chi_{A_i} : X \rightarrow \{0, 1\}$, która przypisuje elementom $x \in X$ należącym do zbioru

A_i (a więc elementom, posiadającym własność etykietowaną tym zbiorem) wartość $\chi_{A_i}(x) = 1$, a pozostałym elementom – wartość $\chi_{A_i}(x) = 0$.

W przypadku zbiorów rozmytych dopuszcza się, że elementy zaliczane do zbioru A_i posiadają interesującą nas własność jedynie w pewnym stopniu⁵. Ten stopień kwantyfikuje funkcja przynależności $\mu_{A_i}: X \rightarrow [0, 1]$. Jeżeli $\mu_{A_i}(x) = 1$ oznacza to, że element x bezwzględnie spełnia warunek pozwalający zaliczyć go do zbioru A_i , a jeżeli $\mu_{A_i}(x) = 0$, że element x absolutnie nie posiada danej własności. Istnieje jednak pokaźna liczba elementów spełniających warunek A_i jedynie w pewnym stopniu. Można np. powiedzieć, że jeżeli $\mu_{A_i}(x) > \frac{1}{2}$, to ten element *raczej* spełnia warunek A , natomiast te elementy, dla których $\mu_{A_i}(x) < \frac{1}{2}$ – raczej nie spełniają tego warunku.

Wprowadźmy pewne podstawowe definicje:

1. O funkcji przynależności zakłada się zazwyczaj, że istnieje przynajmniej jeden element x^* taki, że $\mu_A(x^*) = 1$. Mówimy wówczas, że dany zbiór rozmyty jest *normalny* (tzn. znormalizowany). Ogólnie, jeżeli przez $h(A) = \sup_{x \in X} \mu_A(x)$ oznaczymy *wysokość* zbioru rozmytego A , to zbiorem normalnym jest zbiór o wysokości $h(A) = 1$.
2. Zbiór $\{x \in X: \mu_A(x) > 0\}$ nazywamy nośnikiem zbioru A_i i oznaczamy symbolem $\text{supp}(A)$.
3. Zbiór rozmyty A jest pusty, co zapisujemy $A = \emptyset$, wtedy i tylko wtedy, gdy $\mu_A(x) = 0$ dla każdego $x \in X$. Zauważmy, że gdy $A = \emptyset$, to $\text{supp}(A) = \emptyset$.
4. Zbiór rozmyty A zawiera się w zbiorze rozmytym B , co zapisujemy $A \subset B$, wtedy i tylko wtedy, gdy $\mu_A(x) \leq \mu_B(x)$ dla każdego $x \in X$. Jeżeli $\mu_A(x) = \mu_B(x)$, to oba zbiory są identyczne, co oznaczamy $A \equiv B$.
5. α -przekrojem, $0 \leq \alpha \leq 1$, zbioru rozmytego nazywamy ostry zbiór $A_\alpha = \{x \in X: \mu_A(x) \geq \alpha\}$. Jeżeli $\alpha_1 < \alpha_2$ to $A_{\alpha_1} \supseteq A_{\alpha_2}$.
6. Zbiór $\text{core}(A) = A_1 = \{x \in X: \mu_A(x) = 1\}$ nazywamy jądrem zbioru rozmytego A . Jeżeli A jest zbiorem normalnym, to jego jądro zawiera przynajmniej jeden element.
7. Zbiór rozmyty $A \subset \mathbb{R}$ nazywamy wypukłym jeżeli

$$(\forall x_1, x_2 \in \mathbb{R})(\forall \lambda \in [0, 1])\mu_A[\lambda x_1 + (1 - \lambda)x_2] \geq \min[\mu_A(x_1), \mu_A(x_2)]$$

Podobnie, jeżeli

$$(\forall x_1, x_2 \in \mathbb{R})(\forall \lambda \in [0, 1])\mu_A[\lambda x_1 + (1 - \lambda)x_2] \leq \max[\mu_A(x_1), \mu_A(x_2)]$$

to zbiór rozmyty $A \subset \mathbb{R}$ nazywamy wklęsłym.

Jeżeli A jest zbiorem wypukłym (wklęsłym) to wszystkie jego α -odcinki są zbiorami wypukłymi (wklęsłymi).

8. Zbiór rozmyty $A \in \mathbb{R}$, który jednocześnie jest normalny, wypukły i którego funkcja przynależności jest funkcją przedziałami ciągłą nazywamy liczbą rozmytą.

⁵Aby odróżniać zbiory ostre od rozmytych, te ostatnie oznacza się często np. jako \tilde{A}_i . Aby nie komplikować notacji rezygnujemy z takiej konwencji.

2.1.2 Funkcja przynależności

Typowe postacie funkcji przynależności to (dla prostoty pomijamy indeks specyfikujący zbiór rozmyty):

1. Gaussowska funkcja przynależności $\mu(x) = \exp(-\frac{(x-x_0)^2}{2\sigma^2})$
2. Podobny kształt mają funkcje opisane równaniami

$$\mu(x) = \left[1 + \left(\frac{x-x_0}{\sigma}\right)^2\right]^{-1}$$

$$\mu(x) = 1 - \exp\left[-\left(\frac{\sigma}{x_0-x}\right)^a\right]$$

Dodatkowy parametr a użyty w drugim równaniu pozwala sterować nachyleniem ramion krzywej.

3. S -funkcja (przypominająca kształtem funkcję sigmoidalną); jej parametrami są lewy (x_l) i prawy (x_r) punkt przegięcia.

$$\mu^S(x; x_l, x_r) = \begin{cases} 0 & \text{gdy } x < x_l \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{x-x_l}{x_r-x_l}\pi\right) & \text{gdy } x_l \leq x \leq x_r \\ 1 & \text{gdy } x > x_r \end{cases}$$

Jej odbiciem jest Z -funkcja zdefiniowana następująco

$$\mu^Z(x; x_l, x_r) = \begin{cases} 1 & \text{gdy } x < x_l \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{x-x_l}{x_r-x_l}\pi\right) & \text{gdy } x_l \leq x \leq x_r \\ 0 & \text{gdy } x > x_r \end{cases}$$

Złożeniem obu tych funkcji przynależności jest tzw. π -funkcja,

$$\mu^\pi(x; x_1, x_2, x_3, x_4) = \min[\mu^S(x; x_1, x_2), \mu^Z(x; x_3, x_4)]$$

Duże znaczenie mają trapezoidalne funkcje przynależności opisane następująco

$$\mu(x; a, b, c, d) = \begin{cases} 0 & \text{gdy } x \leq a \\ \frac{x-a}{b-a} & \text{gdy } a \leq x \leq b \\ 1 & \text{gdy } b \leq x \leq c \\ \frac{d-x}{d-c} & \text{gdy } c \leq x \leq d \\ 0 & \text{gdy } x \geq d \end{cases} \quad (2)$$

W przypadku gdy $b = c$ otrzymujemy *trójkątną* funkcję przynależności opisaną trójką liczb (a, b, c) (dla prostoty przyjmuje się, że $b = \frac{a+c}{2}$, chociaż ogólnie powinien być spełniony jedynie warunek $a \leq b \leq c$). Stopień przynależności elementu x do zbioru rozmytego A opisanego taką funkcją przynależności μ_A oblicza się następująco:

$$\mu_A(x) = \begin{cases} 0 & \text{gdy } x \leq a \\ \frac{x-a}{b-a} & \text{gdy } a \leq x \leq b \\ \frac{c-x}{c-b} & \text{gdy } b \leq x \leq c \\ 0 & \text{gdy } x \geq c \end{cases} \quad (3)$$

O ile trójka (a, b, c) , gdzie $|ab| < \infty$, specyfikuje rozmyty przedział, o tyle trójki $(-\infty, b, c)$, (a, b, ∞) określają rozmyte półprzedziały. Np. gdy zbiór B ma postać $(-\infty, b, c)$, wówczas

$$\mu_B(x) = \begin{cases} 1 & \text{gdy } x \leq b \\ \frac{c-x}{c-b} & \text{gdy } b \leq x \leq c \\ 0 & \text{gdy } x \geq c \end{cases} \quad (4)$$

Praktyka pokazuje, że wybór kształtu funkcji przynależności jest zdecydowanie mniej istotny w porównaniu z prawidłowym określeniem jej nośnika. Instrukcja obsługi pakietu `TILShell` [7] rekomenduje co następuje:

- Rozpoczynaj od trójkątnych funkcji przynależności. Wszystkie używane funkcje powinny być symetrycznymi trójkątami o identycznej długości podstawy. Skrajny lewy trójkąt powinien mieć postać $\Delta(-\infty, a, b)$, zaś skrajny prawy – $\Delta(a, b, \infty)$.
- Ramiona sąsiadujących trójkątów powinny przecinać się w punkcie o współrzędnych $(\xi, 0.5)$. Pokrycie dziedziny każdej zmiennej powinno być tak dobierane, aby każdy jej element, z wyjątkiem skrajnych punktów, należał do (przynajmniej) dwóch zbiorów rozmytych.

2.2 Baza reguł

Baza reguł, nazywana też *modelem lingwistycznym* lub rozmytą pamięcią asocjacyjną (*FAM – Fuzzy Associative Memory*) składa się z rozmytych reguł $R^{(k)}$, $k = 1, \dots, N$, postaci

$$\begin{aligned} R^{(k)}: & \text{if } [(x_1 \text{ is } A_1^{(k)}) \text{ and } \dots \text{ and } (x_n \text{ is } A_n^{(k)})] \\ & \text{then } [(u_1 \text{ is } B_1^{(k)})] \text{ and } \dots \text{ and } (u_m \text{ is } B_m^{(k)}) \end{aligned} \quad (5)$$

gdzie $A_i^{(k)} \subset X_i \subset \mathbb{R}$, $i = 1, \dots, n$ są rozmytymi podzbiorami etykietującymi obserwowane wartości sygnałów wejściowych $x_i \in X_i$, a $B_j^{(k)} \subset Y_j \subset \mathbb{R}$, $j = 1, \dots, m$ są rozmytymi podzbiorami etykietującymi obserwowane wartości sygnałów wyjściowych $u_j \in Y_j$.

Zmienne wejściowe $(x_1, \dots, x_n)^T$ tworzą wektor obserwacji $\mathbf{x} \in \mathbf{X} = X_1 \times \dots \times X_n$; podobnie zmienne wyjściowe $(u_1, \dots, u_m)^T$ tworzą wektor sterowań $\mathbf{u} \in \mathbf{Y} = Y_1 \times \dots \times Y_m$.

Tworząc zbiór reguł (5) należy pokrywać dziedziny zmiennych X_1, \dots, X_n w taki sposób, aby zbiory rozmyte charakteryzujące etykiety używane w regułach pokrywały każdą obserwację, tzn. aby

$$(\forall x_i \in X_i)(\exists j \in \{1, \dots, k_i\}): \mu_{i,j}(x_i) \geq \delta$$

gdzie k_i to liczba zbiorów rozmytych pokrywających dziedzinę X_i , δ to tzw. stopień kompletności pokrycia dziedziny. Uwzględniając uwagi z poprzedniego punktu, przyjmujemy $\delta = 0.5$. O j -tym zbiorze rozmytym pokrywającym dziedzinę X_i zakłada się ponadto, że jest on normalny i wypukły.

Podobnie jak w klasycznej bazie wiedzy, przyjmuje się, że reguły $R^{(k)}$, $k = 1, \dots, N$ są powiązane ze sobą za pomocą spójnika logicznego **lub**. Ponadto zakładamy, że wyjścia u_1, \dots, u_m są wzajemnie niezależne. Bez utraty ogólności można wówczas rozpatrywać reguły z atomową (elementarną) konkluzją postaci

$$R^{(k)}: \text{if } [(x_1 \text{ is } A_1^{(k)}) \text{ and } \dots \text{ and } (x_n \text{ is } A_n^{(k)})] \\ \text{then } (u \text{ is } B^{(k)}) \quad (6)$$

W regule (6) część **if** nazywamy poprzednikiem, a część **then** następnikiem lub konkluzją. Poprzednik jest tzw. przesłanką *złożoną* stanowiącą połączenie za pomocą operatorów logicznych (w tym przypadku – operatora **and**) przesłanek *prostych* postaci $(x_i \text{ is } A_i^{(k)})$, $i = 1, \dots, n$.

Oznaczając przez $A^{(k)}$ iloczyn kartezjański zbiorów $A_1^{(k)}, \dots, A_n^{(k)}$ możemy regule (6) nadać postać rozmytej implikacji

$$R^{(k)}: A^{(k)} \rightarrow B^{(k)} \quad (7)$$

Regułę (6) można też interpretować jako relację rozmytą określoną na zbiorze $\mathbf{X} \times Y$ o funkcji przynależności

$$\mu_{R^{(k)}}(\mathbf{x}, u) = \mu_{A^{(k)} \rightarrow B^{(k)}}(\mathbf{x}, u) \quad (8)$$

Często blok zawierający definicje funkcji przynależności bywa nazywany bazą danych, a baza reguł wraz z bazą danych tworzą w takim kontekście bazę wiedzy.

2.3 Blok rozmywania

Skoro przesłanki reguły $R^{(k)}$ mają postać rozmytą, to konkretnej obserwacji $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n)$ należy przydzielić stopnie przynależności do odpowiednich zbiorów rozmytych wymienionych w przesłance reguły.

Najczęściej zakłada się, że obserwowane są dokładne wartości zmiennych wejściowych. W takim przypadku wartość $\mu_{A^{(k)}}(\bar{\mathbf{x}})$ interpretujemy jako stopień zgodności wektora obserwacji z rozmytą przesłanką k -tej reguły. Np. jeżeli zbiór $A^{(k)}$ opisany jest trójkątną funkcją przynależności (3) postaci $(-10, 10, 30)$ to obserwacji $\bar{x}_i = 25$ przypisany zostanie stopień $\mu_{A^{(k)}}(25) = 0.25$ kwantyfikujący prawdziwość stwierdzenia " $x_i = 25$ is $A^{(k)}$ ".

Znając zgodność obserwacji \bar{x}_i z elementarną przesłanką, tzn. warunkiem zadany w postaci zbioru rozmytego $A_i^{(k)}$ trzeba jeszcze wyznaczyć stopień spełnialności złożonej przesłanki $A^{(k)}$ przez wektor obserwacji $\bar{\mathbf{x}}$. Proces wyznaczania stopnia zgodności dla przesłanki złożonej w oparciu o znajomość stopni zgodności przesłanek prostych nazywa się *agregacją*. Jeżeli przesłanka jest złożona z n przesłanek prostych połączonych spójnikiem **and**, jak w

przypadku reguły (6), to dla danych wartości argumentów $x_i = \bar{x}_i$ stopień jej prawdziwości obliczamy jako stopień przynależności do relacji $A^{(k)}$:

$$\begin{aligned}\mu_{A^{(k)}}(\bar{x}_1, \dots, \bar{x}_n) &= \mu_{A_1^{(k)} \cap \dots \cap A_n^{(k)}}(\bar{x}_1, \dots, \bar{x}_n) \\ &= \mu_{A_1^{(k)}}(\bar{x}_1) * \dots * \mu_{A_n^{(k)}}(\bar{x}_n)\end{aligned}\quad (9)$$

gdzie $*$ jest symbolem normy trójkątnej (w skrócie T-normy).

Przypominamy, że funkcję $T: [0, 1] \times [0, 1] \rightarrow [0, 1]$ nazywamy T-normą, jeżeli dla dowolnych $a, b, c, d \in [0, 1]$ spełnia ona następujące warunki [14], [12], [10]:

- (a) $T(a, 0) = 0, T(a, 1) = a$ (ograniczoność)
- (b) $T(a, b) = T(b, a)$ (przemienność)
- (c) $T(T(a, b), c) = T(a, (b, c))$ (łączność)
- (d) $T(a, c) \leq T(b, d)$ gdy $a \leq b, c \leq d$ (monotoniczność)

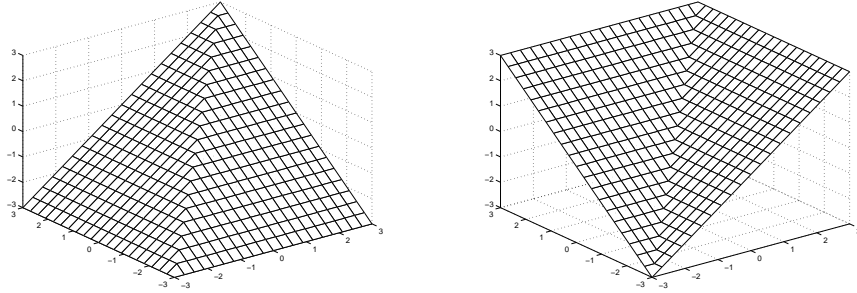
Dowolna T-norma jest ograniczona w sposób następujący

$$T_w(a, b) \leq T(a, b) \leq \min(a, b) \quad (10)$$

gdzie T_w jest T-normą postaci

$$T_w(a, b) = \begin{cases} a & \text{gdy } b = 1 \\ b & \text{gdy } a = 1 \\ 0 & \text{gdy } a, b \neq 1 \end{cases} \quad (11)$$

T-normę $\min(a, b)$ przedstawia lewa część rysunku 3. W teorii zbiorów rozmytych często zamiast $T(a, b)$ piszemy $a * b$.



Rysunek 3: Wykresy T-normy $\min(a, b)$ (lewy rysunek) oraz dualnej S-normy $\max(a, b)$ (prawy rysunek).

W przypadku (rzadziej stosowanych) przesłanek stanowiących alternatywę przesłanek prostych, stopień prawdziwości obliczamy jako stopień przynależności do relacji $A^{(k)} = A_1^{(k)} \text{ or } \dots \text{ or } A_n^{(k)}$:

$$\begin{aligned}\mu_{A^{(k)}}(\bar{x}_1, \dots, \bar{x}_n) &= \mu_{A_1^{(k)} \cup \dots \cup A_n^{(k)}}(\bar{x}_1, \dots, \bar{x}_n) \\ &= \mu_{A_1^{(k)}}(\bar{x}_1) \diamond \dots \diamond \mu_{A_n^{(k)}}(\bar{x}_n)\end{aligned}\quad (12)$$

gdzie \diamond jest symbolem ko-normy trójkątnej, nazywanej w skrócie S-normą.

Funkcję dwóch zmiennych $S: [0, 1] \times [0, 1] \rightarrow [0, 1]$ nazywamy S-normą jeżeli jest ona nierosnąca względem obu argumentów, spełnia ona wymienione wcześniej warunki przemienności i łączności oraz następujące warunki brzegowe

$$S(a, 0) = a, S(a, 1) = 1 \quad (13)$$

Dowolna S-norma jest ograniczona w sposób następujący

$$\max(a, b) \leq S(a, b) \leq S_w(a, b) \quad (14)$$

gdzie S_w jest S-normą postaci

$$S_w(a, b) = \begin{cases} a & \text{gdy } b = 0 \\ b & \text{gdy } a = 0 \\ 1 & \text{gdy } a, b \neq 0 \end{cases} \quad (15)$$

a S-normę $\max(a, b)$ przedstawiono w prawej części rysunku 3.

Działanie S-normy na argumentach a i b będziemy oznaczać $S(a, b) = a \diamond b$.

Każdej T-normie odpowiada dualna do niej S-norma, zgodnie z poniższym równaniem

$$a * b = 1 - [(1 - a) \diamond (1 - b)] \quad (16)$$

Najczęściej stosowane T-normy i S-normy przedstawiono w tabelicy 1.

Nr	$T(a, b)$	$S(a, b)$	parametry
1	$\min(a, b)$	$\max(a, b)$	
2	ab	$a + b - ab$	
3	$\max(a + b - 1, 0)$	$\min(a + b, 1)$	
4	$T_w(a, b)$	$S_w(a, b)$	
5	$\frac{ab}{\gamma + (1 - \gamma)(a + b - ab)}$	$\frac{a + b - (2 - \gamma)ab}{1 - (1 - \gamma)ab}$	$\gamma > 0$
6	$\frac{ab}{\max(a, b, \alpha)}$	$\frac{a + b - ab - \min(a, b, 1 - \alpha)}{\max(1 - a, 1 - b, \alpha)}$	$\alpha \in [0, 1]$
7	$\max\left(\frac{a + b - 1 + \lambda ab}{1 + \lambda}, 0\right)$	$\min(a + b + \lambda ab, 1)$	$\lambda \geq -1$

Tabela 1: Normy i ko-normy trójkątne

gdzie $*$ jest symbolem T-normy.

2.4 Blok wnioskowania

Zakładamy, że znamy zbiór rozmyty $\tilde{A} \subset \mathbf{X}$ i chcemy wyznaczyć zbiór rozmyty \tilde{B} odpowiadający przesłance reguły $R^{(k)}$. Do tego celu stosuje się uogólniony schemat *Modus Ponens* postaci

$$\begin{array}{c|c} \text{fakt} & \bar{\mathbf{x}} \text{ is } \tilde{A} \\ \text{reguła} & \text{if } (\mathbf{x} \text{ is } A^{(k)}) \text{ then } (u \text{ is } B^{(k)}) \\ \hline \text{konkluzja} & (\bar{u} \text{ is } \tilde{B}) \end{array} \quad (17)$$

Warunkiem jego stosowalności jest posiadanie przez regułę $R^{(k)}$ *własności interpolacyjnych* umożliwiających jej stosowanie także w tych przypadkach gdy obserwacja x jest *mniej więcej* równa $A^{(k)}$; wówczas wyciągane wnioski są także mniej więcej zgodne z rzeczywistością [10].

Formalizacją schematu (17) jest równanie

$$\mu_{\tilde{B}}(u) = \sup_{\mathbf{x} \in \mathbf{X}} [\mu_{A^{(k)}}(\mathbf{x}) * \mu_{A^{(k)} \rightarrow B^{(k)}}(\mathbf{x}, u)] \quad (18)$$

Najczęściej spotykane definicje rozmytej implikacji przedstawiono w tabelicy 2.

Nr	Nazwa	Definicja
1	reguła Mamdaniego	$\mu_{A \rightarrow B}(x, u) = \min[\mu_A(x), \mu_B(u)]$
2	reguła typu iloczyn (Larsena)	$\mu_{A \rightarrow B}(x, u) = \mu_A(x) \cdot \mu_B(u)$
3	reguła Łukasiewicza	$\mu_{A \rightarrow B}(x, u) = \min[1, 1 - \mu_A(x) + \mu_B(u)]$
4	reguła typu max-min	$\mu_{A \rightarrow B}(x, u) = \max \left\{ \min[\mu_A(x), \mu_B(u)], 1 - \mu_A(x) \right\}$
5	reguła probabilistyczna	$\mu_{A \rightarrow B}(x, u) = \min[1, 1 - \mu_A(x) + \mu_A(x)\mu_B(u)]$
6	reguła ograniczonej sumy	$\mu_{A \rightarrow B}(x, u) = \min[1, \mu_A(x) + \mu_B(u)]$
7	reguła binarna	$\mu_{A \rightarrow B}(x, u) = \max[1 - \mu_A(x), \mu_B(u)]$
8	reguła Goguena	$\mu_{A \rightarrow B}(x, u) = \min[1, \frac{\mu_B(x)}{\mu_A(u)}]$
9	reguła Gödela	$\mu_{A \rightarrow B}(x, u) = \begin{cases} 1 & \text{gdy } \mu_A(x) \leq \mu_B(u) \\ \mu_B(u) & \text{gdy } \mu_A(x) > \mu_B(u) \end{cases}$
10	regułą Sharpa	$\mu_{A \rightarrow B}(x, u) = \begin{cases} 1 & \text{gdy } \mu_A(x) \leq \mu_B(u) \\ 0 & \text{gdy } \mu_A(x) > \mu_B(u) \end{cases}$

Tabela 2: Najczęściej spotykane definicje rozmytej implikacji

Podsumowując materiał tej sekcji można sformułować następującą procedurę wnioskowania w oparciu o zbiór N reguł i wektor obserwacji $\bar{\mathbf{x}}^T = (\bar{x}_1, \dots, \bar{x}_n)^T$.

1. Dla każdej obserwacji \bar{x}_i wyznaczyć stopień zgodności $\tau_i^{(k)} = \mu_{A_i^{(k)}}(\bar{x}_i)$ tej obserwacji ze zbiorem rozmytym występującym w i -tej elementarnej przesłance k -tej reguły.
2. (*Agregacja*) Dla każdej reguły wyznaczyć stopień zgodności obserwacji $\bar{\mathbf{x}}^T$ z przesłanką k -tej reguły, $\tau^{(k)} = \tau_1^{(k)} * \dots * \tau_n^{(k)}$.
3. (*Aktywacja*) Dla każdej reguły wyznaczyć, w oparciu o wybraną definicję implikacji, postać zbioru rozmytego opisującego k -tą konkluzję.
4. (*Akumulacja*) Zastosować wybraną S -normę (najczęściej max) do utworzenia wynikowej funkcji przynależności opisującej rozmytą konkluzję.

Odnotujmy, że wynikowa funkcja przynależności, o której mowa w punkcie (4) może być uzyskiwana na dwa sposoby [2]:

- Złożenie FITA (ang. *First Inference Then Aggregate*), gdzie najpierw każda reguła jest aktywowana wejściem rozmytym, a następnie poszczególne wyniki są łączone, celem stworzenia wynikowej funkcji przynależności opisującej końcową konkluzję.
- Złożenie FATI (ang. *First Aggregate Then Inference*), gdzie najpierw łączy się wszystkie reguły bazy wiedzy, a następnie aktywuje je wejściem rozmytym.

W [2] wskazano, że wyniki uzyskiwane metodą FITA są bardziej dokładne, a ponadto ta metoda akumulacji jest prostsza w implementacji. Operatorem umożliwiającym akumulację może być zarówno S -norma jak i uogólniony operator uśredniania:

$$\bigsqcup_{i=1}^N(\alpha, \xi_i) = \left(\frac{1}{N} \sum_{i=1}^N \xi_i^\alpha \right)^{\frac{1}{\alpha}} = \begin{cases} \min_{i=1, \dots, N} \xi_i & \text{gdy } \alpha \rightarrow -\infty \\ \left(\frac{1}{N} \sum_{i=1}^N \frac{1}{\xi_i} \right)^{-1} & \text{gdy } \alpha \rightarrow -1 \\ \left(\prod_{i=1}^N \xi_i \right)^{1/N} & \text{gdy } \alpha \rightarrow 0 \\ \frac{1}{N} \sum_{i=1}^N \xi_i & \text{gdy } \alpha = 1 \\ \sqrt{\frac{1}{N} \sum_{i=1}^N \xi_i^2} & \text{gdy } \alpha = 2 \\ \max_{i=1, \dots, N} \xi_i & \text{gdy } \alpha \rightarrow \infty \end{cases} \quad (19)$$

Poniżej przedstawiamy konkretyzacje reguły (18) dla wybranych definicji implikacji. We wszystkich przykładach będziemy zakładać, że obserwowany wektor wejść ma postać $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2)$ a rozmyty sterownik opisują dwie reguły postaci

$$\begin{aligned} R^{(1)}: & \text{ if } [(x_1 \text{ is } A_1^{(1)}) \text{ and } (x_2 \text{ is } A_2^{(1)})] \text{ then } (u \text{ is } B^{(1)}) \\ R^{(2)}: & \text{ if } [(x_1 \text{ is } A_1^{(2)}) \text{ and } (x_2 \text{ is } A_2^{(2)})] \text{ then } (u \text{ is } B^{(2)}) \end{aligned}$$

W tym przypadku reguła (18) przyjmuje postać

$$\mu_{B^{(k)}}(u) = \sup_{X_1, x_2 \in X_1 \times X_2} [\mu_{A_1' \times A_2'}(x_1, x_2) * \mu_{R^{(k)}}(x_1, x_2, u)] \quad (20)$$

a funkcję przynależności wynikowego zbioru \tilde{B} wyznaczymy z równania

$$\mu_{\tilde{B}}(u) = \mu_{B^{(1)}}(u) \diamond \mu_{B^{(2)}}(u) \quad (21)$$

Przykład 2.1. Niech $*$ = min, \diamond = max i niech do zdefiniowania rozmytej implikacji posłuży reguła Mamdaniego. W takim przypadku reguła (9) przybiera postać (agregacja)

$$\mu_{A^{(k)}}(\bar{x}_1, \bar{x}_2) = \mu_{A_1^{(k)} \times A_2^{(k)}}(\bar{x}_1, \bar{x}_2) = \min[\mu_{A_1^{(k)}}(\bar{x}_1), \mu_{A_2^{(k)}}(\bar{x}_2)] \quad (22)$$

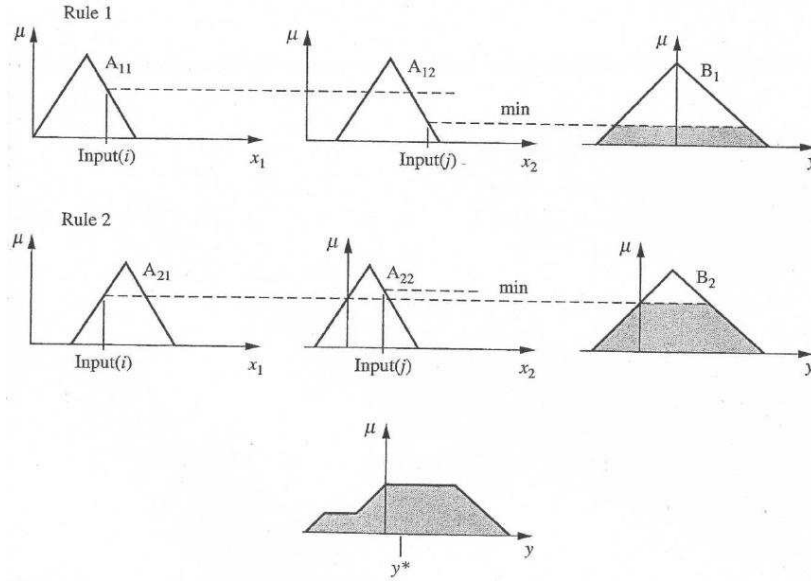
Stosując regułę Mamdaniego otrzymujemy (aktywacja)

$$\begin{aligned} \mu_{\tilde{B}^{(k)}}(u) &= \min \left\{ \min[\mu_{A_1^{(k)}}(\bar{x}_1), \mu_{A_2^{(k)}}(\bar{x}_2)], \mu_{B^{(k)}}(u) \right\} \\ &= \min[\mu_{A_1^{(k)}}(\bar{x}_1), \mu_{A_2^{(k)}}(\bar{x}_2), \mu_{B^{(k)}}(u)] \end{aligned} \quad (23)$$

i ostatecznie (akumulacja)

$$\mu_{\tilde{B}}(u) = \max_{k=1,2} \left\{ \min[\mu_{A_1^{(k)}}(\bar{x}_1), \mu_{A_2^{(k)}}(\bar{x}_2), \mu_{B^{(k)}}(u)] \right\} \quad (24)$$

Graficzną ilustrację tego procesu przedstawiono na rysunku 4 □



Rysunek 4: Graficzna reprezentacja procesu wnioskowania w oparciu o dwie reguły, których przesłanki złożone są z dwóch warunków elementarnych. Pierwsze dwie linie reprezentują agregację stopni spełnialności przesłanek oraz aktywację poszczególnych reguł, a w trzeciej linii przedstawiono akumulację uzyskanych wyników.

Przykład 2.2. Załóżmy teraz, że implikację definiujemy za pomocą iloczynu (por. regułę 2 w tablicy 2), a definicje T-normy i S-normy nie ulegają zmianie. Stosując rozumowanie z poprzedniego przykładu równanie (23)

przyjmuje teraz postać

$$\mu_{\tilde{B}^{(k)}}(u) = \mu_{B^{(k)}}(u) \cdot \min[\mu_{A_1^{(k)}}(\bar{x}_1), \mu_{A_2^{(k)}}(\bar{x}_2)] \quad (25)$$

Wyznaczenie kształtu ostatecznego wyjściowego zbioru rozmytego odbywa się jak poprzednio. \square

Przykład 2.3. Przyjmijmy, że zarówno T-norma jak i implikacja zdefiniowane są za pomocą iloczynu (definicja S-normy nie ulega zmianie). Stosując rozumowanie z poprzednich przykładów znajdujemy najpierw konkretyzację reguły (9):

$$\mu_{A^{(k)}}(\bar{x}_1, \bar{x}_2) = \mu_{A_1^{(k)} \times A_2^{(k)}}(\bar{x}_1, \bar{x}_2) = \mu_{A_1^{(k)}}(\bar{x}_1) \cdot \mu_{A_2^{(k)}}(\bar{x}_2) \quad (26)$$

Odnajdujemy, że taki sposób agregacji jest najczęściej stosowany w praktyce. Równanie 2 z tablicy 2 przyjmuje teraz postać

$$\mu_{\tilde{B}^{(k)}}(u) = \mu_{B^{(k)}}(u) \cdot \mu_{A_1^{(k)}}(\bar{x}_1) \cdot \mu_{A_2^{(k)}}(\bar{x}_2) \quad (27)$$

Wyznaczenie kształtu ostatecznego wyjściowego zbioru rozmytego odbywa się jak w poprzednich przykładach. \square

Przykład 2.4. Dotychczas rozważaliśmy dokładne pomiary \bar{x}_1, \bar{x}_2 . W praktyce dopuszczalnie jest ich rozmywanie. Załóżmy więc, że obie obserwacje zadane są w postaci zbiorów rozmytych A_1, A_2 i powróćmy do sytuacji rozpatrywanej w przykładzie 2.1.

Tym razem kwantyfikacja stopnia prawdziwości faktu x_i is $A_i^{(k)}$ wyznaczana jest z równania

$$\tau_{A_i^{(k)}} = \sup_{x_i \in X_i} \min[\mu_{A_i^{(k)}}(x_i), \mu_{A_i}(x_i)] \quad (28)$$

a stopień prawdziwości przesłanki jest równy

$$\tau_{A^{(k)}} = \sup_{x_1 \in X_1} [\mu_{A_1^{(k)}}(x_1) \wedge \mu_{A_1}(x_1)] \wedge \sup_{x_2 \in X_2} [\mu_{A_2^{(k)}}(x_2) \wedge \mu_{A_2}(x_2)] \quad (29)$$

gdzie $a \wedge b = \min(a, b)$. Ostateczny kształt funkcji przynależności wynikowego zbioru rozmytego \tilde{B} wyznaczamy z równania

$$\mu_{\tilde{B}}(u) = \max_{k=1,2} \left\{ \tau_{A^{(k)}} \wedge \mu_{B^{(k)}}(u) \right\} \quad (30)$$

Definiując – jak w poprzednim przykładzie – T-normę za pomocą iloczynu, powyższe równanie przyjmuje postać

$$\mu_{\tilde{B}}(u) = \max_{k=1,2} \left\{ \sup_{x_1} [\mu_{A_1^{(k)}}(x_1) \mu_{A_1}(x_1)] \cdot \sup_{x_2} [\mu_{A_2^{(k)}}(x_2) \mu_{A_2}(x_2)] \cdot \mu_{B^{(k)}}(u) \right\} \quad (31)$$

\square

2.5 Blok wyostrozania

Pod pojęciem wyostrozania rozumiemy zastąpienie zbioru rozmytego o funkcji przynależności $\mu_{\tilde{B}}$ pojedynczą wartością $\bar{u} \in Y$, która będzie wyznaczonym sterowaniem. Istnieje wiele metod wyostrozania. Poniżej omawiamy najbardziej typowe. Inne metody omawiane są np. w pracy [10].

2.5.1 Sterownik Mamdaniego

Przypominamy, że w przypadku sterownika Mamdaniego konkluzja ma postać stwierdzenie "u is \tilde{B} ", a zadanie polega na znalezieniu pojedynczej wartości \bar{u} , która należy przyjąć za wartość sterowania.

Najczęściej stosowaną metodą wyostrozania jest metoda środka ciężkości (*COG – center of gravity*), w której za wartość \bar{u} przyjmuje się odciętą współrzędną środka ciężkości powierzchni zawartej pod wykresem funkcji przynależności $\mu_{\tilde{B}}$:

$$\bar{u} = \frac{\int \mu_{\tilde{B}}(u) \cdot u du}{\int \mu_{\tilde{B}}(u) du} \quad (32)$$

W przypadku dyskretnej funkcji przynależności powyższe równanie przyjmuje postać $\mu_{\tilde{B}}$:

$$\bar{u} = \frac{\sum_i \mu_{\tilde{B}}(u_i) \cdot u_i}{\sum_i \mu_{\tilde{B}}(u_i)} \quad (33)$$

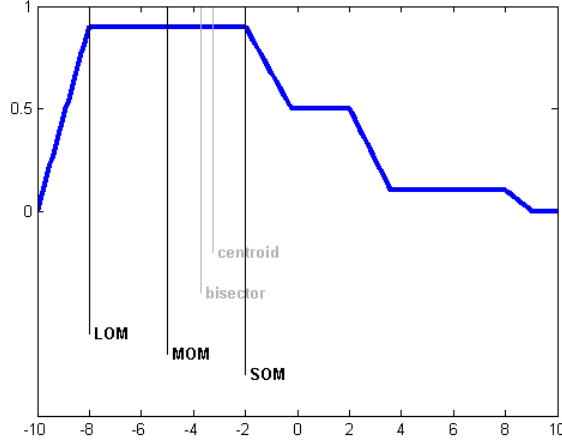
Zaletą tej metody jest uwzględnienie wpływu wszystkich funkcji przynależności występujących w konkluzjach zaktywizowanych reguł. Jej wadą jest stosunkowo duży nakład obliczeniowy związany z całkowaniem powierzchni o nieregularnym kształcie. W przypadku trapezoidalnych lub trójkątnych funkcji przynależności można stosować proste metody całkowania numerycznego do wyznaczenia wartości \bar{u} .

Stosunkowo proste – z numerycznego punktu widzenia są metody polegające na wyborze takiej wartości \bar{u} , dla której funkcja $\mu_{\tilde{B}}$ osiąga swoje maksimum. Wyróżnia się tu trzy warianty:

- (a) metodę pierwszego maksimum (*FOM – first of maxima* lub *SOM – smallest of maxima*), w której za \bar{u} przyjmuje się najmniejszą wartość u odpowiadającą maksymalnemu stopniowi przynależności $\mu_{\tilde{B}}(u)$;
- (b) metodę ostatniego maksimum (*LOM – last, lub largest of maxima*), w której za \bar{u} przyjmuje się największą wartość u odpowiadającą maksymalnemu stopniowi przynależności $\mu_{\tilde{B}}(u)$;
- (c) metodę średniego maksimum (*MOM – middle of maxima* wybierającą centralny punkt zbioru tych wartości u , dla których funkcja przynależności $\mu_{\tilde{B}}(u)$ przyjmuje wartość maksymalną.

Wadą tych metod jest uwzględnianie wyłącznie konkluzji najbardziej zaktywizowanej reguły.

Graficzne porównanie wymienionych tu metod przedstawiono na rysunku 5. Inne metody wyostrzania, to:



Rysunek 5: Porównanie metod wyostrzania.

1. Metoda środka sum (*CSD - center of sums defuzzification*). Wartość \bar{u} obliczana jest tu następująco

$$\bar{u} = \frac{\int u \sum_{k=1}^N \mu_{\tilde{B}^{(k)}}(u) du}{\int \sum_{k=1}^N \mu_{\tilde{B}^{(k)}}(u) du} \quad (34)$$

2. Metoda bisekcji (*BOA - bisector of area*), w której za \bar{u} wybiera się taką wartość odciętej, która dzieli powierzchnię pod krzywą $\mu_{\tilde{B}}$ na połowy, tzn.

$$\int_{u_{min}}^{\bar{u}} \mu_{\tilde{B}}(u) du = \int_{\bar{u}}^{u_{max}} \mu_{\tilde{B}}(u) du \quad (35)$$

3. Metoda *CAD (center average defuzzification)*. Wartość \bar{u} wyznacza się ze wzoru

$$\bar{u} = \frac{\sum_{k=1}^N \mu_{\tilde{B}^{(k)}}(\bar{u}^k) \cdot \bar{u}^k}{\sum_{k=1}^N \mu_{\tilde{B}^{(k)}}(\bar{u}^k)} \quad (36)$$

gdzie $\bar{u}^k = \operatorname{argmax}_u \mu_{\tilde{B}^{(k)}}(u)$. Punkt \bar{u}^k nazywany jest środkiem zbioru rozmytego $\tilde{B}^{(k)}$. Nie zależy on ani od kształtu ani od nośnika funkcji przynależności $\mu_{\tilde{B}^{(k)}}$.

4. Metoda *MFP (maksimum funkcji przynależności)*. Jeżeli $\mu_{\tilde{B}}$ jest funkcją jedno-modalną, to wartość \bar{u} wyznacza się ze wzoru

$$\bar{u} = \sup_u \mu_{\tilde{B}}(u) \quad (37)$$

2.5.2 Sterownik Takagi-Sugeno

W przypadku tego sterownika konkluzja k -tej reguły ma postać $u = f^{(k)}(x_1, \dots, x_n)$. W celu wyznaczenia wartości sterowania \bar{u} postępujemy następująco:

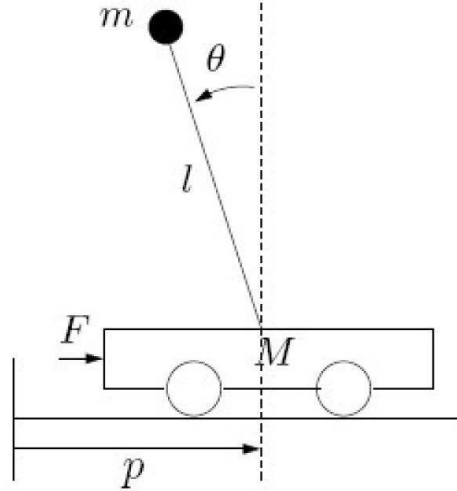
1. Dla każdej reguły wyznaczamy stopień zgodności $w^{(k)}$ obserwacji $(\bar{x}_1, \dots, \bar{x}_n)$ z przesłankami k -tej reguły, $w^{(k)} = \mu_{A_1^{(k)}}(\bar{x}_1) * \dots * \mu_{A_n^{(k)}}(\bar{x}_n)$, gdzie $*$ jest symbolem T-normy (jest to zazwyczaj min lub iloczyn). Obliczamy także wartość $\bar{u}^{(k)} = f^{(k)}(x_1, \dots, x_n)$.
2. Poszukiwaną wartość \bar{u} wyznaczamy z równania

$$\bar{u} = \frac{\sum_{k=1}^N w^{(k)} \bar{u}^{(k)}}{\sum_{k=1}^N w^{(k)}} \quad (38)$$

2.6 Przykłady

2.6.1 Odwrócone wahadło

Układ składa się z wózka i zamocowanego na elastycznym przegubie pionowego wahadła o masie m i długości l – por. rysunek 6.



Rysunek 6: Problem sterowania odwróconym wahadłem.

Zachowanie systemu opisuje równanie różniczkowe

$$-ml^2 \frac{d^2\theta}{dt^2} + mlg \cdot \sin \theta = \tau = u(t)$$

gdzie $\tau = u(t)$ jest sterowaniem (moment obrotowy), które należy przyłożyć w kierunku przeciwnym do wychylenia θ aby utrzymać wahadło w pozycji pionowej.

Oznaczając $x_1 = \theta$, $x_2 = d\theta/dt$ sprowadzamy powyższe równanie do układu równań opisującego zachowanie badanego układu w przestrzeni stanów

$$\begin{aligned} dx_1/dt &= x_2 \\ dx_2/dt &= (g/l) \sin x_1 - (1/ml^2)u(t) \end{aligned}$$

Dla małych wartości θ mamy $\sin \theta \approx \theta$, co pozwala zlinearyzować powyższy układ

$$\begin{aligned} dx_1/dt &= x_2 \\ dx_2/dt &= (g/l)x_1 - (1/ml^2)u(t) \end{aligned}$$

Jeżeli x_1 jest mierzone w stopniach, a x_2 w stopniach na sekundę (*sns*), to przyjmując $l = g$, $m = 180/(\pi \cdot g^2)$ uzyskujemy następujący różnicowy odpowiednik powyższego układu

$$\begin{aligned} x_1(t+1) &= x_1(t) + x_2(t) \\ x_2(t+1) &= x_1(t) + x_2(t) - u(t) \end{aligned}$$

Przyjmijmy, że $-2^\circ \leq x_1 \leq 2^\circ$, $-5sns \leq x_2 \leq 5sns$. Konstrukcja sterownika składa się z następujących kroków.

1. Dziedzinę zmiennej wejściowej x_1 pokrywamy trzema zbiorami rozmytymi: $\mu_{NEG} = \Delta(-\infty, -2, 0)$, $\mu_{Zero} = \Delta(-2, 0, 2)$, $\mu_{POS} = \Delta(0, 2, \infty)$. Podobnie, trzema zbiorami rozmytymi, lecz o innych nośnikach, pokrywamy dziedzinę zmiennej x_2 : $\mu_{NEG} = \Delta(-\infty, -5, 0)$, $\mu_{Zero} = \Delta(-5, 0, 5)$, $\mu_{POS} = \Delta(0, 5, \infty)$.
2. Pokrywamy dziedzinę zmiennej sterującej u pięcioma funkcjami przynależności: $\mu_{NB} = \Delta(-\infty, -24, -16)$, $\mu_N = \Delta(-16, -8, 0)$, $\mu_Z = \Delta(-8, 0, 8)$, $\mu_P = \Delta(0, 8, 16)$, $\mu_{PB} = \Delta(16, 24, \infty)$.
3. Tworzymy reguły sterujące. Wygodną formą ich zapisu jest tzw. rozmyta pamięć asocjacyjna (*FAM*) – tablica 3: Wiersze tablicy odpo-

	NEG	Zero	POS
NEG	NB	N	Z
Zero	N	Z	P
POS	Z	P	PB

Tabela 3: Baza rozmytych reguł

wiadają rozmytym wartościom zmiennej x_1 , a kolumny – rozmytym wartościom zmiennej x_2 . Element znajdujący się na przecięciu wiersza i kolumny odpowiada konkluzji reguły, której przesłankami są odpowiednie wartości lingwistyczne zmiennych wejściowych. Przykładowo, przecięciu pierwszego wiersza i trzeciej kolumny odpowiada reguła **if** x_1 **is** NEG **and** x_2 **is** Pos **then** u **is** Z.

4. Właściwe sterowanie. Przypuśćmy, że w chwili $t = 0$ wahadło znajduje się w punkcie $x_1 = 1^\circ$, $x_2 = -4\text{stn}$. Punkt ten uaktywnia następujące reguły

if x_1 is Pos and x_2 is Zero then u is P $\min(0.5, 0.2) = 0.2$ (P)
 if x_1 is Pos and x_2 is Neg then u is Z $\min(0.5, 0.8) = 0.5$ (Z)
 if x_1 is Zero and x_2 is Zero then u is Z $\min(0.5, 0.2) = 0.2$ (Z)
 if x_1 is Zero and x_2 is Neg then u is N $\min(0.5, 0.8) = 0.5$ (N)

Ostatecznie konkluzji N przypisujemy stopień 0.5, konkluzji Z stopień $\max(0.5, 0.2) = 0.5$ i konkluzji P stopień 0.2. Wyostrzając wynikowy zbiór rozmyty stwierdzamy, że $u = -2$. Możemy więc wyznaczyć kolejne położenie wahadła w przestrzeni stanów

$$\begin{aligned}
 x_1(1) &= x_1(0) + x_2(0) = 1 - 4 = -3 \\
 x_2(1) &= x_1(0) + x_2(0) - u(0) = 1 - 4 - (-2) = -1
 \end{aligned}$$

Powtarzamy teraz procedurę wyznaczania kolejnego sterowania.

2.6.2 Automatyczny pilot

Rozważmy teraz symulację procesu lądowania samolotu. Interesują nas dwie zmienne: wysokość samolotu h i jego prędkość v . Równania opisujące dynamikę procesu wyprowadza się stosując następujące rozumowanie. Ciało o masie m poruszające się z prędkością v posiada moment $p = mv$. Przy braku działania zewnętrznych sił, ciało to będzie poruszać się nadal ze stałą prędkością v . Dodatkowe oddziaływanie siłą f w przedziale czasu Δt skutkuje zmianą prędkości $\Delta v = f\Delta t/m$. Przyjmując $\Delta t = 1.0$ oraz $m = 1.0$ (i dobierając odpowiednio jednostki miary), otrzymujemy $\Delta v = f$, co oznacza, że zmiana prędkości jest proporcjonalna do użytej siły. Otrzymujemy więc następujący układ równań:

$$\begin{aligned}
 v(t+1) &= v(t) + f(t) \\
 h(t+1) &= h(t) + v(t)
 \end{aligned}$$

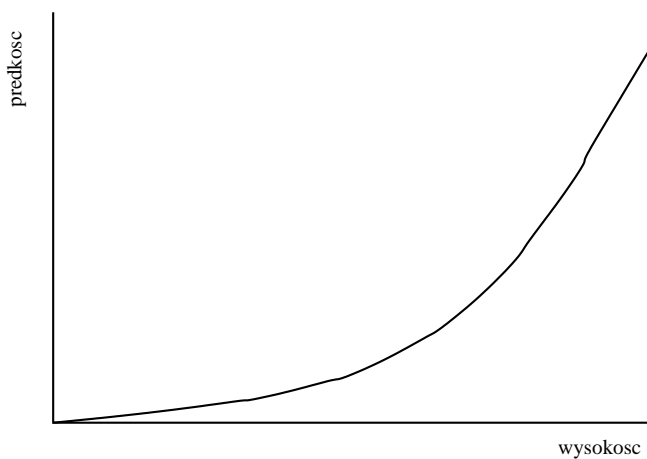
Przyjmijmy, że wysokość samolotu zmienia się w przedziale $[0, 1000]$. Pokrywamy dziedzinę zmiennej h czterema zbiorami rozmytymi: **Near Zero** o funkcji przynależności $\Delta(0, 0, 500)$, **Small** o funkcji przynależności $\Delta(-200, 300, 800) \cdot \mathbb{1}(x)$, gdzie $\mathbb{1}(x)$ jest funkcją przyjmującą wartość 0 dla $x \leq 0$ i 1 gdy $x > 0$, **Medium** o funkcji przynależności $\Delta(300, 800, 1300) \cdot (1 - \mathbb{1}(x - 1000))$ oraz **Large** o trójkątnej funkcji $\Delta(500, 1000, 1000)$.

Założmy dalej, że prędkość samolotu zmienia się w przedziale $[-30, 30]$ pokrytym pięcioma zbiorami rozmytymi, których definicje przedstawiono w Tablicy 4. Podobnie zmienia się wartości siły f (sterowania), tzn. $f \in [-30, 30]$. Dlatego dziedzinę tej funkcji pokryto również pięcioma zbiorami rozmytymi o identycznych, jak poprzednio, funkcjach przynależności.

etykieta	definicja
Down Large (DL)	$\Delta(-\infty, -20, -10)$
Down Small (DS)	$\Delta(-20, -10, 0)$
Zero (Z)	$\Delta(-10, 0, 10)$
Up Small (US)	$\Delta(0, 10, 20)$
Up Large (UL)	$\Delta(10, 20, \infty)$

Tabela 4: Pokrycie zbiorami rozmytymi dziedziny zmiennej v

Należy tak sterować samolotem, aby jego trajektoria ruchu wyglądała jak na poniższym rysunku.



Rysunek 7: Pożądana trajektoria lotu w przestrzeni fazowej (h, v)

Proponowana rozmyta pamięć asocjacyjna (FAM), reprezentująca reguły sterujące zapewniające uzyskanie powyższej trajektorii, ma następującą postać:

wysokość h	prędkość v				
	DL	DS	Zero	US	UL
L	Z	DS	DL	DL	DL
M	US	Z	DS	DL	DL
S	UL	US	Z	DS	DL
NZ	UL	UL	Z	DS	DS

Tabela 5: Rozmyta pamięć asocjacyjna

Dla przykładu założmy, że w chwili $t = 0$ mamy $h = 1000$, $v = -20$. Wyznaczając wartość siły f za pomocą metody z poprzedniego punktu i wyostrzając wynikowy zbiór rozmyty metodą środka ciężkości znajdujemy, że $f = 5.8$.

Problem: Przeprowadzić symulację lądowania samolotu rozpoczynając od podanych wyżej wartości wysokości i prędkości. Wykonać stosowny wykres. Porównać wyniki z tymi, uzyskanymi przez zastosowanie metody średniego maksimum (*MOM*).

3 Podstawowy algorytm generowania reguł rozmytych

Pochodzi on z pracy [16] (jego opis znajduje się także w książce [12]). Jej autorzy zauważyli, że sterowanie rozmyte jest znakomitą metodą pozwalającą korzystać z opisowej wiedzy o problemie (model lingwistyczny). Z drugiej strony podejścia bazujące na zastosowaniu sieci neuronowych w sterowaniu odwołują się do wiedzy reprezentowanej w postaci par (*wejście*, *wyjście*), gdzie *wejście* to zbiór obserwowanych wartości wejściowych, a *wyjście* – odpowiadające im wartości zmiennych wyjściowych. A więc sterowniki rozmyte korzystają z reguł lingwistycznych, a sterowniki neuronowe – z danych numerycznych.

Przedstawiane dalej podejście stawia sobie za cel połączenie obu paradygmatów, tzn. dysponując danymi numerycznymi konwertuje się je do postaci reguł rozmytych, które mogą być z kolei integrowane ze skonstruowanym wcześniej modelem lingwistycznym w celu utworzenia wspólnej bazy wiedzy pozwalającej generować wartości sterowań.

W dalszym ciągu będziemy zakładać, jak w obu cytowanych pracach, że chcemy utworzyć bazę reguł dla układu rozmytego o dwóch wejściach i jednym wyjściu. Dane uczące zadane są w postaci m trójek

$$\begin{aligned} & (x_1^{(1)}, x_2^{(1)}, u^{(1)}) \\ & (x_1^{(2)}, x_2^{(2)}, u^{(2)}) \\ & \dots \\ & (x_1^{(m)}, x_2^{(m)}, u^{(m)}) \end{aligned} \tag{39}$$

gdzie $x_1^{(i)}, x_2^{(i)}$ to wartości sygnałów wejściowych zarejestrowanych w i -tym przykładzie, a $u^{(i)}$ to odpowiadająca im wartość sygnału wyjściowego.

Poniżej omawiamy główne kroki algorytmu.

3.1 Podział wejściowej i wyjściowej przestrzeni na obszary

Dla każdego sygnału wejściowego i wyjściowego wyznaczamy jego dziedzinę $[x_{i,min}, x_{i,max}]$, $[u_{min}, u_{max}]$, $i = 1, 2$.

Każdy z wyznaczonych w ten sposób przedziałów dzielony jest na $2N + 1$ obszarów, oznaczających nośniki N zbiorów rozmytych. Liczba N może być różna dla każdego z sygnałów, a nośniki mogą być różnej długości. Z każdym

zbiorem rozmytym kojarzymy etykietę oddającą intencję jego wprowadzenia. Przykładowe określenia to *NB* (*Negative Big*), *NM* (*Negative Medium*), *NS* (*Negative Small*), *Z*, *PS* (*Positive Small*), *PM* (*Positive Medium*), *PB* (*Positive Big*).

Kształt funkcji przynależności może być dowolny, jednak najchętniej stosowanymi są trójkątne funkcje przynależności opisane równaniem (3).

3.2 Tworzenie reguł rozmytych na podstawie przykładów

Przypuśćmy, że zakres zmienności sygnału x_1 pokryto zbiorami rozmytymi $A^j, j = 1, \dots, j_{max}$, sygnału x_2 – zbiorami $B^k, k = 1, \dots, k_{max}$, a dziedzinę sygnału wyjściowego pokryto zbiorami $C^l, l = 1, \dots, l_{max}$.

Dla każdego przykładu $(x_1^{(i)}, x_2^{(i)}, u^{(i)})$ wyznaczmy teraz stopnie przynależności do odpowiedniego obszaru: $\mu_{A^j}(x_1^{(i)}), \mu_{B^k}(x_2^{(i)}), \mu_{C^l}(u^{(i)})$. Ostatecznie konkretną wartość sygnału zaliczamy do obszaru o najwyższym stopniu przynależności. Jeżeli np. obserwacja $x_1^{(i)}$ należy do zbioru *PB* w stopniu 0.8, do zbioru *PS* w stopniu 0.3, a jej przynależność do pozostałych zbiorów rozmytych jest równa zeru, to zaobserwowaną wartość zaliczamy do zbioru *PB*.

Ogólnie, jeżeli obserwacja $x_1^{(i)}$ należy w maksymalnym stopniu do zbioru A^{j*} , obserwacja $x_2^{(i)}$ należy w maksymalnym stopniu do zbioru B^{k*} , a obserwacja $u^{(i)}$ należy w maksymalnym stopniu do zbioru C^{l*} , to przykładowi $(x_1^{(i)}, x_2^{(i)}, u^{(i)})$ odpowiada reguła

$$\text{if } (x_1 \text{ is } A^{j*}) \text{ and } (x_2 \text{ is } B^{k*}) \text{ then } (u \text{ is } C^{l*}) \quad (40)$$

3.3 Przyporządkowanie stopni prawdziwości regułom

W przypadku, gdy reguły o identycznych przesłankach mają różne konkluzje – mamy do czynienia z sytuacją konfliktową, a reguły o takiej własności tworzą zbiór sprzecznych reguł.

Najprostsza metoda rozwiązywania konfliktów polega na przypisaniu każdej regule (ze zbioru sprzecznych reguł) tzw. stopnia prawdziwości i wybrania tylko tej reguły, która ma najwyższy taki stopień. W ten sposób nie tylko rozwiązujemy konflikt, ale i redukujemy liczbę reguł. Dla reguły R postaci (40) stopień prawdziwości τ_R obliczamy następująco:

$$\tau_R = \mu_{A^{j*}}(x_1) \cdot \mu_{B^{k*}}(x_2) \cdot \mu_{C^{l*}}(u) \quad (41)$$

Problem: Zbadać, co będzie gdy występujący w powyższym równaniu iloczyn zastąpi się operatorem \wedge (min).

W praktyce posiadamy pewną *aprioryczną* wiedzę o rozwiązywanym problemie. Np. parze danych $(x_1^{(i)}, x_2^{(i)}, u^{(i)})$ można przypisać subiektywny stopień $\beta^{(i)}$ charakteryzujący jej wiarygodność, typowość czy też użyteczność w

<i>PB</i>					
<i>PM</i>					
<i>PS</i>		<i>PB</i>			
<i>Z</i>					
<i>NS</i>					
<i>NM</i>					
<i>NB</i>					
	<i>NM</i>	<i>NS</i>	<i>Z</i>	<i>PM</i>	<i>PB</i>

Tabela 6: Postać bazy rozmytych reguł

procesie formowania bazy wiedzy. W takim przypadku stopień prawdziwości przyznany tworzonej regule będzie miał postać

$$\tau_R = \mu_{A^{j*}}(x_1^{(i)}) \cdot \mu_{B^{k*}}(x_2^{(i)}) \cdot \mu_{C^{l*}}(u^{(i)})\beta^{(i)} \quad (42)$$

3.4 Utworzenie bazy rozmytych reguł

Bazę reguł (wiedzy) stanowi tablica tworzona w sposób następujący: tworzymy tablicę, której boki odpowiadają wartościom rozmytym sygnałów wejściowych. Dla reguły postaci (40) rozmytą wartość sygnału wyjściowego umieszczamy na przecięciu wiersza odpowiadającego rozmytej wartości A^{i*} sygnału x_1 i kolumny odpowiadającej rozmytej wartości B^{j*} sygnału x_2 . W ten sposób otrzymujemy tablicę 6. Reprezentuje ona regułę **if** (x_1 *is* NS) **and** (x_2 *is* PS) **then** (u *is* PB).

3.5 Wyostrzanie

W oparciu o uzyskaną tablicę reguł poszukujemy teraz odwzorowania $g: (x_1, x_2) \rightarrow \bar{u}$, gdzie \bar{u} jest wartością wyjściową zwracaną przez sterownik.

Niech

$$\alpha^{(k)} = \mu_{A^{(k)}}(x_1) \cdot \mu_{B^{(k)}}(x_2) \quad (43)$$

oznacza stopień aktywności k -tej reguły.

Poszukiwana wartość \bar{u} wyznaczamy korzystając z jednej z metod wyostrzania, np. stosując metodę *center average defuzzification* otrzymujemy

$$\bar{u} = \frac{\sum_{k=1}^N \alpha^{(k)} \bar{u}^{(k)}}{\sum_{k=1}^N \alpha^{(k)}} \quad (44)$$

gdzie $\bar{u}^{(k)}$ jest punktem, w którym funkcja przynależności występująca w konkluzji k -tej reguły osiąga swoje maksimum.

W przypadku, gdy przesłanki reguł zawierają n przesłanek elementarnych, w których występują zbiory $A_1^{(k)}, \dots, A_n^{(k)}$, stopień aktywności k -tej

reguły jest równy

$$\alpha^{(k)} = \prod_{i=1}^n \mu_{A_i^{(k)}}(\bar{x}_i) \quad (45)$$

Niech \mathfrak{F} będzie zbiorem funkcji $f: [a_1, b_1] \times \dots \times [a_n, b_n] \rightarrow \mathbb{R}$, gdzie $|a_i b_i| < \infty$, $i = 1, \dots, n$. Wang i Mendel pokazali w [16], że odwzorowanie (44) jest uniwersalnym aproksymatorem, tzn. może aproksymować z dowolną dokładnością każdą funkcję $f \in \mathfrak{F}$.

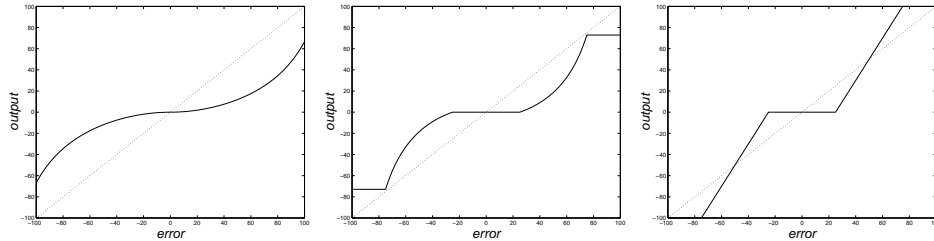
4 Metody alternatywne

4.1 Generowanie uproszczonych reguł

Pomysł ten pochodzi z pracy [9], w której autorzy zajmują się regułami postaci

$$R^{(k)}: \text{if } [(x_1 \text{ is } A_1^{(k)}) \text{ and } \dots \text{ and } (x_n \text{ is } A_n^{(k)})] \text{ then } (u \text{ is } b^{(k)}) \quad (46)$$

gdzie $b^{(k)} \in \mathbb{R}$ jest pewną liczbą. W dalszym ciągu reguły postaci (46) będziemy nazywać *prostymi* regułami rozmytymi w odróżnieniu od reguł postaci (5). Zaletą omawianych tu reguł jest to, że w procesie wnioskowania pomija się etap wyostrzania. W marę potrzeb można je także konwertować do typowego modelu lingwistycznego, co pokazujemy w sekcji 4.1.3. Ponadto – podobnie jak w przypadku ogólnym – reguły te posiadają dobre własności aproksymacyjne.



Rysunek 8: Zależność sygnału wyjściowego od wejściowego w proporcjonalnym sterowniku rozmytym. Na lewym i środkowym rysunku przesłanki i konkluzje są zbiorami rozmytymi o: (a) trójkątnych, (b) trapezoidalnych funkcjach przynależności. Linia przerywaną zaznaczono rezultat uzyskany za pomocą prostych reguł sterowania z przesłankami zdefiniowanymi przez trójkątne funkcje przynależności. Prawy rysunek przedstawia wynik generowany przez proste reguły z przesłankami zdefiniowanymi przez trapezoidalne funkcje przynależności.

Różnicę między prostymi a "normalnymi" regułami sterowania przedstawiono na rysunkach 8. Zakładamy, że zachowanie serownika proporcjonalnego opisują trzy reguły:

1. **if** error **is** *NEG* **then** output **is** *NEG*
2. **if** error **is** *Zero* **then** output **is** *Zero*
3. **if** error **is** *POS* **then** output **is** *POS*

Przyjęto też, że zbiory rozmyte *NEG*, *Zero* i *POS* opisane są za pomocą

- trójkątnych funkcji przynależności, odpowiednio: $(-\infty, -100, 0)$, $(-100, 0, 100)$, $(0, 100, \infty)$,
- trapezoidalnych funkcji przynależności, odpowiednio: $(-\infty, -100, -75, -25)$, $(-75, -25, 25, 75)$, $(25, 75, 100, \infty)$.

W przypadku gdy zarówno sygnał wejściowy jak i wyjściowy charakteryzowany jest przez powyższe funkcje przynależności, kształt wyostrzonego sygnału wyjściowego ma postać jak na rysunkach 8a oraz 8b. W obu przypadkach linią przerywaną zaznaczono sygnał wyjściowy uzyskany za pomocą prostych reguł sterowania. W tym ostatnim przypadku wejście etykietowane jest zbiorami z trójkątnymi funkcjami przynależności, a wyjście charakteryzują trzy liczby: -100, 0 oraz 100. Dla porównania, rysunek 8c odpowiada sytuacji, gdy przesłanki prostych reguł opisane są trapezoidalnymi funkcjami przynależności.

W przypadku sterownika Mamdaniego⁶ (pierwszy wiersz na rysunku 8) wynikowa relacja między sygnałem wejściowym a wyjściowym pokazana jest w ostatniej kolumnie pierwszego wiersza. Jak widać, nie jest ona liniowa, a co istotniejsze – nie cały zakres sterowania jest tu wykorzystywany.

Ten ostatni problem znika w przypadku zastosowania prostych reguł rozmytych. Tutaj etykietcie *NEG* odpowiada wartość -100, etykietcie *Zero* – wartość 0, a etykietcie *POS* – wartość 100. W takim przypadku otrzymujemy liniową relację wejście-wyjście.

Można sprawdzić, że np. w przypadku trapezoidalnych funkcji przynależności, relacja wejście-wyjście ma także charakter liniowy, ale z *plateau* w okolicy zera. Eksperymenty wskazują, że przy projektowaniu sterownika najlepiej jest stosować się do następujących reguł:

1. używać trójkątnych funkcji przynależności przecinających się w punkcie $\mu = 0.5$
2. modelować spójnik **and** operatorem mnożenia
3. konkluzje reguł powinny być numeryczne, tzn. należy stosować proste reguły sterowania wyznaczania sygnału wyjściowego stosować wariant wyostrzania metodą środków ciężkości opisany równaniem (47).

⁶W obliczeniach przyjęto, że T-normę operator mnożenia, że S-normę – operator max, a wyostrzanie przeprowadzono metodą środka ciężkości.

4.1.1 Wnioskowanie z użyciem prostych reguł rozmytych

Niech $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n)^T$ będzie wektorem obserwacji. Odpowiadającą mu wartość $u(\bar{\mathbf{x}})$ wyznaczamy z równania

$$u(\bar{\mathbf{x}}) = \frac{\sum_{k=1}^N \prod_{j=1}^n \mu_{A_j^{(k)}}(\bar{x}_j) b^{(k)}}{\sum_{k=1}^N \prod_{j=1}^n \mu_{A_j^{(k)}}(\bar{x}_j)} \quad (47)$$

Tutaj, jak poprzednio, iloczyn $\prod_{j=1}^n \mu_{A_j^{(k)}}(\bar{x}_j)$ określa stopień zgodności obserwacji $\bar{\mathbf{x}}$ z przesłankami k -tej reguły. Jest to wariant wyostrzania metodą środków ciężkości.

4.1.2 Identyfikacja prostych reguł rozmytych

Zakładamy, że dane mają postać identyczną z (39). Zakładamy też, że dokonano odpowiedniego pokrycia dziedzin zmiennych X_1, \dots, X_n zbiorami rozmytymi. Jeżeli A_i^k oznacza k_i -ty rozmyty podzbiór dziedziny zmiennej X_i to przez

$$w_{k_1 \dots k_n}(\bar{\mathbf{x}}) = \left(\prod_{j=1}^n \mu_{A_j^{k_j}}(\bar{x}_j) \right)^\alpha \quad (48)$$

oznacza stopień zgodności obserwacji $\bar{\mathbf{x}}$ z rozmytym obszarem wyznaczonym przez wybraną rodzinę podzbiorów. Zauważmy, że z równania tego wynika, że jeżeli dziedzinę zmiennej X_i pokryto K_i zbiorami rozmytymi, to dane uczące pozwalają wygenerować co najwyżej $K = K_1 \dots K_n$ reguł rozmytych. Jeżeli istotnie wygenerowano K reguł, mówimy wówczas o *kompletnej* bazie wiedzy.

Występujący w tym równaniu wykładnik $\alpha \in (0, 1)$ steruje stopniem dopasowania wynikowych reguł do danych trenujących.

Dysponując wagami (48), regule o indeksie $k = k_1 \dots k_n$ przypisuje się następującą numeryczną konkluzję:

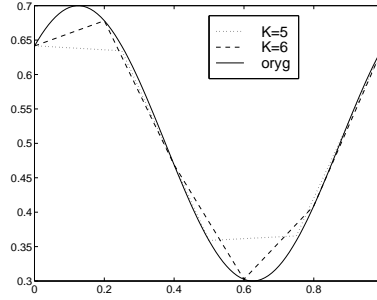
$$b_{k_1 \dots k_n} = \frac{\sum_{j=1}^m w_{k_1 \dots k_n}(\bar{\mathbf{x}}_j) \cdot u_j}{\sum_{j=1}^m w_{k_1 \dots k_n}(\bar{\mathbf{x}}_j)} \quad (49)$$

tzn. $b_{k_1 \dots k_n}$ jest ważoną sumą m obserwacji zmiennej zależnej.

Oczywiście, jeżeli $w_{k_1 \dots k_n} = 0$ oznacza to niemożność wygenerowania reguły dla wybranej kombinacji zbiorów rozmytych etykietujących elementarne przesłanki. Jest to sygnał wskazujący na konieczność redefinicji rozmytych pokryć.

Decydując się na wybór konkretnego pokrycia, jak również na wybór wartości α można kierować się wartościami wskaźnika dopasowania

$$e = \sqrt{\sum_{j=1}^m \left(u(\bar{\mathbf{x}}_j) - u_j \right)^2} \quad (50)$$



Rysunek 9: Porównanie jakości aproksymacji zadanej funkcji (przedstawionej linią ciągłą) dla dwóch różnych pokryć dziedziny $[0, 1]$.

Przykład 4.1. Celem ilustracji omówionej procedury rozważmy problem aproksymacji statycznego nieliniowego systemu opisanego równaniem

$$u = 0.2 \sin(2\pi x + \pi/4) + 0.5, 0 \leq x \leq 1$$

Zbiór trenujący utworzono wybierając 11 punktów o odciętych $x_{i+1} = x_i + 0.1$, $x_0 = 0.0$. Rozważono też dwa alternatywne pokrycia zbioru $[0, 1]$. Pierwsze to pięć trójkątnych funkcji przynależności postaci $\Delta_1(-\infty, 0.0, 0.25)$, $\Delta_2(0.0, 0.25, 0.50)$, $\Delta_3(0.25, 0.50, 0.75)$, $\Delta_4(0.50, 0.75, 1.0)$, $\Delta_5(0.75, 1.0, \infty)$. Drugie pokrycie obejmuje sześć trójkątnych funkcji przynależności postaci $\Delta_1(-\infty, 0.0, 0.2)$, $\Delta_2(0.0, 0.2, 0.4)$, $\Delta_3(0.2, 0.4, 0.6)$, $\Delta_4(0.4, 0.6, 0.8)$, $\Delta_5(0.6, 0.8, 1.0)$, $\Delta_6(0.8, 1.0, \infty)$.

W pierwszym przypadku jakość rekonstrukcji oryginalnej funkcji mierzona błędem średniokwadratowym jest równa $e = 0.003268$, a w drugim $e = 0.001972$. Zrekonstruowane krzywe przedstawiono na rysunku 4.1.2. Jeżeli w danych występują zakłócenia, np. *outliery*, to sensowny wybór wartości α pozwala zredukować ich wpływ na jakość rekonstrukcji. \square

4.1.3 Konwersja prostych reguł rozmytych

Przypuśćmy, że dany jest zestaw prostych reguł, a na dziedzinie zmiennej u określono rodzinę rozmytych podzbiorów $\mathbb{B} = \{B_1, \dots, B_N\}$. Niech $B_{j*}^{(k)} \in \mathbb{B}$ będzie zbiorem rozmytym o najwyższym stopniu zgodności z konkluzją $b^{(k)}$ tej reguły. Podobnie, niech $B_{j**}^{(k)} \in \mathbb{B} \setminus \{B_{j*}^{(k)}\}$ będzie zbiorem rozmytym o najwyższym (w zbiorze $\mathbb{B} \setminus \{B_{j*}^{(k)}\}$) stopniu zgodności z konkluzją $b^{(k)}$ tejże reguły. Wówczas regułę (46) można zastąpić dwiema regułami lingwistycznymi

$$\begin{aligned} R^{(k*)}: & \quad \text{if } [(x_1 \text{ is } A_1^{(k)}) \text{ and } \dots \text{ and } (x_n \text{ is } A_n^{(k)})] \text{ then } (u \text{ is } B_{j*}^{(k)}) \\ R^{(k**)}: & \quad \text{if } [(x_1 \text{ is } A_1^{(k)}) \text{ and } \dots \text{ and } (x_n \text{ is } A_n^{(k)})] \text{ then } (u \text{ is } B_{j**}^{(k)}) \end{aligned} \quad (51)$$

W szczególnym przypadku, gdy stosujemy zbiory rozmyte o trójkątnych funkcjach przynależności przecinających się parami na wysokości $\mu = 0.5$ zauważamy co następuje:

- dowolna konkluzja $b^{(k)}$ jest pokrywana przez dwa sąsiadujące ze sobą zbiory $B_{j^*}^{(k)}$ oraz $B_{j^{**}}^{(k)}$;
- $\mu_{B_{j^*}^{(k)}}(u) + \mu_{B_{j^{**}}^{(k)}}(u) = 1$, $u \in [\hat{B}_{j^*}^{(k)}, \hat{B}_{j^{**}}^{(k)}]$, przy czym $\hat{B}_{j^*}^{(k)}$ oznacza środek zbioru $B_{j^*}^{(k)}$, a ponadto $B_{j^*}^{(k)} < B_{j^{**}}^{(k)}$;
- dowolną wartość $u \in [\hat{B}_{j^*}^{(k)}, \hat{B}_{j^{**}}^{(k)}]$ można przedstawić w postaci sumy ważonej $u = \mu_{B_{j^*}^{(k)}}(u) \cdot \hat{B}_{j^*}^{(k)} + \mu_{B_{j^{**}}^{(k)}}(u) \cdot \hat{B}_{j^{**}}^{(k)}$, tzn. dowolną wartość $u \in [\hat{B}_{j^*}^{(k)}, \hat{B}_{j^{**}}^{(k)}]$ można przedstawić jako kombinację wypukłą zbiorów B_{j^*} oraz $B_{j^{**}}$.

Ta ostatnia własność oznacza, że istotnie liczbę $b^{(k)}$ można przedstawić lingwistycznie za pomocą dwóch zbiorów rozmytych.

4.2 Zastosowanie algorytmów genetycznych

4.2.1 Genetyczna regresja

Przypuśćmy, że zbiór danych postaci (39) chcemy opisać równaniem liniowym $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$. Klasyczną metodą jest w tym wypadku metoda najmniejszych kwadratów polegająca na wyznaczeniu wektora $\hat{\mathbf{a}}$ oraz wyrazu wolnego \hat{b} minimalizujących wskaźnik jakości

$$q(\mathbf{a}, b) = \frac{1}{m} \sum_{i=1}^m [\mathbf{a}^T \mathbf{x}_i + b - u^{(i)}]^2 \quad (52)$$

Jest to zadanie regresji liniowej. Do jego rozwiązania można też wyznaczyć algorytm genetyczny. W tym przypadku chromosomem jest wektor o składowych reprezentujących poszczególne elementy wektora \mathbf{a} , natomiast (52) jest naturalną funkcją jakości. Ideę tę wykorzystano m.in. do wyboru "optymalnych" parametrów funkcji przynależności.

Pomysł ten wykorzystali Karr i Gentry [8] zakładając, że dany jest zbiór obserwacji (39) oraz zbiór reguł sterowania postaci (5).

Przykład 4.2. [11] Załóżmy, że zachowanie układu opisuje FAM pokazana w lewej części tablicy 7, a dane empiryczne przedstawiono w jej prawej części.

x	S	L	1	2	3	4	5
y	S	VL	1	4	9	16	25

Tabela 7: Lingwistyczny opis systemu wraz z danymi empirycznymi

Przypuśćmy, że terminy lingwistyczne S modelowane są przez trójkątne funkcje przynależności postaci $(-\infty, 0, c_i^1)$ a terminy L oraz VL przez funkcje postaci (a_i^2, ξ_i^2, ∞) , $i = 1, 2$, $\xi_1^2 = 5$, $\xi_2^2 = 25$. Załóżmy, że dziedziną zmiennej x jest przedział $[0, 5]$, a $y \in [0, 25]$ i przyjmijmy, że pojedynczy chromosom $\mathbf{v} = (\delta_1^1, \delta_2^1, \delta_1^2, \delta_2^2)$ reprezentuje rozpiętości odpowiednich zbiorów rozmytych, tzn. $\delta_1^i = c_i^1$, $\delta_2^i = \xi_i^2 - a_i^2$. Dla każdej wartości x_j z lewej części tablicy 7 wyznaczamy odpowiadającą jej wartość $\hat{y}_j = \mu_{R_k}^{-1}(\mu_{L_k}(x_j))$, gdzie R_k (L_k) oznacza prawą (lewą) stronę reguły uaktywnionej przez obserwację x_j ; wyznaczamy też kwadrat błędu $e_j = (\hat{y}_j - y_j)^2$. Funkcja dostosowania ma więc postać (52), tzn. $f(\mathbf{v}) = \sum_j e_j$.

Przykładowo, jeżeli chromosom ma postać $(0.56, 1, 59, 8.73, 20.24)$ to p. wartość $x_4 = 4$ aktywuje regułę z drugiej kolumny lewej części tablicy 7 i $\mu_L(x) = 0.37$. Wartości tej odpowiada $\hat{y} = \mu_{VL}^{-1}(0.37) = 12.25$, a zatem błąd aproksymacji jest równy w tym wypadku $(16 - 12.25)^2 = 14.06$. \square

4.2.2 Uczenie reguł

Przedstawiony powyżej pomysł modyfikowano na wiele sposobów. Por. np. punkt 6.2.2 w [10], a także następujące prace:

- Ishibuchi, H. Nozaki, K., Yamamoto, N. Selecting fuzzy rules by genetic algorithm for classification problems. In: *Second IEEE International Conference on Fuzzy Systems*, pp. 1119–1124, San Francisco, CA., 1993
- Park, Y.J., Lee, S.Y., Cho, H.S. A genetic algorithm-base fuzzy control of an electro-hydraulic fin position servo system. In: *Proc. of 1999 IEEE Internat. Conference on Fuzzy Systems*, Seoul, Korea 1999, pp.III-1361-III-1366
- Porter L.L., Passino K.M., Genetic adaptive observers, *Engineering Applications of Artificial Intelligence*, **8**, 1995, pp. 261–269
- Kim, H.S., Roschkeb, P.N. Design of fuzzy logic controller for smart base isolation system using genetic algorithm. *Engineering Structures*, **28**, 2006, pp. 84–96

Poniżej przedstawiono algorytm zaproponowany w [6].

Założmy, że do modelowania zbiorów rozmytych wykorzystujemy trapezoidalne funkcje przynależności, tzn. i -ta przesłanka k -tej reguły wykorzystuje funkcję $(a_i^{(k)}, b_i^{(k)}, c_i^{(k)}, d_i^{(k)})$. Oczywiście, gdy $b_i^{(k)} = c_i^{(k)}$, otrzymujemy trójkątną funkcję przynależności.

Skoro reguła składa się z n przesłanek i pojedynczej konkluzji, to do jej reprezentacji potrzebujemy wektora \mathbf{v} o wymiarze $4(n+1)$ zbudowanego z $(n+1)$ czwórek postaci (a_i, b_i, c_i, d_i) , $i = 1, \dots, n+1$.

Do modyfikacji chromosomów można stosować np. mutację gaussowską, lub mutację niejednostajną postaci

$$v'_k = \begin{cases} v_k + \Delta(t, v_{k,max} - v_k) & \text{gdy } r < 0.5 \\ v_k - \Delta(t, v_k - v_{k,min}) & \text{wpp} \end{cases} \quad (53)$$

gdzie r to liczba losowa, $[v_{k,min}, v_{k,max}]$ to przedział zmienności k -tego parametru, t jest numerem generacji, natomiast

$$\Delta(t, \xi) = \xi(1 - r_1^{(1-t/T)^b})$$

przy czym r_1 jest (inną) liczbą losową, T to maksymalna liczba generacji, a b jest parametrem (np. $b = 5$).

Do wymiany informacji między osobnikami \mathbf{v} i \mathbf{w} zastosowano wariant krzyżowania arytmetycznego polegającego na wygenerowaniu czterech osobników

$$\begin{aligned} \mathbf{z}_1 &= \alpha \mathbf{v} + (1 - \alpha) \mathbf{w} \\ \mathbf{z}_2 &= \alpha \mathbf{w} + (1 - \alpha) \mathbf{v} \\ \mathbf{z}_3 &\text{ gdzie } z_i = \min\{v_i, w_i\} \\ \mathbf{z}_4 &\text{ gdzie } z_i = \max\{v_i, w_i\} \end{aligned}$$

i wybraniu dwóch najlepszych. Oczywiście $\alpha \in (0, 1)$.

Do oceny osobników wykorzystano funkcję uwzględniającą pięć aspektów

- (a) (*Intensywność*) Jeżeli przez $\tau_k(e_j)$ oznaczmy zgodność j -tego przykładu z k -tą regułą (por. równanie 41), to intensywność reguły opisuje funkcja

$$f_1(R^{(k)}) = \frac{1}{m} \sum_{j=1}^m \tau_k(e_j)$$

- (b) (*Dodatnie pokrycie*) Mówimy, że j -ty element zbioru uczącego jest przykładem pozytywnym, jeżeli $\tau_k(e_j) \geq p$, gdzie p jest zadaną i ustaloną wartością progową. Niech $E^+(R^{(k)}) = \{e_j \in E : \tau_k(e_j) \geq p\}$ będzie zbiorem przykładów pozytywnych, a $n_k^+ = |E^+(R^{(k)})|$ jego liczebnością. Wówczas uśrednione pokrycie zbioru dodatnich przykładów określa funkcja

$$f_2(R^{(k)}) = \frac{1}{n_k^+} \sum_{e_j \in E^+(R^{(k)})} \tau_k(e_j)$$

- (c) (*Negatywne pokrycie*) Przykład e_j nazywamy negatywnym, jeżeli jego zgodność z przesłanką jest dodatnia, ale zgodność z konkluzją jest zerowa. A zatem istnieje zapewne reguła z inną konkluzją pokrywająca ten przykład. Niech n_k^- będzie liczebnością zbioru przykładów negatywnych. Definiujemy kolejną funkcję

$$f_3(R^{(k)}) = \begin{cases} 1 & \text{gdy } n_k^- \leq 5 \\ (n_k^- - 5 + \exp(1))^{-1} & \text{wpp} \end{cases}$$

- (d) (*Rozpiętość reguły*) Niech $s_i = d_i - a_i$, $i = 1, \dots, n+1$, będzie rozpiętością zbioru opisującego i -tą elementarną przesłankę reguły, $m_i = c_i - b_i$ określa długość jej mediany (rozmiar jądra) tejże przesłanki, a dom_i będzie rozpiętością jej dziedziny. Określmy następujące wartości:

$$\rho_1 = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{s_i}{dom_i}, \quad \rho_2 = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{m_i}{dom_i}$$

Rozpiętość reguły wyznaczamy ze wzoru

$$f_4(R^{(k)}) = g_1(\rho_1)g_2(\rho_2)$$

gdzie przyjmujemy $g_i(\xi) = \exp(1 - a\xi)$ gdy preferowane są niewielkie wartości ilorazów oraz $g_i(\xi) = \exp(-|1 - a\xi|)$ gdy ważna jest ustalona relacja między odpowiednimi wielkościami.

- (d) (*Symetria*) Niech $\delta_{i,1} = b_i - a_i$, $\delta_{i,2} = d_i - c_i$, $d_i = \max(\delta_{i,1}/\delta_{i,2}, \delta_{i,2}/\delta_{i,1})$ dla $i = 1 \dots, n+1$. Wówczas

$$f_5(R^{(k)}) = \left(\max_{i=1, \dots, n+1} d_i \right)^{-1}$$

Ostatecznie jakość reguły $R^{(k)}$ wyznaczamy mnożąc wartości powyższych pięciu funkcji. Zadanie polega na maksymalizacji globalnej funkcji jakości $f(R^{(k)})$.

Algorytm generowania "optymalnych" reguł przyjmuje następującą postać:

0. (*Inicjalizacja*). Wczytać zbiór przykładów E o liczebności m . Ustalić wartości progowe p oraz ϵ . Utworzyć tablicę $CV[j] = 0$, $j = 1, \dots, m$. Podstawić $\mathfrak{R} = \mathbf{nil}$ (\mathfrak{R} jest listą reguł).
1. Zastosować algorytm genetyczny do zbioru przykładów.
2. Wybrać najlepszy chromosom, a reprezentowaną przez niego regułę R^* dodać do zbioru reguł \mathfrak{R} .
3. Wyznaczyć zbiór pozytywnych przykładów $E^+(R^*)$. Dla każdego $e_j \in E^+(R^*)$: $CV[j] = CV[j] + \tau_*(e_j)$. Jeżeli $CV[j] \geq \epsilon$ usunąć e_j ze zbioru przykładów.
4. Jeżeli $E = \mathbf{nil}$ to STOP, w przeciwnym razie powrót do kroku (1).

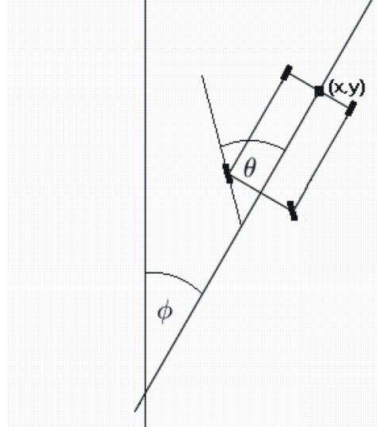
Parametry algorytmu genetycznego: populacja złożona z 60 osobników, prawdopodobieństwo krzyżowania $p_c = 0.6$, parametr $\alpha = 0.35$, prawdopodobieństwo modyfikacji chromosomu $p_m = 0.6$, a prawdopodobieństwo modyfikacji genu $p_g = p_m/[4 * (n+1)]$. Warunek zatrzymania to zadana z góry liczba iteracji. **uzupełnić** ...

5 Przykładowe problemy

Jednym z popularniejszych zadań testowych z zakresu sterowania jest zadanie parkowania ciężarówki; omówiono je w sekcji 5.1. Popularnym zadaniem testowym z zakresu prognozowania, tzn. przewidywania przyszłych wartości szeregu czasowego jest prognozowanie dla szeregu generowanego przez równanie Mackey-Glassa, które omówiono w sekcji 5.2.

5.1 Zadanie parkowania ciężarówki

Ciężarówka porusza się wewnątrz prostokąta $[-150, 150] \times [0, 300]$. Należy sterować kątem skręcenia przednich kół θ w taki sposób aby zaparkować ciężarówkę tyłem do rampy znajdującej się w punkcie $(0, 0)$. Położenie ciężarówki jest określone przez trzy współrzędne: x, u (określają one środek tylnej części ciężarówki) oraz kąt ϕ pod jakim ciężarówka (wektor równoległy do niej) znajduje się względem osi u – por. rysunek 10. Jeżeli $\phi = 0$ – ciężarówka ulokowana jest "tyłem" do rampy. Przyjmujemy, że $\theta \in [-45^\circ, 45^\circ]$, a $\phi \in [-180^\circ, 180^\circ]$.



Rysunek 10: Parametry wejściowe, (x, u, ϕ) oraz parametr sterujący θ w zadaniu parkowania ciężarówki.

Równania opisujące (przybliżoną) dynamikę ciężarówki mają postać

$$\begin{aligned} x(t+1) &= x(t) + \sin[\theta(t) + \phi(t)] - [\sin \theta(t)][\cos \phi(t)] \\ u(t+1) &= u(t) - \cos[\theta(t) + \phi(t)] - [\sin \theta(t)][\sin \phi(t)] \\ \phi(t+1) &= \phi(t) - \arcsin\left[\frac{2 \sin \theta(t)}{b}\right] \end{aligned} \quad (54)$$

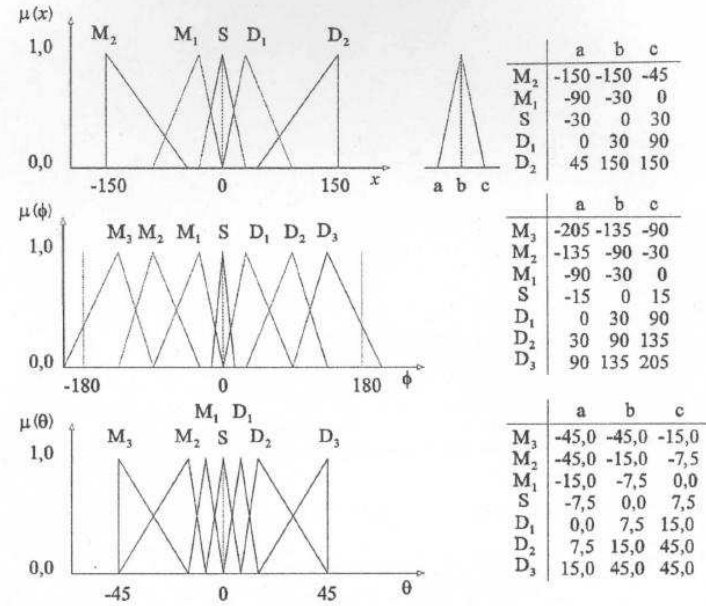
gdzie $b = 20$ jest długością ciężarówki.

Dla uproszczenia zakładamy, że współrzędna u nie będzie uwzględniana jako wejście sterownika, tzn. dane uczące mają postać $(x(t), \phi(t), \theta(t))$. Jeżeli ciężarówka znajdzie się na osi OY i będzie umiejscowiona w stosunku do

niej równolegle – zadanie można uznać za zakończone. Symulację należy też przerwać jeżeli ciężarówka znajdzie się poza "placem" (porażka).

Dane uczące przygotowywane są w następujący sposób: dla każdego położenia ciężarówki $(x(t), \phi(t))$ określa się sterowanie, tzn. kąt skręcenia kół ciężarówki, $\theta(t)$ opierając się na własnej intuicji. Po kilku próbach wybiera się dane odpowiadające najbardziej płynnej trajektorii jazdy (oczywiście zakończonej powodzeniem). Aby wygenerować dane uczące wybrano szesnaście położen początkowych ciężarówki: $(-100, -90^\circ)$, $(-100, 0^\circ)$, $(-100, 90^\circ)$, $(-100, 180^\circ)$, $(-50, -90^\circ)$, $(-50, 0^\circ)$, $(-50, 90^\circ)$, $(-50, 180^\circ)$, $(50, -90^\circ)$, $(50, 0^\circ)$, $(50, 90^\circ)$, $(50, 180^\circ)$, $(100, -90^\circ)$, $(100, 0^\circ)$, $(100, 90^\circ)$, $(100, 180^\circ)$.

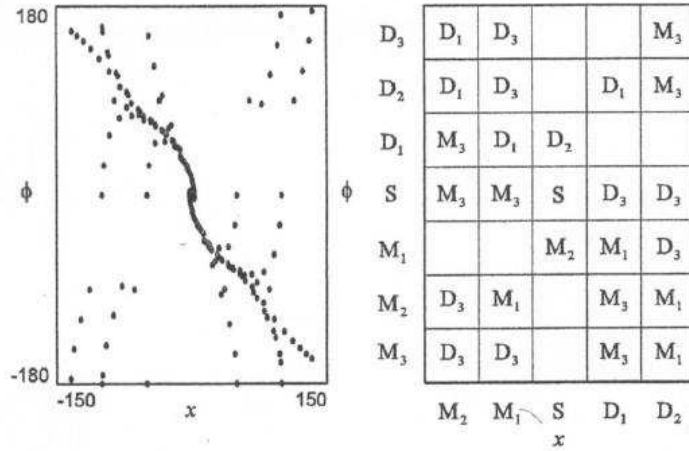
W książce [12] zastosowano do opisu badanego obiektu zbiory rozmyte, których definicje przedstawiono na rys. 11. W oparciu o nie oraz o zebrane dane empiryczne wygenerowano reguły przedstawione na rysunku 12



Rysunek 11: Przykładowe definicje zbiorów rozmytych wykorzystane do sterowania procesem parkowania ciężarówki (źródło: [12]).

5.2 Równanie Mackey-Glass'a

Jak podaje Wikipeda "Prognozowanie (predykcja) jest naukowym sposobem przewidywania, w jaki sposób będą kształtowały się w przyszłości procesy lub zdarzenia". Najczęściej przyjmuje się, że dysponujemy danymi uczącymi postaci $(x_i, x_{i+1}, \dots, x_{i+m-1}; x_{i+m})$, $i = 1, 2, \dots$, gdzie $m > 0$ jest parametrem określającym rozmiar obserwacji. W oparciu o m zaobserwowanych



Rysunek 12: Dane empiryczne wraz z wygenerowanym zestawem reguł (źródło: [12]).

wartości należy określić kolejną (przyszłą) wartość x_{i+m} . Jest to tzw. predykcja jedno-krokowa w odróżnieniu od predykcji k -krokowej polegającej na przewidzeniu k kolejnych wartości $x_{i+m}, x_{i+m+1}, \dots, x_{i+m+k-1}$. Podstawowe miary jakości predykcji jedno-krokowej to:

- błąd średniokwadratowy (czyli estymator wariancji prognozy)

$$MSE = \frac{1}{t} \sum_{i=1}^t (x_{i+m} - \hat{x}_{i+m})^2$$

gdzie t oznacza liczbę wykonanych prognoz, a \hat{x}_{i+m} jest wartością prognozowaną

- średni bezwzględny błąd prognozy

$$MAE = \frac{1}{t} \sum_{i=1}^t |x_{i+m} - \hat{x}_{i+m}|$$

Dla małych (mniejszych od 1) wartości zmiennej zależnej, błąd średniokwadratowy jest mniejszy od błędu absolutnego, natomiast dla dużych (tzn. większych od 1) liczb zależność jest odwrotna. Dlatego lepszą miarą jakości predykcji dla stosunkowo niewielkich wartości jest MAE, a dla większych od jeden MSE.

Aby móc wnioskować na podstawie próby losowej należy przyjąć dwa założenia:

1. prawidłowość z okresu $(1, \dots, m)$ również obowiązuje w okresie prognostycznym w przypadku danych czasowych, albo prawidłowość zachodząca dla próby również obowiązuje w całej populacji,

2. składnik losowy w przedziale prognozy zachowuje sie tak samo jak w okresie próby.

Wygodną formą zapewnienia spełnienia warunku (1) jest generowanie nietrywialnego, syntetycznego, szeregu czasowego. Do realizacji tego celu może służyć równanie zaproponowane przez Mackey'a i Glass'a⁷ postaci:

$$\frac{dx(t)}{dt} = \frac{b \cdot x(t - \tau)}{1 + x^c(t - \tau)} - ax(t), \tau \geq 17 \quad (55)$$

gdzie zazwyczaj $a = 0.1$, $b = 0.2$, $c = 10$.

Modeluje ono wiele procesów fizjologicznych, np. dynamikę produkcji białych ciałek krwi (limfocytów), czy bezdech centralny związany z oddychaniem typu Cheyne-Stokes'a.

Występujący w tym równaniu parametr τ jest opóźnieniem; gdy $\tau \geq 17$, szereg czasowy wykazuje własności chaotyczne⁸.

Poniższe równanie jest różnicowym odpowiednikiem równania (55)

$$x(t + 1) = x(t)(1 - a) + \frac{bx(t - \tau)}{1 + x^c(t - \tau)} \quad (56)$$

W przypadku, gdy $\tau = 17$ można przyjąć następujące wartości początkowe dla generowanego szeregu⁹: 1.097410, 1.128060, 1.145130, 1.147970, 1.139460, 1.123850, 1.105310, 1.087330, 1.072550, 1.062700, 1.058500, 1.059560, 1.064340, 1.070190, 1.073410, 1.069520, 1.054130.

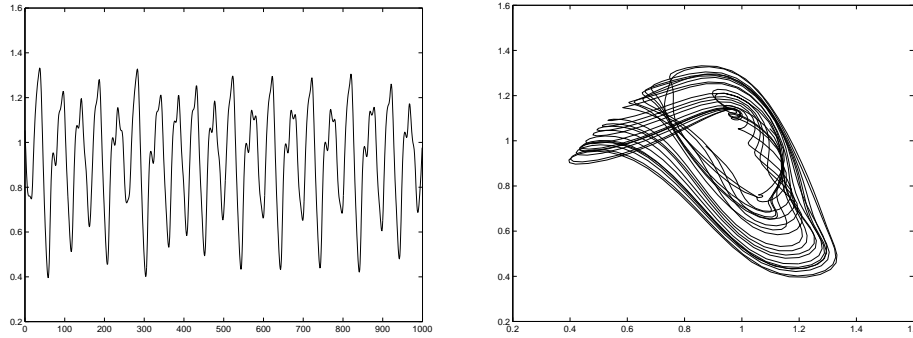
Generowany tą metodą szereg czasowy wykorzystywany był w wielu pracach poświęconych predykcji z zastosowaniem sieci neuronowych, np.

1. Crowder, R.S., Predicting the Mackey-Glass timeseries with cascade-correlation learning. In: D. Touretzky, G. Hinton, and T. Sejnowski (eds), *Proc. of the 1990 Connectionist Models summer School*. Carnegie Mellon University, 1990, pp. 117-123
2. 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
3. Lapedes, A.S., Farber, R. Nonlinear signal processing using neural networks. Tech. Rep. No. LA-UR-87-2662). Los Alamos, NM: Los Alamos National Laboratory, 1987

⁷Por. M.C. Mackey, L. Glass. Oscillation and chaos in physiological control systems. *Science*, **197**(4300), 1977, pp. 287-289, URL: http://www.cnd.mcgill.ca/bios/mackey/differential_equations.htm

⁸Przyjazne wprowadzenie do tych zagadnień można znaleźć w I. Steward, *Czy Bóg gra w kości*, PWN, Warszawa 1994.

⁹Por. B.S. Mulloy, R.L. Riolo, R.S. Savit. Dynamics of genetic programming and chaotic time series prediction. In: J.R. Koza, *et al.* (eds.) *Genetic Programming. Proc. 1-st Annual Conference*. The MIT Press, Cambridge, MA. 1996, pp. 166-174



Rysunek 13: Lewy rysunek: chaotyczny szereg uzyskany przez iterowanie równania różnicowego (56). Prawy rysunek: rekonstrukcja dziwnego atraktora metodą Packarda-Takensa.

4. Moody, J. Fast learning in multi-resolution hierarchies. *Advances in neural information processing systems*. Morgan Kaufmann Publishers Inc. 1989, pp. 29 - 39
5. Moody, J., Darken, C. Fast learning in networks of locally-tuned processing units. *Neural Computation*, **1**, 1989, pp. 289-303
6. Sanger, T.D. A tree-structured adaptive network for function approximation in high-dimensional spaces. *IEEE Transactions on Neural Networks*, **2**, 1991, pp. 285-293

Wyniki rozmytej predykcji przedstawiono na rys. 14

6 Związki między rozmytością a prawdopodobieństwem

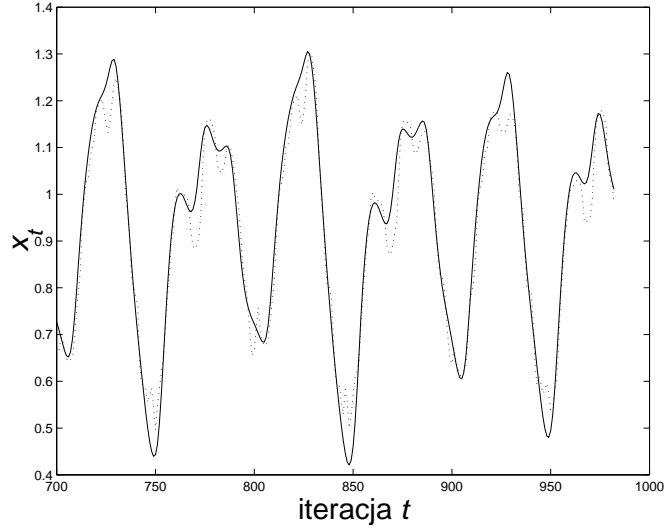
6.1 Zbiory losowe

Niech Ω, Θ będą dyskretnymi zbiorami, a $\Gamma: \Omega \rightarrow 2^\Theta$ będzie multifunkcją, tzn. odwzorowaniem przypisującym elementom $\omega \in \Omega$ podzbiory $\Gamma(\omega) \subseteq \Theta$. Przeciwobrazami zbioru $H \subseteq \Theta$ w odwzorowaniu Γ są

1. *górnny* przeciwobraz $H^* = \{\omega \in \Omega: \Gamma(\omega) \cap H \neq \emptyset\}$; wskazuje on np. na te obserwacje, które nie przeczą hipotezie H ,
2. *dolny* przeciwobraz $H_* = \{\omega \in \Omega: \Gamma(\omega) \subseteq H, \Gamma(\omega) \neq \emptyset\}$; wskazuje on np. na te obserwacje, które implikują (potwierdzają) hipotezę H ,
3. $H^{-1} = \{\omega \in \Omega: \Gamma(\omega) = H\}$

Zachodzi relacja

$$H^* = \Theta \setminus (\Theta \setminus H)_*$$



Rysunek 14: Prognozowanie wartości x_{701}, \dots, x_{1000} szeregu chaotycznego w oparciu o 155 reguł wygenerowanych z danych x_1, \dots, x_{700} . Dane oryginalne zaznaczono linią ciągłą, a prognozę – linią kropkowaną. Błąd prognozy: $MSE = 0.002476$, $MAE = 0.037110$

Jeżeli Γ jest odwzorowaniem przestrzeni probabilistycznej (Ω, \mathcal{B}, P) na przestrzeń mierzalną $\mathfrak{K} = (2^\Theta, 2^{2^\Theta})$, to indukuje ono w \mathfrak{K} pewną miarę $P_{\mathfrak{K}}$, którą charakteryzują następujące funkcje zbioru [15]:

- $m(H) = Pr(\mathfrak{K} = H) = \sum_{\omega \in \Omega: \Gamma(\omega) = H} P(\omega)$ – pełni rolę funkcji gęstości masy prawdopodobieństwa; oczywiście $(\forall H \subseteq \Theta): m(H) \in [0, 1]$ oraz $\sum_{H \subseteq \Theta} m(H) = 1$. Na ogół wymaga się, aby $m(\emptyset) = 0$. Jeżeli $m(H) > 0$ to H jest *elementem ogniskowym* (fokalnym).
- $Bel(H) = \sum_{A \subseteq H, A \neq \emptyset} m(A)$ – funkcja *przekonania*
- $Pl(H) = \sum_{A \cap H \neq \emptyset} m(A)$ – funkcja *domniemania*
- $Q(H) = \sum_{A \supseteq H} m(A)$ – funkcja *pokrycia*

Funkcje te charakteryzują się następującymi własnościami [15], [5]:

- $Bel(\emptyset) = Pl(\emptyset) = 0$,
- $Bel(\Theta) = Pl(\Theta) = 1$,
- monotoniczne funkcje zbioru, tzn. $(\forall H_1, H_2 \in \Theta)(H_1 \subset H_2) \Rightarrow (f(H_1) \leq f(H_2))$, gdzie $f \in \{Bel, Pl\}$,
- $(\forall H \in \Theta): Bel(H) \leq Pl(H)$,
- $(\forall H \in \Theta): Bel(H) = 1 - Pl(H^c)$,
- Przedział $[Bel(\cdot), Pl(\cdot)]$ można traktować jako oszacowanie (nieznane-

- go) prawdopodobieństwa na podzbiorach zbioru Θ ,
- $(\forall H_1, H_2 \in \Theta): Pl(H_1 \cup H_2) \geq \max[Pl(H_1), Pl(H_2)]$,
- $(\forall H_1, H_2 \in \Theta): Bel(H_1 \cap H_2) \leq \min[Bel(H_1), Bel(H_2)]$

Przykład 6.1. (Brak kompletny wiedzy.)

$$m(A) = \begin{cases} 0 & \text{gdy } A \neq \Theta \\ 1 & \text{gdy } A = \Theta \end{cases}$$

$$Bel(A) = \begin{cases} 0 & \text{gdy } A \neq \Theta \\ 1 & \text{gdy } A = \Theta \end{cases} \quad Pl(A) = \begin{cases} 0 & \text{gdy } A = \emptyset \\ 1 & \text{gdy } A \neq \emptyset \end{cases}$$

□

Przykład 6.2. (Deterministyczna funkcja przekonania.) W tym przykładzie zakładamy, że osoba podejmująca decyzję, ma pewność, że prawdziwa jest hipoteza $H \subset \Theta$).

$$m(A) = \begin{cases} 0 & \text{gdy } A \neq H \\ 1 & \text{gdy } A = H \end{cases}$$

$$Bel(A) = \begin{cases} 1 & \text{gdy } A \supseteq H \\ 0 & \text{w p.p.} \end{cases} \quad Pl(A) = \begin{cases} 1 & \text{gdy } A \cap H \neq \emptyset \\ 0 & \text{w p.p.} \end{cases}$$

□

6.2 Konsonantne zbiory losowe

Założmy teraz, że elementy ogniskowe tworzą monotoniczną rodzinę: $F_1 \subset F_2 \subset \dots \subset F_n$, $n \leq |\Theta|$. Wówczas:

- $Pl(\{\theta\}) = Q(\{\theta\}) = \mu(\theta)$,
- $Pl(H) = \max_{\theta \in H} \mu(\theta)$, tzn. funkcja domniemania jest tzw. *miarą możliwości* oznaczaną symbolem $\Pi(H)$, [18], [5]
- $\mathfrak{N}(H) = 1 - \Pi(H^c)$ jest *miarą konieczności* spełniającą warunek $\mathfrak{N}(H_1 \cap H_2) = \min[\mathfrak{N}(H_1), \mathfrak{N}(H_2)]$

Niech teraz dana będzie funkcja przynależności $\mu: \Theta \rightarrow [0, 1]$ i niech $U = \{u_1, \dots, u_n\}$ będzie zbiorem różnych wartości funkcji przynależności uporządkowanym rosnąco, tzn. $u_j \leq u_{j+1}$. Niech dalej $F_j = \{\theta \in \Theta: \mu(\theta) \geq u_j\}$. Wówczas (transformacja *możliwość* \rightarrow *prawdopodobieństwo*)

$$m(A) = \begin{cases} 1 - u_n & \text{gdy } A = \emptyset \\ u_j - u_{j-1} & \text{gdy } A = F_j \\ 0 & \text{w p.p.} \end{cases}$$

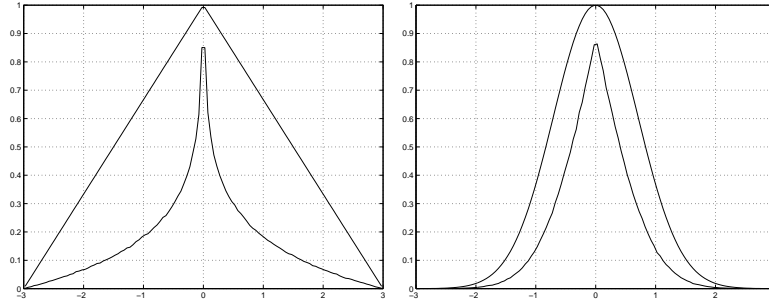
przy czym $u_0 = 0$.

6.3 Próbkowanie ze zbioru rozmytego (losowego)

Ostatnia obserwacja dała asumpt do probabilistycznego traktowania funkcji przynależności, w szczególności do generowania obserwacji w oparciu o zadaną funkcję przynależności. Problem ten był badali m.in. Yager [17] a także Dubois i Prade [3, 4]. W szczególności, Yager [17] zaproponował następującą procedurę generowania obserwacji, jeżeli dana jest funkcja przynależności:

- (a) Wybierz losowo α -przekrój F_α
- (b) Wybierz losowo element $\theta \in F_\alpha$

Poniższe rysunki przedstawiają wyniki symulacji dla różnych kształtów funkcji przynależności



Rysunek 15: 1,000,000 punktów losowanych dla zbioru rozmytego o: (a) trójkątnej, (b) wykładniczej funkcji przynależności

W przypadku wielowymiarowym szczególnie proste jest generowanie obserwacji z gaussowskiej funkcji przynależności postaci

$$\mu(x_1, \dots, x_n) = \exp\left(-\frac{(x_i - m_i)^2}{2\sigma_i^2}\right) \quad (57)$$

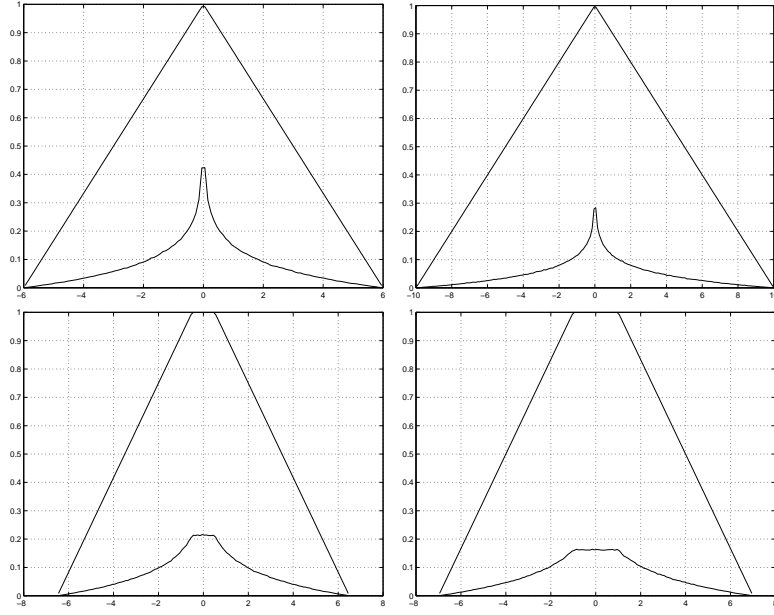
W tym przypadku, α -przekrojami są hiper-elipsoidy o równaniu

$$\left(\frac{x_1 - m_1}{\sqrt{-2\ln(\alpha)}\sigma_1}\right)^2 + \dots + \left(\frac{x_n - m_n}{\sqrt{-2\ln(\alpha)}\sigma_n}\right)^2 = 1 \quad (58)$$

Aby wybrać losowo punkt ξ z takiego obszaru, generuje się n liczb losowych $u_i \sim U(-1, 1)$ takich, że suma ich kwadratów nie przekracza wartości 1, a następnie oblicza się jego współrzędne:

$$\xi_i = m_i + \sqrt{-2\ln(\alpha)}\sigma_i \cdot u_i, i = 1, \dots, n \quad (59)$$

Wyniki symulacji przedstawiono na rysunku 18.



Rysunek 16: 1,000,000 punktów losowanych dla zbioru rozmytego o: (a) trójkątnych, (b) trapezowych f-kcjach przynależności o różnych nośnikach i rdzeniach.

6.4 Losowe eksperymenty z nieprecyzyjnymi obserwacjami

Przykład 6.3. Chcemy oszacować interpretację kategorii *tall* rozumiejąc, że zaliczamy do niej osoby, których wzrost przekracza wartość θ . Przyjmijmy, że dziedziną zmiennej lingwistycznej *tall* jest $\Theta = [0.5, 2.5]$ metra. Rozbijmy Θ na n rozłącznych przedziałów $S_i, i = 1, \dots, n$.

Wykonujemy pomiary dużej grupy osób traktowanych jako *wysokie*. Każdy pomiar $\hat{\theta}$ to oszacowanie wartości θ sugerujące formułę $tall = [\hat{\theta}, 2.5]$. Prawdopodobieństwa $p(S_i)$ są więc przypisywane do przedziałów $[s_i, 2.5]$ gdzie $s_i = \inf S_i$, a multifunkcja Γ ma postać $\Gamma(s) = [s_i, 2.5]$ gdy $s \in S_i$, czyli $m([s_i, 2.5]) = p(S_i)$. Indukuje to zbiór rozmyty o funkcji przynależności

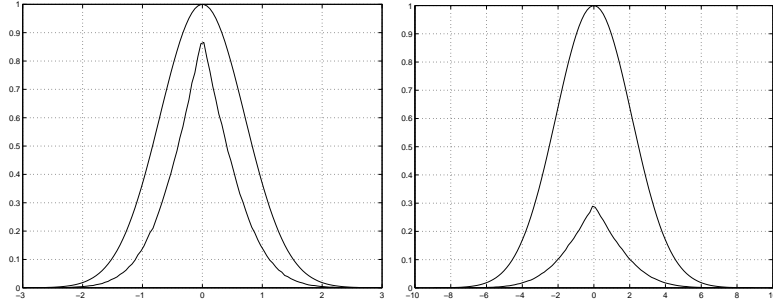
$$\mu_{tall}(s) = \Pi(\{s\}) = \sum_{s_i \leq s} p(S_i)$$

□

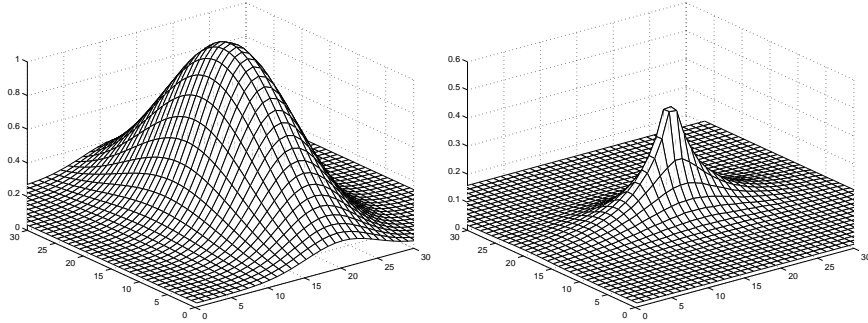
W ogólnym przypadku wymagamy, aby nieznaną wartość szacować za pomocą (ostrego) przedziału $[a_i, b_i], i = 1, \dots, q$, przy czym stawiamy wymaganie aby $\bigcap_{i=1}^q [a_i, b_i] = [a, b] \neq \emptyset$.

Niech $[A, B] = \bigcup_{i=1}^q [a_i, b_i]$ i zdefiniujmy monotoniczną rodzinę

$$[a, b] \subset I_1 \cdots \subset I_n = [A, B]$$



Rysunek 17: 1,000,000 punktów losowanych dla zbioru rozmytego o wykładniczej funkcji przynależności $\mu_A(u) = \exp(-u^2/a)$ dla $a = 1, 3$.



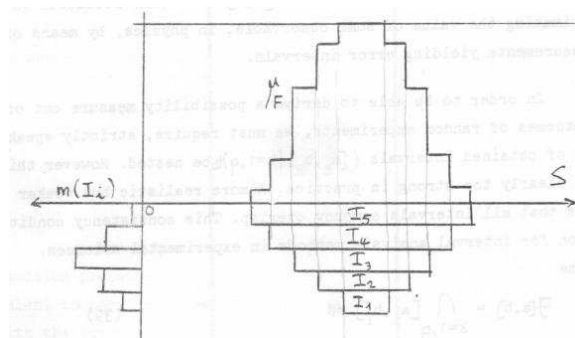
Rysunek 18: Rozkład gęstości 3,000,000 obserwacji wygenerowanych z odwołaniem się do gaussowskiej funkcji przynależności o parametrach $m_1 = m_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 2$. W prawej części rysunku przedstawiono wykres funkcji gęstości, a w lewej – funkcji przynależności.

Przedziały I_i odgrywają podobną rolę do przedziałów S_i . Każde oszacowanie $[a_j, b_j]$ zaliczamy do najmniejszego zawierającego go przedziału I_i . Następnie zliczamy jak wiele oszacowań przypisano każdemu I_i , tzn. wartości $m(I_i)$ i wyznaczamy funkcję przynależności

$$\mu(x) = \Pi(\{x\}) = \sum_{x \in I_i} m(I_i) \begin{cases} 0 & \text{gdy } x \notin I_n \\ \sum_{j=i}^n m(I_j) & \text{gdy } x \in [\inf I_i, \inf I_{i-1}) \cup (\sup I_{i-1}, \sup I_i] \\ 1 & \text{gdy } x \in I_1 \end{cases}$$

7 Systemy neuro-rozmyte

Wiedza zawarta w bazie wiedzy sterownika rozmytego ma postać zdań warunkowych zapisanych w języku bliskim naturalnemu. Baza wiedzy nie stanowi jednak narzędzia umożliwiającego pozyskiwanie takiej wiedzy. Z drugiej strony, sieci neuronowe stanowią narzędzie pozwalające na uczenie się z przykładów. Jednakże wiedza reprezentowana przez zbiór wag przypisanych



Rysunek 19: Ilustracja sposobu wyznaczania funkcji przynależności w oparciu o losowe eksperymenty (źródło [3]).

połączeniom jednostek tworzących daną sieć ma zdecydowanie nieintuicyjny charakter. Połączeniem obu formalizmów są systemy neuro-rozmyte posiadające zarówno zdolności uczenia się jak też reprezentujące wiedzę w sposób przyjazny dla użytkownika.

Literatura

- [1] Castro, J. L., Delgado, M., Herrera, F. A learning method of fuzzy reasoning by genetic algorithms, pp. 804-809, URL: citeseer.ist.psu.edu/article/castro93learning.html
- [2] Czogała, E., Łęski, J., *Fuzzy and Neuro-fuzzy Intelligent Systems*. Physica-Verlag, A Springer-Verlag Company, Heidelberg, New York, 1999
- [3] Dubois, D., Prade, H. Fuzzy sets and statistical data. *European J. of Op. Res.*, **25**, 1986, 345-356
- [4] Dubois, D., Prade, H. The statistical approach to membership functions: Random experiments with imprecise outcomes. *Journ. of Organizat. Behaviour & Stat.*
- [5] Dubois, D., Prade, H. *Possibility Theory. An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988
- [6] Herrera, F., Lozano, M., Verdegay, J.L. Generating fuzzy rules from examples using genetic algorithms. In: *Proc. of 5th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'94*, pp. 675-680, URL: <http://citeseer.ist.psu.edu/347754.html>

- [7] Hill, G., Horstkotte, E., Teichrow, J. *Fuzzy-C development system – user’s manual*. Togai Infralogic, 30 Corporate Park, Irvine, CA 92714, USA
- [8] Karr, C., Gentry, E. Fuzzy control of ph using genetic algorithms, *IEEE Transactions on Fuzzy Systems*, **1**, 1993, pp. 46–53
- [9] Nozaki, K., Ishibuchi, H., Tanaka, H. A simple but powerful heuristic method for generating fuzzy rules from numerical data. *Fuzzy Sets and Systems*, textbf86, 1997, 251-270
- [10] Piegat, A. *Modelowanie i sterowanie rozmyte*. Akademicka Oficyna Wydawnicza EXIT, warszawa, 1999
- [11] Ross, T.J. *Fuzzy Logic with Engineering Applications* J. Wiley, 2004
- [12] Rutkowska, D., Piliński, M. Rutkowski, L. *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*, PWN, Warszawa 1997
- [13] Rutkowski, L. *Metody i techniki sztucznej inteligencji. Inteligencja obliczeniowa*. PWN, Warszawa, 2005
- [14] Schweizer, B., Sklar, A. *Probabilistic Metric Spaces*. Elsevier & North-Holland, 1983
- [15] Shafer, G. *A Mathematical Theory of Evidence*, Princeton Univ. Press, N.J., 1976
- [16] Wang, L.X., Mendel, J.M. Generating fuzzy rules by learning from examples. *IEEE Trans. on Systems, Man and Cybernetics*, **22**, 1992, 1414-1427, URL: <http://sipi.usc.edu/~mendel/publications/Wang&Mendel\%20SMC\%2011-12-92.pdf>
- [17] Yager, R.R. level sets for membership evaluation of fuzzy subsets. In: R.R. Yaer, ed. *Fuzzy Sets and Possibility Theory: Recent Developments*, Pergamon Press, Oxford, 1982, pp. 90-97
- [18] Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* **1**: 3–28