

# Sieci Neuronowe

## Wykład 11 Algorytmy genetyczne; Systemy rozmyte

wykład przygotowany na podstawie.

S. Osowski, "Sieci Neuronowe w ujęciu algorytmicznym", Rozdz. 4, PWNT, Warszawa 1996.  
W. Duch, J. Korbicz, L. Rutkowski i R. Tadeusiewicz, "Sieci Neuronowe", Biocybernetyka i inżynieria biomedyczna 2000, t.6, ACW Exit, Warszawa 2000.

# Algorytmy genetyczne

---

Algorytmy genetyczne oraz ich modyfikacje nazywane algorytmami ewolucyjnymi to metody rozwiązywania różnych problemów w naukach technicznych, głównie zagadnień optymalizacyjnych, wzorowane na naturalnej ewolucji.

- ✓ Nie przetwarzają bezpośrednio parametrów zadania, lecz ich zakodowaną (w formie “chromosomów”) postać.
- ✓ Dokonują przeszukania przestrzeni rozwiązań startując nie z pojedynczego punktu ale z pewnej populacji takich punktów.
- ✓ Wykorzystują jedynie funkcje celu, nie wymagając znajomości jej pochodnych (tak jak metody gradientowe) ani też innych informacji pomocniczych.
- ✓ Stosują probabilistyczne a nie deterministyczne reguły wyboru.

# Algorytmy genetyczne

---

Podstawowy algorytm genetyczny, określany też jako klasyczny, powstał na przełomie lat 60-tych/70-tych, dzięki pracom *Hollanda*.

Holland zainteresowany cechami naturalnej ewolucji, m.in. faktem że odbywa się ona na chromosomach a nie na żywych istotach, wierzył że odpowiedni program komputerowy może zrealizować algorytm rozwiązywania wielu problemów w sposób naśladujący ewolucję. Rozpoczął więc pracę nad algorytmami symulowanej ewolucji przebiegającej na chromosomach, stanowiących ciąg cyfr binarnych.

Algorytmy te, wykorzystując analogiczne jak w naturalnej ewolucji mechanizmy selekcji oraz reprodukcji, zmierzały do rozwiązania danego problemu poprzez poszukiwanie najlepszego chromosomu na podstawie oceny jego przystosowania.

# Algorytmy genetyczne

---

Algorytmy nie wymagały znajomości rozwiązywanego problemu (nie ma konieczności modelowania), a ich działanie ograniczało się do znalezienia chromosomu najlepiej przystosowanego. Rozwiązanie problemu za pomocą takiego algorytmu rozpoczyna się od zakodowania jego potencjalnych rozwiązań w postaci ciągów zer-jedynek, nazywanych chromosomami oraz określenia odpowiedniej funkcji ich przystosowania, oceniającej te chromosomy. W trakcie działania algorytmu wykorzystywane są wspomniane wcześniej mechanizmy selekcji i reprodukcji.

Selekcja *stymulowanej ewolucji* naśladuje zasadę doboru naturalnego i przetrwania osobników najlepiej przystosowanych w procesie ewolucji organizmów żywych.

# Algorytmy genetyczne a sieci neuronowe

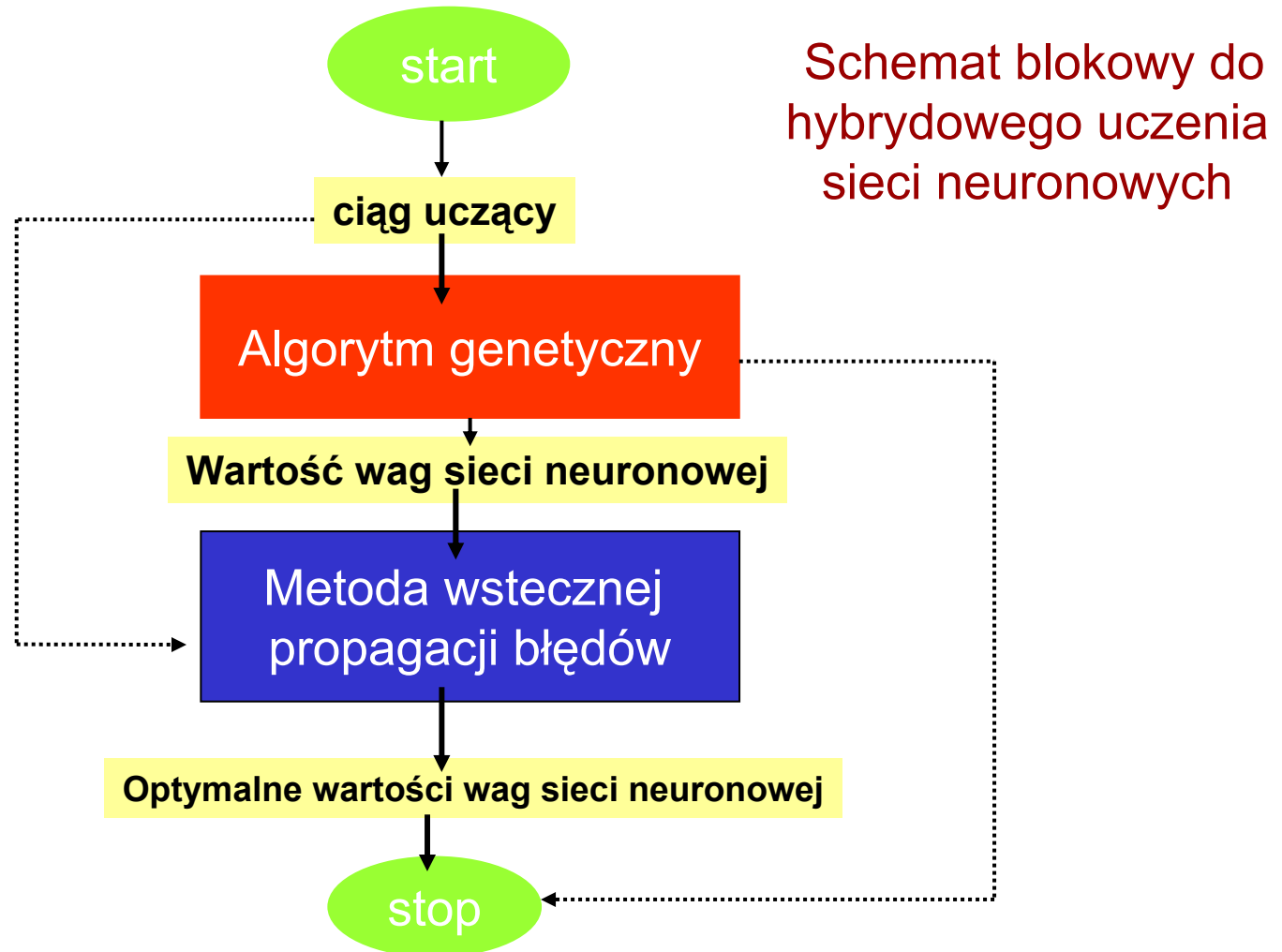
---

Możliwe jest połączenie wspomagające i współdziałające algorytmów genetycznych i sieci neuronowych:

→ *wspomagające* : oznacza, że jedna z metod wspomaga drugą, np. służy do przygotowywania danych przygotowywanych przez drugą metodę. Np.: algorytmu genetycznego używano do wyboru cech lub transformacji przestrzeni cech wykorzystywanych przez sieć neuronową; algorytm genetyczny służył do wyboru reguły uczenia lub parametrów sterujących uczeniem w sieci neuronowej.

→ *współdziałające* : oznacza że obie metody stosuje się jednocześnie. Do tej grupy kombinacji zalicza się ewolucyjne uczenie sieci neuronowych, to znaczy optymalny dobór wag sieci za pomocą algorytmu genetycznego. Ponadto w tej grupie mieści się również wykorzystanie algorytmów genetycznych do określenia topologii (architektury) sieci neuronowych, a także systemy łączące adaptacyjne strategie obu tych metod.

# Achemat blokowy hybrydowego uczenia sieci neuronowych



# Systemy rozmyte

---

Systemy rozmyte charakteryzują się przetwarzaniem informacji danej w postaci symbolicznej i wiedzy zapisanej w formie rozmytych reguł. Mogą być stosowane jako komplementarne metody stosowane do rozwiązywania tego samego problemu.

Działanie wynika z zastosowania *teorii zbiorów rozmytych* i *logiki rozmytej*.

Pozwalają opisywać zjawiska o charakterze nieprecyzyjnym i wieloznacznym, których nie potrafiła ująć teoria klasyczna i logika dwuwartościowa.

*Zbiór rozmyty*: to zbiór elementów dla których zdefiniowana jest funkcja przynależności elementu do zbioru

$$A = \{ (x, \mu_A(x)): x \in X \}, \mu_A(x); X \rightarrow [0, 1]$$

*Logika wielowartościowa* (przybliżone wnioskowanie) to aparat przy pomocy którego opisujemy operacje dla zbiorów rozmytych.

# Informacja numeryczna oraz lingwistyczna

---

Omawiane dotychczas sieci neuronowe w procesie zarówno uczenia jak i testowania wykorzystywały jedynie informacje numeryczne, wiążące w sposób ścisły i jednoznaczny dane wejściowe systemu z jego odpowiedzią.

Drugim typem informacji który jest niezrozumiały dla klasycznych sieci neuronowych, to informacja lingwistyczna, operująca pojęciami opisowymi, nie związanymi ściśle z wartościami numerycznymi. Przykładem takim są deskryptory „mały”, „duży”, „bardzo mały”, „bardzo duży”, które są zrozumiałe dla człowieka ale trudne do zaadoptowania w postaci numerycznej. Co więcej, każde z tych określeń lingwistycznych, wobec niezbyt precyzyjnie zdefiniowanych różnic między nimi, stanowi pojęcie nieostre, rozmyte (fuzzy). Co więcej, co dla jednego człowieka jest duże w ocenie drugiego może być bardzo duże, natomiast dla trzeciego przeciwnie - małe.



# Informacja numeryczna oraz lingwistyczna

---

*Wielkościom lingwistycznym* można przypisać określone zmienne numeryczne, dla których można wyznaczyć odpowiednie funkcje opisujące i zakres zmienności parametrów z nimi związanych.

W ten sposób można stworzyć reguły określające w sposób ścisły systemy tego typu, wytworzyć metody ich adaptacji (uczenia) oraz testowania.

Struktura układowa odpowiadająca tym regułom jest podobna do sieci neuronowych (struktura warstwowa, wiele neuronów w warstwie, równoległe przetwarzanie informacji), a sposób uczenia i testowania również oparty na metodach stosowanych w sieciach neuronowych.

# Informacja numeryczna oraz lingwistyczna

---

Sieci podlegające uczeniu i wykorzystujące przy tym pojęcia i zależności logiki rozmytej są nazywane *sieciami neuronowymi rozmytymi* (neurofuzzy).

W procesie uczenia korzystają one zarówno z *informacji lingwistycznych* jak i *numerycznych*, przy czym wykorzystuje się różne sposoby połączeń obu typów informacji. Najczęściej stosowanym podejściem jest wykorzystanie informacji lingwistycznej na etapie wstępnego doboru parametrów, opierając właściwe uczenie na informacji numerycznej.

# Podstawowe pojęcia systemów rozmytych

---

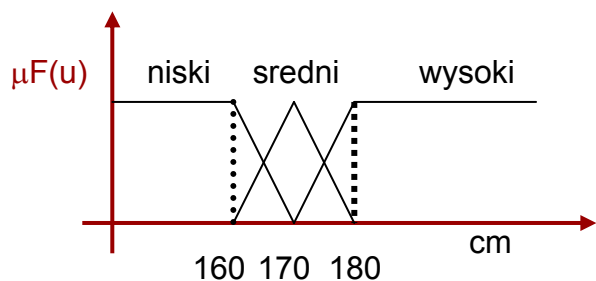
Podstawowym pojęciem teorii zbiorów, obok samego zbioru  $U$ , jest relacja przynależności elementu do tego zbioru. W odróżnieniu od zbiorów o nierozmytej przynależności, w zbiorach rozmytych nie ma ostrej granicy między elementami, które do danego zbioru należą, a tymi które należą do innych zbiorów.

Określa się stopień przynależności elementu do zbioru, czyli liczbę z przedziału  $[0,1]$ . Każdej zmiennej  $u$  można przyporządkować inną funkcję przynależności, określającą stopień przynależności zmiennej do zbioru rozmytego  $F$ .

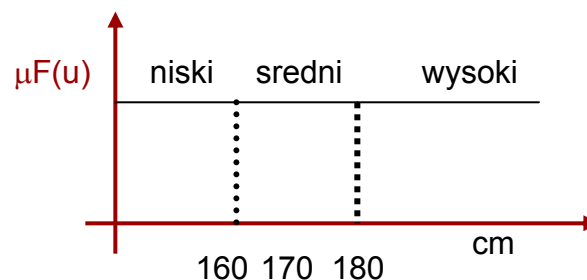
# Podstawowe pojęcia systemów rozmytych

Warunek  $\mu_F(u) = 1$  oznacza pełną przynależność  $u$  do zbioru  $F$ , wartość  $\mu_F(u)=0$  oznacza brak tej przynależności. Wartości pośrednie  $\mu_F(u)$  wyrażają przynależność częściową  $u$  do zbioru  $F$ .

W odróżnieniu od *algebry Boole'a*, która jednoznacznie klasyfikuje dane wejściowe przyporządkowując im w sposób nierozmyty odpowiedni zbiór, w przypadku *algebry rozmytej* tej samej wartości zmiennej wejściowej można przyporządkować różne zbiory z określoną wartością funkcji przynależności do zbioru.



system rozmyty



system przynależności ostrej

# Podstawowe pojęcia systemów rozmytych

---

Funkcje przynależności można przedstawić również w postaci zapisu macierzowego.

W przypadku systemu rozmytego w postaci o dwu wejściach  $u_1$ ,  $u_2$  oraz jednym wyjściu  $v$ , opisanego trzema regułami rozmytymi, wszystkie informacje dotyczące systemu mogą być opisane za pomocą wektorów  $u_1=[\dots]$ ,  $u_2=[\dots]$ ,  $v=[\dots]$  oraz trzech macierzy przynależności:  $A_{u_1}$ ,  $A_{u_2}$ ,  $A_v$  gdzie każdy wiersz odpowiada realizacji jednej z reguł rozmytych. Liczba kolumn macierzy odpowiada rozmiarowi wektorów wejściowych i wyjściowych.

Opis macierzowy funkcji rozmytych umożliwia uproszczenie organizacji obliczeń numerycznych i sprowadzenie ich do operacji algebraicznych na macierzach.

# Podstawowe pojęcia systemów rozmytych

---

Podstawową cechą zbiorów rozmytych, decydującą o ich znaczeniu praktycznym, jest wnioskowanie. Ogólny zapis wnioskowania ma postać:

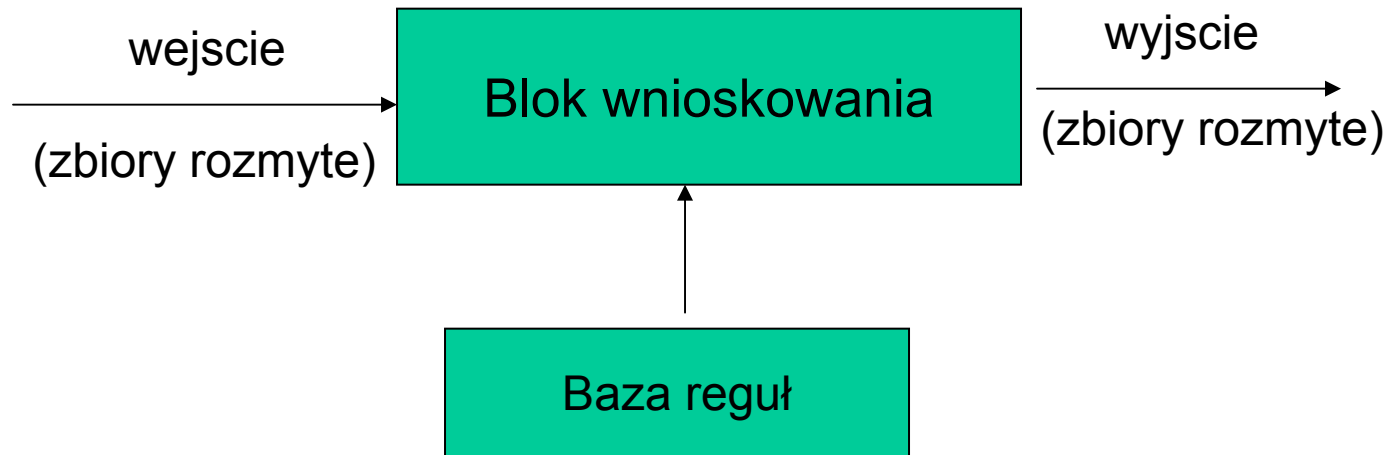
*jeśli przesłanka logiczna to konkluzja*

Istotnym problemem jest sposób tworzenia reguł wnioskowania.

Jednym z rozwiązań jest zastosowanie *eksperta* który na podstawie zdobytego wcześniej doświadczenia określi sposób postępowania dla poszczególnych przypadków, które mogą się zdarzyć w trakcie procesu. Zadaniem eksperta będzie konstrukcja samej reguły wnioskowania jak i doboru funkcji przynależności dla każdego przypadku.

Innym rozwiązaniem jest konstrukcja reguł wnioskowania na podstawie szeregu wykonanych badań eksperymentalnych. W takim przypadku podstawą modelu podejmowania decyzji są wyniki numeryczne eksperymentów, określające zarówno reguły wnioskowania, jak i funkcje przynależności.

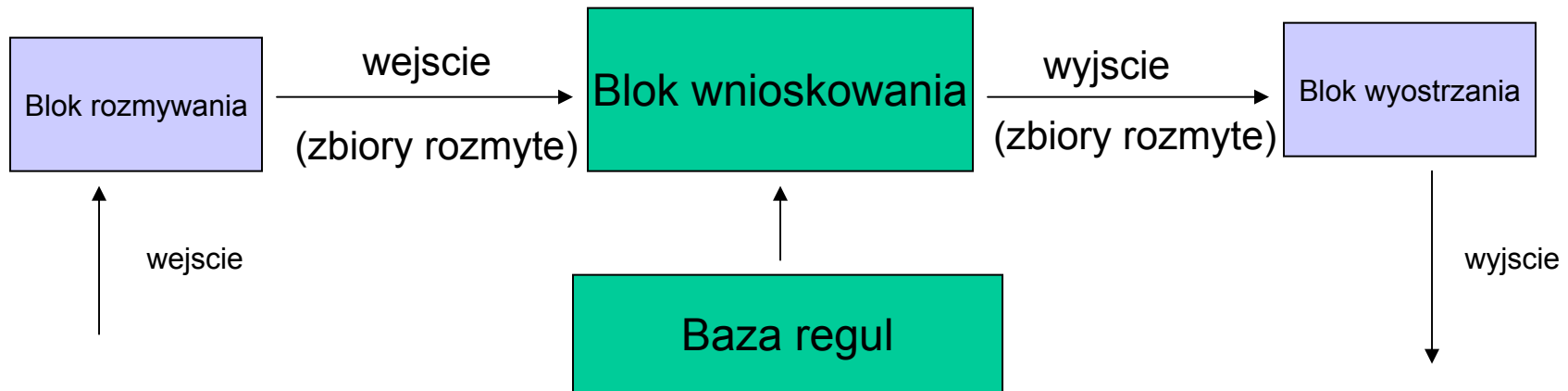
# Systemy rozmyte



W bloku wnioskowania dokonuje się wyliczenia rozmytych wartości wyjściowych na podstawie reguł zawartych w bazie reguł i rozmytych wartości na wejściu, korzystając z logiki rozmytej

W praktycznych zastosowaniach systemów rozmytych, głównie do sterowania, mamy do czynienia z wartościami numerycznymi 'ostrymi', podawanymi na wejście sterownika, którego zadaniem jest wskazanie wartości wyjściowych, również 'ostrych'.

# Systemy rozmyte



Algorytm sterowania jest oparty o zbiory rozmyte i logikę rozmytą. Wprowadzane i wyprowadzane są dane ‘ostre’, dodatkowe bloki realizują funkcje “rozmywania” i “wyostrzania”.

Blok wnioskowania może być realizowany przy pomocy sieci neuronowych pracujących z wartościami rozmytymi. Takie systemy nazywamy też *systemami rozmyto-neuronowymi*.



# Blok rozmywania i wyostrzania

---

Blok rozmywania ( *fuzyfikikator* ) przekształca n-wymiarowy wektor wejściowy  $x = [x_1, x_2, \dots, x_N]^T$  w zbiór rozmyty  $F$ , scharakteryzowany funkcją przynależności. Z punktu widzenia praktycznego największą popularność uzyskały funkcje gaussowskie oraz funkcje trójkątne i trapezoidalne.

Blok wyostrzania ( *defuzyfikikator* ) dokonuje przekształcenia zbioru rozmytego na ściśle zdeterminowany punkt rozwiązania  $y$ . Najczęściej stosuje się defuzyfikikator wg. Średnich wartości centrów (środek ciężkości) albo wg. ważonych średnich wartości centrów co uwzględnia również informacji na temat kształtu funkcji.

Zdolność funkcji rozmytych do reprezentacji dowolnej ciągłej funkcji nieliniowej  $f(x)$  umożliwia ich zastosowanie w aproksymacji funkcji wielu zmiennych i zbudowanie w oparciu o nie odpowiedniej sieci rozmytej realizowanej w postaci układu. To zdecydowało o możliwościach zastosowań praktycznych.

# Sieci neuronowe o logice rozmytej

---

Reprezentacja dowolnej funkcji nieliniowej wielu zmiennych za pomocą funkcji rozmytych jednej zmiennej, scharakteryzowanych przez funkcje przynależności  $\mu(x_i)$ , uzasadnia idee zastosowania funkcji rozmytych do odwzorowania dowolnych procesów nieliniowych ciągłych i stanowi alternatywne podejście do klasycznych sieci neuronowych jednokierunkowych.

Wyrażenie funkcji aproksymującej  $f(x)$  za pomocą relacji przynależności umożliwia ich implementację numeryczną w strukturze równoległej wielowarstwowej, identycznie jak w przypadku sieci neuronowych sigmoidalnych lub radialnych.

Struktury takie nazywane są sieciami o logice rozmytej lub sieciami neuronowymi rozmytymi. Stwarza to możliwość adaptacji metod uczenia sieci neuronowych klasycznych do sieci rozmytych.

# Sieci neuronowe o logice rozmytej

---

Podstawowe strategie uczenia:

- metoda gradientowa oparta na propagacji wstecznej;
- metoda samoorganizacji
- metoda tablicowa ( nie ma odpowiednika w metodach dla sieci klasycznej)

## *Metod tablicowa:*

w przypadku sieci neuronowych rozmytych jest możliwe bezpośrednio wykorzystanie reguł wnioskowania do sterowania procesem. Zamiast dopasowania wartości numerycznych wag, umożliwiającą określenie sygnału wyjściowego sieci , jest możliwe wyznaczenie tych wartości bezpośrednio na podstawie tzw. *tabeli przejść*.

Na podstawie żadanego ciągu danych uczących  $(x^{(1)}, d^{(1)})$ ,  $(x^{(2)}, d^{(2)})$ , .... Jest tworzony zbiór reguł logicznych typu *przesłanka* → *konkluzja* określających w sposób jednoznaczny pełne zachowanie układu przy żądanych warunkach początkowych.

# Sieci neuronowe o logice rozmytej

---

Kolejne zastosowanie tych reguł umożliwia przeprowadzenie zmiennej wyjściowej od położenia początkowego do pożądanego położenia końcowego.

Najistotniejszym etapem w tej metodzie jest generacja odpowiednich reguł logicznych na podstawie danych numerycznych i reguł lingwistycznych, określających kolejne etapy wnioskowania, umożliwiające przeprowadzenie procesu od warunków początkowych do warunków końcowych zadania.

Etapy tej metody to:

- podział przestrzeni danych wejściowych i wyjściowych na podzbiory rozmyte
- generacja reguł rozmytych na podstawie danych uczących
- hierarchizacja reguł
- tabela reguł wynikowych podejmowania decyzji
- strategia defuzyfikacji

# Sieci neuronowe o logice rozmytej

---

Przy zastosowaniu przedstawionych zasad postępowania kolejne etapy sterowania polegają na wielokrotnym wykorzystaniu tabeli podejmowania decyzji, za każdym razem przyporządkowujące wektorowi wejściowemu  $x$  odpowiedni zakres i przypisaną mu regułę logiczną.

Tabela dostarcza informacji o zakresie i odpowiadającej mu regule logicznej dotyczącej zmiennej wyjściowej.

Taki algorytm postępowania, dość złożony na etapie przygotowywania reguł wynikowych podejmowania decyzji, umożliwia bezpośrednie sterowanie systemami w fazie odtworzeniowej bez potrzeby uczenia.

Cała informacja ucząca jest zawarta w tabeli reguł podejmowania decyzji, stanowiącej *bazę wiedzy* o systemie.

# Inteligentny system obliczeniowy

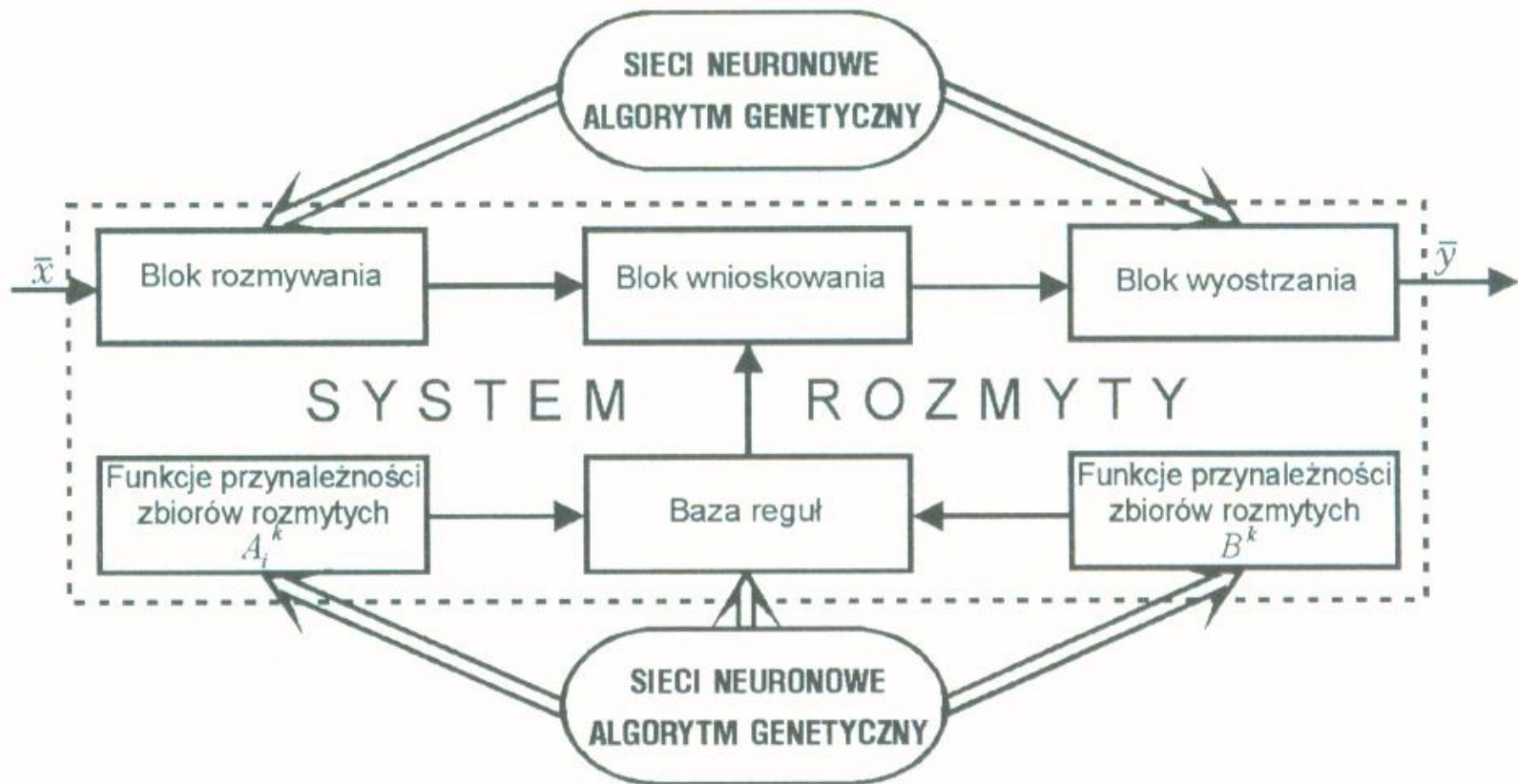
---

Inteligentny system obliczeniowy, łącząc w sobie cechy systemów rozmytych, sieci neuronowych i algorytmów genetycznych, charakteryzuje się zdolnością uczenia sieci neuronowych, regułowa reprezentacja wiedzy przedstawionej w sposób symboliczny (lingwistyczny) – typowa dla systemów rozmytych oraz możliwością globalnej optymalizacji parametrów systemu i reguł za pomocą algorytmu genetycznego.

Najważniejszą funkcję spełnia tu blok wnioskowania, który dokonuje wnioskowania na podstawie wielkości wejściowych i bazy reguł, korzystając z logiki rozmytej. Ten blok nazywamy *rozmytym systemem wnioskującym*.

Rozmyte systemy wnioskujące dokonują poprawnego wnioskowania, jeżeli reguły w bazie reguł, stanowiące wiedzę systemu o rozwiązywanym problemie są dobrane w sposób właściwy. Reguły te mogą być wyznaczane automatycznie, na podstawie ciągu uczącego, zawierającego wzorcowe pary wartości wejściowych i wyjściowych (sieci neuronowe, algorytm genetyczny).

# Inteligentny system obliczeniowy



Rys. 24.1. Ogólna koncepcja inteligentnego systemu obliczeniowego

# Uwagi końcowe

---

- Sztuczne sieci neuronowe mają następujące, podobne do mózgu cechy:
- zdolności pamięciowe, zwłaszcza pamięci adresowanej kontekstowo, skojarzeniowej
  - umiejętność uczenia się na przykładach
  - umiejętność generalizacji
  - odporność na uszkodzenia sieci
  - umiejętność równoległego przetwarzania informacji
  - umiejętność pracy niealgorytmicznej
  - mogą heurystycznie rozwiązywać problemy
  - mogą poprawnie pracować przy pewnym poziomie uszkodzeń sieci

Przyjmuje się powszechnie, że dzisiejsze sztuczne sieci neuronowe są pozbawione wielu cech odpowiadających wyższym czynnościom mózgowym, jak np. umiejętność abstrakcyjnego myślenia i świadomości.