

Technologia informacyjna

Sztuczna Inteligencja

Janusz Uriasz

4. Sztuczna inteligencja

Sztuczna inteligencja (**SI**) - dziedzina informatyki związana z koncepcjami i metodami wnioskowania symbolicznego, wykonywanego przez komputer, oraz symboliczną reprezentacją wiedzy, używaną przy wnioskowaniu; badania w dziedzinie sztucznej inteligencji mają za zadanie dostarczyć komputerom zdolności, które ludzie uznają za inteligencję.

Metody SI:

przetwarzanie danych \Rightarrow przetwarzanie wiedzy

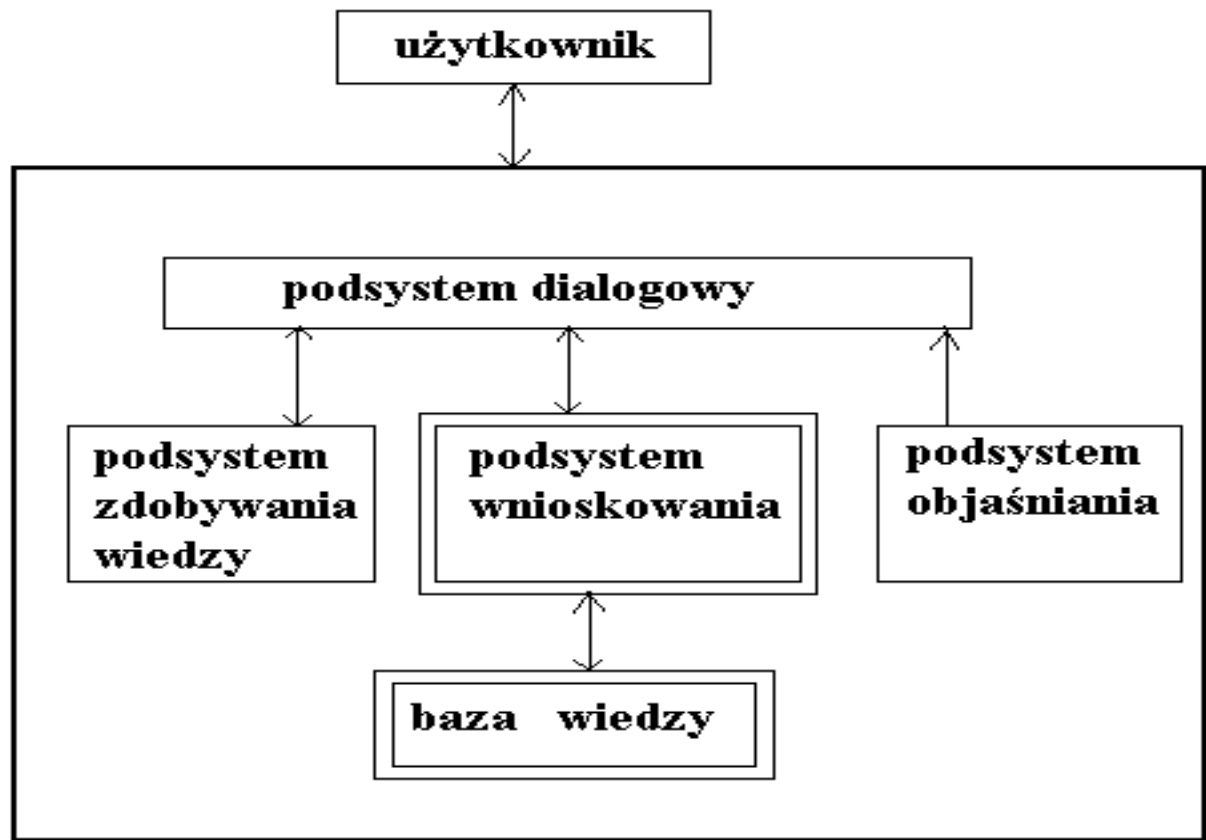
metody algorytmiczne
(przetw. proceduralne) \Rightarrow przeszukiwanie
inteligentne

SI obejmuje:

- rozwiązywanie problemów i strategię przeszukiwań
- teorię gier
- automatyczne dowodzenie twierdzeń
- przetwarzanie języka naturalnego (w tym przetwarzanie mowy)
- SYSTEMY EKSPERTOWE
- robotykę
- procesy percepcji
- uczenie się maszyn
- wyszukiwanie informacji (inteligentne bazy danych)
- programowanie automatyczne

4.1. Systemy ekspertowe

system ekspertowy - system informatyczny posiadający wiedzę z pewnej dziedziny, potrafiący wnioskować w ramach tej dziedziny na poziomie eksperta (wykorzystujący wiedzę eksperta)



podsystem dialogowy - zapewnia dialogowe współdziałanie z użytkownikiem (pytania, pytania uszczegółowiające)

podsystem bazy wiedzy - fakty (dane) wraz z regułami wnioskowania

podsystem wnioskowania -steruje procesami wnioskowania: dedukcyjnym (do przodu) i redukcyjnym (do tyłu)

podsystem objaśniania - umożliwia kontrolowanie procesu wnioskowania (jak rozwiązany został problem i/lub dlaczego została podjęta dana akcja systemu)

Rodzaje systemów ekspertowych

- Interpretacyjne dedukcja opisu sytuacji z obserwacji lub stanu czujników (rozpoznawanie mowy, obrazów, struktur danych)
- Predykcyjne wnioskowanie o przyszłości na podstawie danej sytuacji (prognoza pogody, rozwój choroby)
- Diagnostyczne określanie wad systemu na podstawie obserwacji (medycyna, elektronika, mechanika)
- Kompletowania kompletowanie obiektów w warunkach ograniczeń (konfiguracja systemu komputerowego)
- Planowania podejmowanie działań aby osiągnąć cel (ruchy robota)
- Monitorowania porównywanie obserwacji z ograniczeniami (elektrownie atomowe, medycyna, ruch uliczny)
- Sterowania kierowanie zachowaniem systemu; interpretowanie, produkcja, naprawa i monitorowanie zachowania systemu
- Poprawiania określanie sposobu postępowania w przypadku złego funkcjonowania obiektu, którego dotyczą
- Naprawy harmonogramowanie czynności przy dokonywaniu napraw uszkodzonego obiektu
- Instruowania systemy doskonalenia zawodowego

Sektor Zastosowanie	Bankowość i ubezpieczenia	Przemysł	Handel i usługi	Sektor publiczny i inne
Monitorowanie Sterowanie	obserwowanie trendów	nadzorowanie procesów, sterowanie procesami, raportowanie specjalnych sytuacji	obserwowanie trendów	monitorowanie reaktorów jądrowych oraz dużych sieci (gazowe, wodne)
Projektowanie		projektowanie zakładów i produktów, komputerów	wybór asortymentów, doradztwo dla rolnictwa	sieci (pocztowe, energetyczne)
Diagnostyka	kredyty, pożyczki na nieruchomości, analiza ryzyka, przetwarzanie skarg	wykrywanie uszkodzeń, utrzymanie zdolności produkcyjnej	kredyty, analiza ryzyka	diagnoza medyczna, diagnoza techniczna
Planowanie	analiza ryzyka, planowanie inwestycji	projektowanie funkcji logicznych, planowanie projektu	analiza ryzyka, analiza rynku	planowanie inwestycji, planowanie na wypadek klęski, planowanie dystrybucji

Sposoby realizacji:

- systemy dedykowane
- systemy szkieletowe (*shells*)

Podział SE ze względu na metodę prowadzenia procesu wnioskowania:

- SE z logiką dwuwartościową (Boole'a)
- SE z logiką wielowartościową
- SE z logiką rozmytą

Podział SE ze względu na rodzaj przetwarzanej informacji:

- SE z wiedzą pewną (zdeterminowaną)
- SE z wiedzą niepewną (aparatury probabilistycznej)

Właściwości SE:

- poprawność
- uniwersalność
- złożoność
- autoanaliza
- zdolność udoskonalania bazy wiedzy

Przykład konsultacji z użytkownikiem (I)

materiały niebezpieczne

proszę opisać sytuację

- * plama substancji płynnej w ładowni
- * koniec

czy można ocenić kolor cieczy?

- * tak

podaj kolor cieczy

- * brązowy

czy wyczuwalna woń?

- * tak

czy woń dusząca?

- *tak

załóż środki ochrony dróg oddechowych

....

....

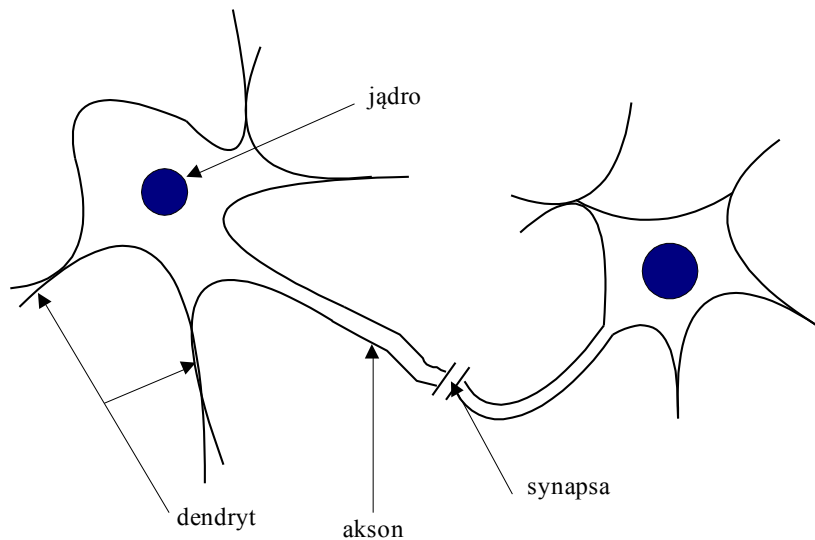
4.2. Sieci neuronowe

Sztuczna sieć neuronowa – abstrakcyjny model rzeczywistego układu nerwowego (uproszczony model mózgu):

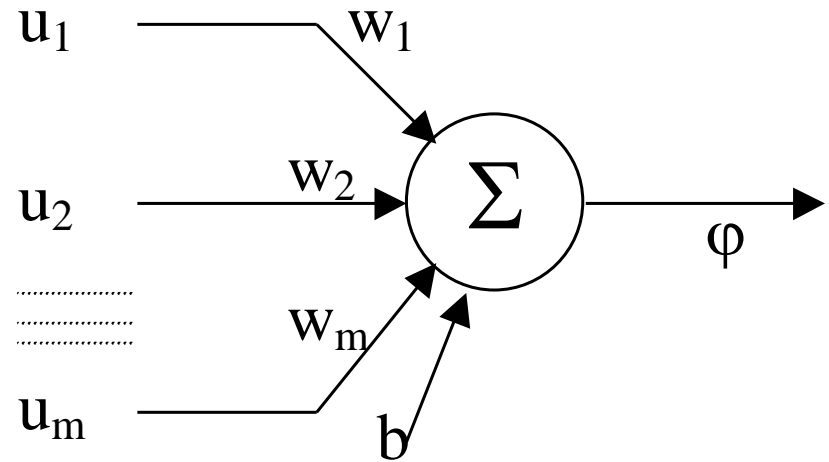
- uniwersalny układ aproksymacyjny odwzorowujący wielowymiarowe zbiory danych;**
- ma zdolność uczenia się i adaptacji do zmieniających się warunków otoczenia;**
- posiada zdolność uogólniania zdobytej wiedzy.**

Właściwości mózgu człowieka:

- Odporny na uszkodzenia.
- Elastyczny. Może z łatwością przystosować się do zmiennego otoczenia przez “uczenie się”
- Potrafi radzić sobie z informacją rozmytą, losową, zaszumioną lub niespójną.
- Równolegle przetwarza sygnały.
- Mały, zwarty, zużywa bardzo mało energii.
- Składa się z około 10^{11} neuronów (komórek nerwowych) różnych typów.



Budowa komórki nerwowej



Budowa neuronu

(ADALINE -ADaptive LINEar-)

Potencjał φ neuronu:

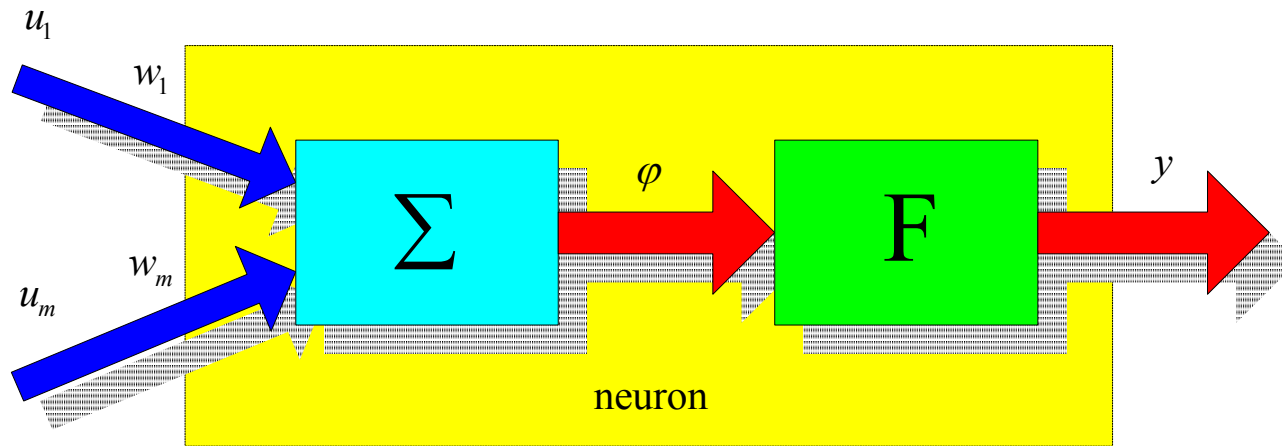
$$\varphi = \sum_{i=1}^m w_i \cdot u_i + b = \mathbf{w}^T \cdot \mathbf{u} + b$$

\mathbf{w} - wektor współczynników wag,

\mathbf{u} - wektor sygnałów wejściowych,

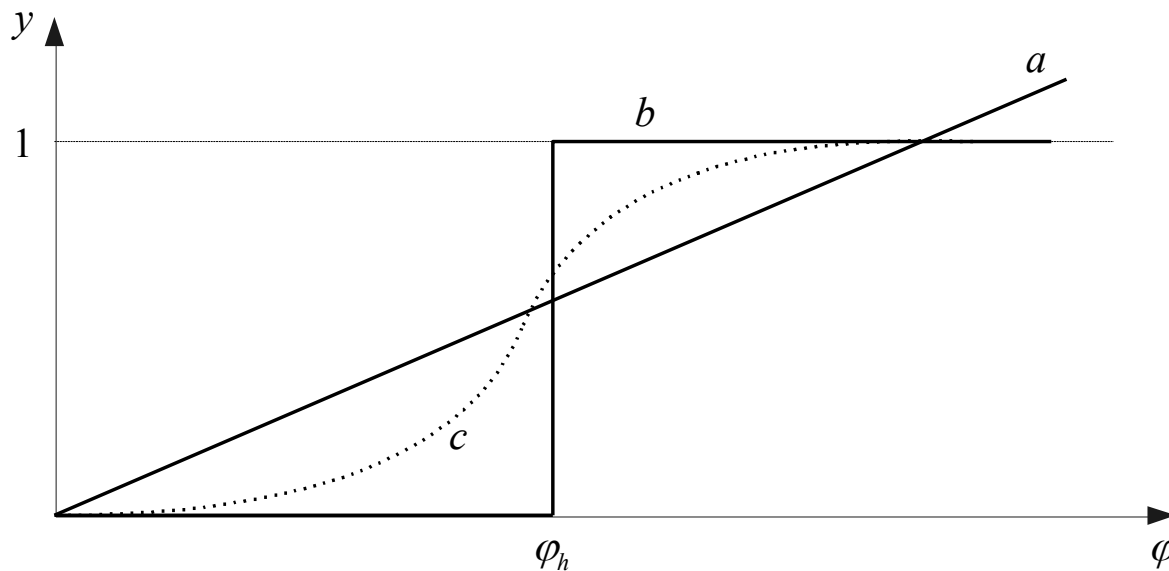
m - liczba wejść neuronu

b – składnik dodatkowy niezależny od sygnałów wejściowych
(bias)



Blok aktywacji F opisany funkcją przejścia (charakterystyka neuronu)

Przetwarzanie sygnałów w neuronie MLP (MultiLinear Perceptron)



Funkcje aktywacji w modelach neuronów: a) funkcja liniowa; b) funkcja skoku jednostkowego (funkcja progowa), c) funkcja sigmoidalna

Elementy sieci neuronowej - neurony (elementy neuropodobne) są określone:

- położeniem w sieci
- stanem wewnętrznym
- układem połączeń z innymi elementami
- stanem pobudzenia
- charakterystyką neuronu (funkcja aktywacji)

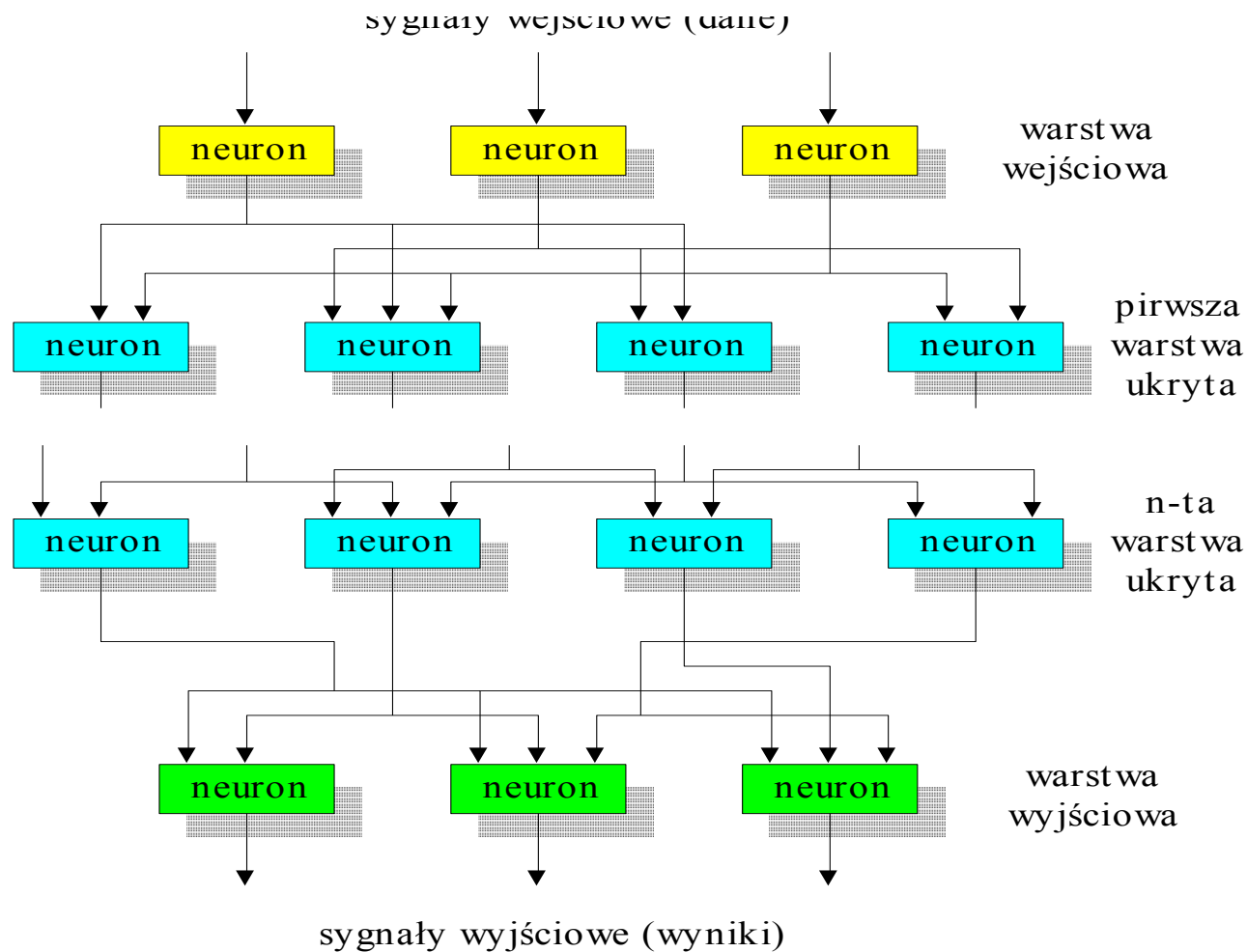
Moc obliczeń neuronowych wynika z połączenia wielu neuronów w sieci tworzące różne struktury. Tak zbudowane sieci neuronowe mają wiele możliwości obliczeniowych i wykazują własności

Cechy sieci:

- zdolność do adaptacji i samoorganizacji
- zmniejszona wrażliwość na uszkodzenia elementów
- zdolność do równoległej pracy
- wygoda w ich programowaniu poprzez uczenie.

Budowa warstwowa sieci neuronowych:

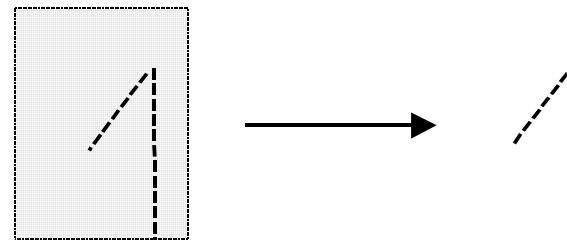
- warstwa wejściowa,
- warstwy ukryte,
- warstwa wyjściowa



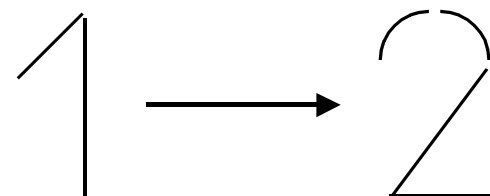
Cztery typy zadań rozwiązywane przy pomocy SSN:

- autoasocjacja
- klasyfikacja
- heteroasocjacja
- detekcja regularności

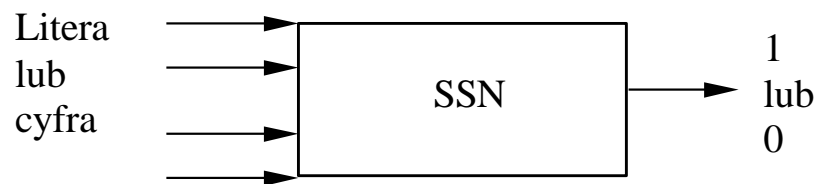
Autoasocjacja. Faktyczny obraz jest rekonstruowany przez sieć z obrazu niekompletnego i (lub) obarczonego zakłóceniami



Heteroasocjacja. Sieć realizuje odwzorowanie wejście-wyjście, np. jeśli na wejściu prezentowany jest obraz 1, to na wyjściu generowany jest obraz 2.



Klasyfikacja. Sieć klasyfikuje, czyli przypisuje do odpowiednich zbiorów (klas) różne obrazy wejściowe.



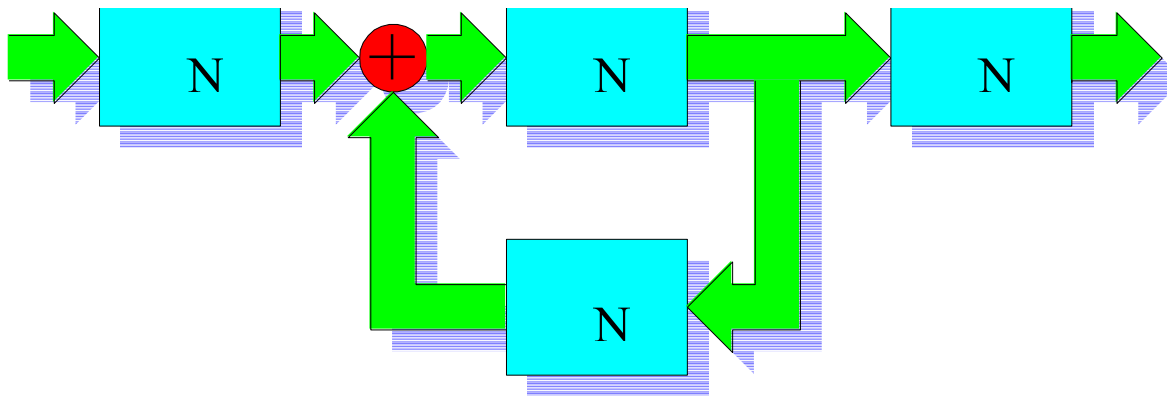
Detekcja regularności. Zadaniem sieci jest wykrywanie statystycznie istotnych cech w obrazie wejściowym. Nie ma zadanego *a priori* zbioru kategorii, według których klasyfikuje się obrazy wejściowe. System samodzielnie tworzy swoją własną reprezentację obrazów wejściowych.

Podział sieci ze względu na architekturę:

- jednokierunkowe, tj. o jednym kierunku przepływu sygnałów (“feedforward networks”)
- sieci rekurencyjne, tj. ze sprzężeniami zwrotnymi (wiązane często z nazwiskiem Hopfielda),
- komórkowe

Sieci rekurencyjne

- sieci o połączeniach dwukierunkowych między elementami przetwarzającymi lub ze sprzężeniami zwrotnymi.



Bloki N_1, \dots, N_4 : - układy wielu sztucznych neuronów połączonych między sobą.

UCZENIE SIECI

Proces uczenia sieci = wyznaczanie współczynników wag połączeń $\{w_{ij}\}$ między i -tymi i j -tymi elementami przetwarzającymi (neuronami)

Techniki uczenia sieci:

- Uczenie nadzorowane lub inaczej uczenie z nauczycielem, np. metoda propagacji wstecznej; zadania 1, 2, 3.
- Uczenie nienadzorowane lub inaczej bez nauczyciela, np. algorytm uczenia konkurencyjnego; zadanie 4.

Uczenie z nauczycielem:

przykłady poprawnego działania (ciąg uczący)

tj. konkretne sygnały wejściowe i wyjściowe (wymagana odpowiedź sieci dla danej konfiguracji danych wejściowych).

Uczenie bez nauczyciela (samouczenia sieci).

Polega na podawaniu na wejście sieci szeregu przykładowych zbiorów danych wejściowych bez podawania jakiegokolwiek informacji dotyczącej pożądaných czy oczekiwanych sygnałów wyjściowych.

Zastosowania

• diagnostyka elektronicznych układów	• interpretacja badań biologicznych
• badania psychiatryczne	• prognozy cen
• prognozy giełdowe	• analiza badań medycznych
• prognozowanie sprzedaży	• planowanie remontów maszyn
• poszukiwanie ropy naftowej	• prognozowanie postępów w nauce
• typowanie na wyścigach konnych	• analiza problemów produkcyjnych
• optymalizacja działalności handlowej	• analiza spektralna
• optymalizacja utylizacji odpadów	• dobór surowców
• selekcja celów śledztwa w kryminalistyce	• dobór pracowników
• sterowanie procesów przemysłowych	• ...

4.3. Logika rozmyta

Cel: przedstawienie wiedzy w formie lingwistycznej

Informacja lingwistyczna - operuje pojęciami opisowymi (deskryptory: duży, mały, średni, bardzo duży, bardzo mały, itp.)

Cecha: nieprecyzyjność typu rozmytości

Klasyczna teoria zbiorów:

podstawowe pojęcia: zbiór, element zbioru

$$x \in A, x \notin A$$

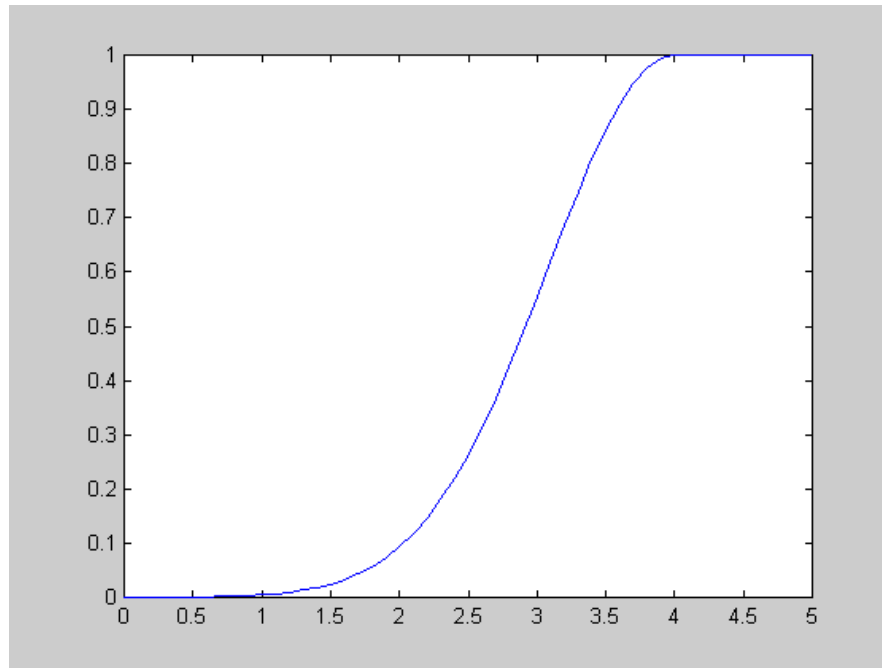
zbiór (A): skończony, policzalny, niepoliczalny
opisany przez wyszczególnienie lub przez
określenie cechy przynależności

Teoria zbiorów rozmytych (fuzzy sets)

Podstawowe pojęcie – relacja przynależności elementu do zbioru ($x \in X$)

Zbiorem rozmytym F określonym w przestrzeni X nazywamy zbiór par:

$$F = \{x, \mu_F(x)\}, \quad \forall x \in X$$



$$F = \{x, \mu_F(x)\}, \quad \forall x \in X$$

$$F = \mu_F(x_1)/x_1 + \dots + \mu_F(x_n)/x_n = \sum_{i=1}^n \mu_F(x_i)/x_i$$

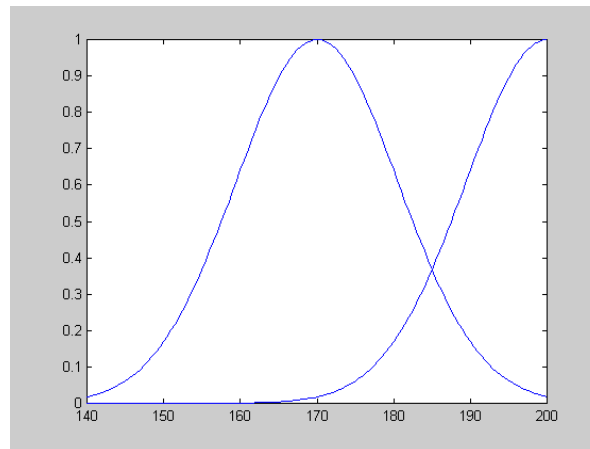
Operacje na zbiorach rozmytych

Przecięcie:

$$F1 \cap F2 \quad \mu_{F1 \cap F2}(x) = \min\{\mu_{F1}(x), \mu_{F2}(x)\} = \mu_{F1}(x) \cap \mu_{F2}(x)$$

Suma:

$$F1 \cup F2 \quad \mu_{F1 \cup F2}(x) = \max\{\mu_{F1}(x), \mu_{F2}(x)\} = \mu_{F1}(x) \cup \mu_{F2}(x)$$



Zmienna liczbowa

Trójka uporządkowana $(X, U, R(x))$

gdzie:

X - nazwa zmiennej,

U – obszar rozważań,

$R(x)$ – ograniczenie nałożone na wartości $u \in U$

Równanie przypisujące: $x=U: R(x)$

Przykład

X – wzrost

$U = \{30, 51, \dots, 250\}$

$R = \{0, 1, 2, \dots, 210\}$

Zmienna rozmyta: analogicznie jak zmienna liczbowa, przy czym stopień, z którym jest spełnione równanie przypisujące, zdefiniowany jest jako:

$$c(u) = \mu_{R(X)}(u), u \in U$$

Przykład:

X – niski wzrost

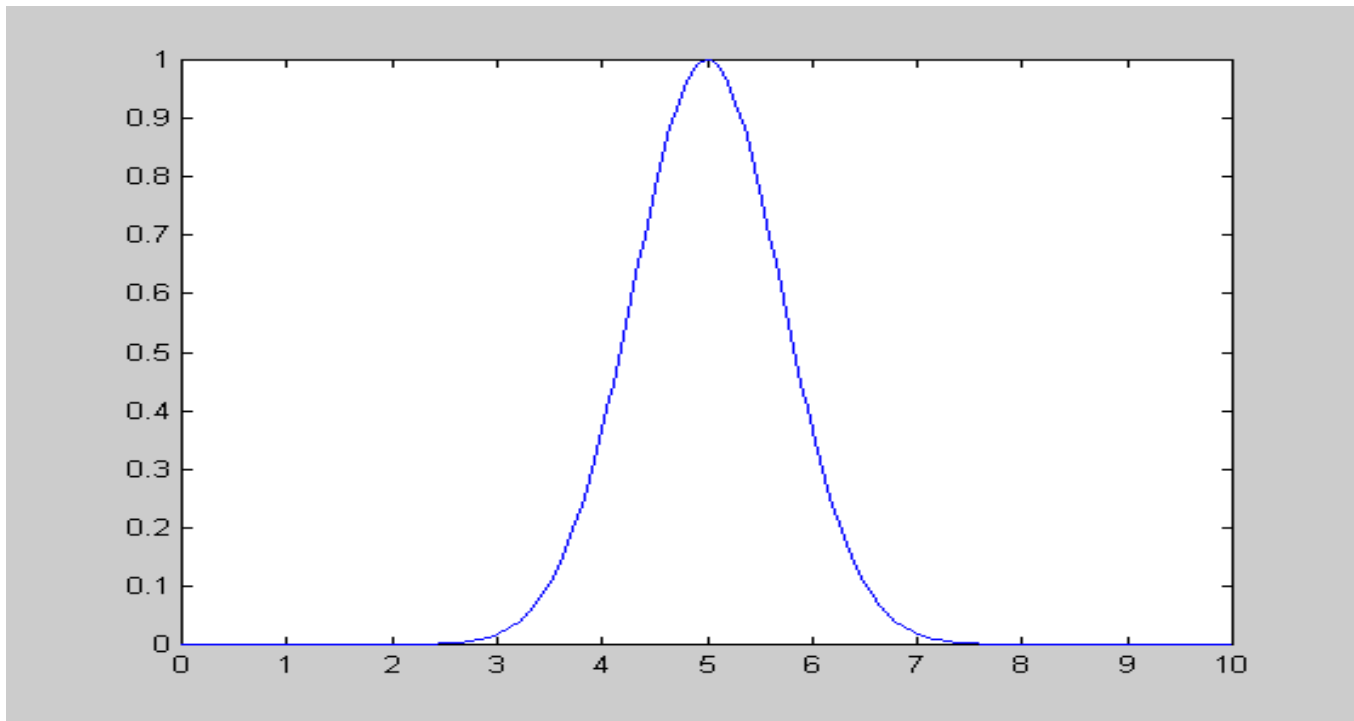
Zmienna lingwistyczna: przyjmuje zmienne rozmyte jako swe wartości

Przykład:

Zmienna lingwistyczna WZROST (bardzo niski, niski średni, wysoki, bardzo wysoki)

Liczby rozmyte

Zbiory rozmyte zdefiniowane w zbiorze liczb rzeczywistych -przestrzeń \mathbb{R} -określane są mianem liczb rozmytych, na przykład liczba rozmyta “około 5”. Liczbą rozmytą A w zbiorze liczb rzeczywistych \mathbb{R} jest zbiór rozmyty o ciągłej funkcji przynależności $\mu_A(x): \mathbb{R} \rightarrow [0,1]$.



Liczba rozmyta “około 5”

Zastosowania:

- wnioskowanie rozmyte
- sterowniki rozmyte
- obliczenia

Wnioskowanie rozmyte (przybliżone)

Jeśli przesłanka logiczna *to* konkluzja

Reguła wnioskowania

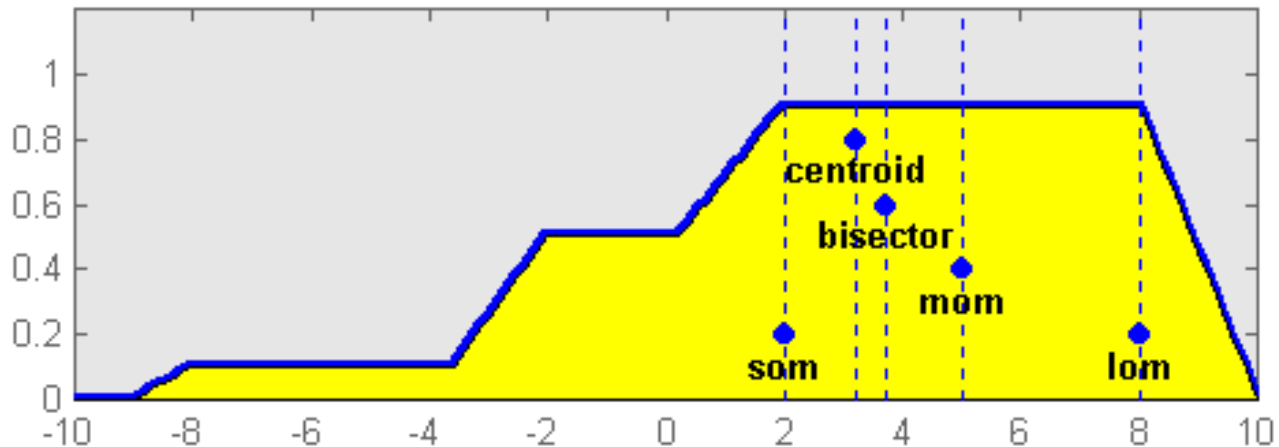
Przesłanka	A
Implikacja	$A \rightarrow B$
Wniosek	B

Przykład (autopilot)

Przesłanka	odchylenie od zadanego kursu jest duże
Implikacja	jeżeli odchylenie od zadanego kursu jest duże to wychylenie steru duże
Wniosek	wykonać: duże wychylenie steru

Metody defuzyfikacji:

1. COA - centroid of area,
2. BOA - bisector of area,
3. MOM - mean of maximum,
4. SOM - smallest of maximum,
5. LOM - largest of maximum.



$$WS = 6.0979e-001$$

4.4. Algorytmy genetyczne

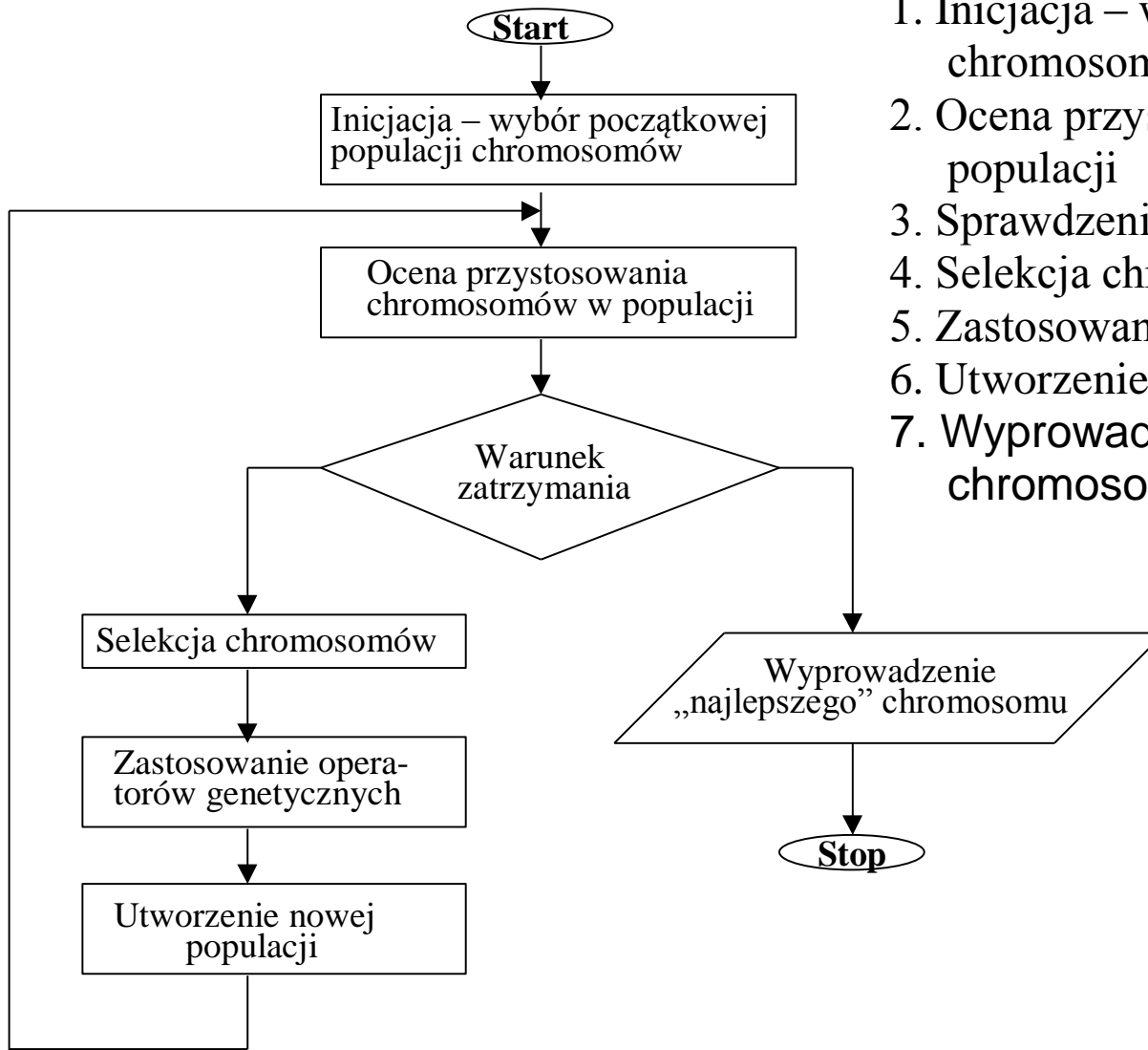
Algorytm genetyczny – metoda rozwiązywania problemów (głównie zagadnień optymalizacyjnych) wzorowana na naturalnej ewolucji

Różnice w stosunku do tradycyjnych metod optymalizacji:

- nie przetwarza bezpośrednio parametrów zadania, a ich zakodowaną postać
- prowadzi przeszukiwanie wychodząc nie z pojedynczego punktu a z pewnej populacji punktów.
- korzysta tylko z funkcji celu , a nie z jej pochodnych lub innych informacji
- stosuje probabilistyczne a nie deterministyczne reguły wyboru.

Podstawowe pojęcia

- populacja - zbiór osobników o określonej liczebności
- osobniki populacji - zakodowane w postaci chromosomów zbiory parametrów zadania (punkty przestrzeni poszukiwań)
- chromosomy - łańcuchy (ciągi kodowe), uporządkowane ciągi genów
- genotyp - (struktura) zespół chromosomów danego osobnika
- fenotyp - zestaw wartości odpowiadający danemu genotypowi
- allel - wartość genu (wartość cechy)
- locus - pozycja, wskazuje położenie genu w łańcuchu czyli chromosomie
- f. przystosowania - stanowi miarę dopasowania (oceny) danego osobnika w (fitness function) populacji, - pozwala wybrać osobniki najlepiej przystosowane)



Klasyczny algorytm genetyczny:

1. Inicjacja – wybór początkowej populacji chromosomów
2. Ocena przystosowania chromosomów w populacji
3. Sprawdzenie warunków zatrzymania
4. Selekcja chromosomów
5. Zastosowanie operatorów genetycznych
6. Utworzenie nowej populacji
7. Wyprowadzenie najlepszego chromosomu

Przykład 1

Funkcja: $f(x) = 2x^2 + 1$

przyjmuje wartości całkowite z przedziału $\langle 0, 31 \rangle$

Zadanie: Znaleźć wartość x , dla której funkcja przyjmuje wartość maksymalną

Parametr zadania: x ;

Zbiór $\langle 0, 31 \rangle$: przestrzeń poszukiwań, rozwiązanie , fenotyp

Kodowanie parametru x (łańcuchy, chromosomy)

00000	00001	00011	00100	00101	00110	...	01111
10000	...	11001	11011	11100	11101	11110	11111

Uwaga: W tym przypadku chromosom jest równocześnie genotypem

Populacja: osobniki wybrane spośród chromosomów np. zbiór 6 chromosomów

Funkcja przystosowania: $f(x)$

Przystosowanie poszczególnych chromosomów określa wartość funkcji $f(x)$ dla x odpowiadających tym chromosomom (fenotypy odpowiadające określonym genotypom) - w przykładzie: $f(x) = 2x^2 + 1$

1. Inicjacja – wybór początkowej populacji chromosomów

Utworzenie populacji początkowej – losowy wybór żądanej liczby chromosomów (osobników) reprezentowanych przez ciągi binarne określonej długości.

Populacja $N=6$ chromosomów wybranych losowo:

$$\begin{array}{lll} \text{ch}_1 = [10011] & \text{ch}_3 = [00111] & \text{ch}_5 = [10101] \\ \text{ch}_2 = [00011] & \text{ch}_4 = [10011] & \text{ch}_6 = [11101] \end{array}$$

Odpowiadające im fenotypy:

$$\begin{array}{lll} \text{ch}_1^* = 19 & \text{ch}_3^* = 7 & \text{ch}_5^* = 8 \\ \text{ch}_2^* = 3 & \text{ch}_4^* = 21 & \text{ch}_6^* = 29 \end{array}$$

2. Ocena przystosowania chromosomów w populacji

Wartości funkcji przystosowania ($f(x)$) dla poszczególnych chromosomów w populacji:

$$F(\text{ch}_1) = 723$$

$$F(\text{ch}_3) = 99$$

$$F(\text{ch}_5) = 129$$

$$F(\text{ch}_2) = 19$$

$$F(\text{ch}_4) = 883$$

$$F(\text{ch}_6) = 1683$$

3. Sprawdzenie warunków zatrzymania

Warunki zatrzymania zależą od konkretnego zastosowania, np.:

- po uzyskaniu żądanej wartości optymalnej
- jeśli dalsze działanie nie poprawia już uzyskanej najlepszej wartości

4. Selekcja chromosomów

Wybranie z puli chromosomów, na podstawie obliczonych wartości funkcji przystosowania, chromosomów, które będą brały udział w tworzeniu potomków do następnego pokolenia - generacji

Metody:

- ruletki
- selekcji turniejowej
- selekcji rankingowej
- ...

Metoda ruletki:

$$v(ch_i) = p_s(ch_i) * 100\%$$

gdzie

$$p_s(ch_i) = \frac{F(ch_i)}{\sum_{i=1}^N F(ch_i)}$$

$p_s(ch_i)$ - prawdopodobieństwo tego, że chromosom ch_i zostanie wybrany
(prawdopodobieństwo selekcji chromosomu ch_i)

$v(ch_i)$ – wycinek koła wyrażony w procentach

5. Zastosowanie operatorów genetycznych

Celem jest utworzenie nowej populacji, stanowiącej populację potomków otrzymanych z wybranej metodą selekcji populacji rodziców.

Operatory:

- krzyżowania
- mutacji

Operator krzyżowania

- kojarzenie w pary (losowo) $(ch_6, ch_4, ch_6, ch_1, ch_4, ch_6)$

ch_1

ch_4

ch_4

ch_6

ch_6

ch_6

•krzyżowanie par chromosomów rodzicielskich

-losowanie pozycji genu (locus) w chromosomie, określającego punkt krzyżowania

$$l_k=3$$

$$l_k=2$$

$$l_k=4$$

-realizacja krzyżowania

$$ch_1 = [10011] \text{ krzyżowanie} \rightarrow [10001]$$

$$ch_4 = [10101] \qquad [10111]$$

$$l_k=3$$

$$ch_4 = [10101] \text{ krzyżowanie} \rightarrow [10101]$$

$$ch_6 = [11101] \qquad [11101]$$

$$l_k=2$$

$$ch_6 = [11101] \text{ krzyżowanie} \rightarrow [11101]$$

$$ch_6 = [11101] \qquad [11101]$$

$$l_k=4$$

6. Utworzenie nowej populacji

Utworzona zostaje nowa populacja o fenotypach:

$$\begin{array}{lll} \text{ch}_1^* = 17 & \text{ch}_3^* = 21 & \text{ch}_5^* = 29 \\ \text{ch}_2^* = 23 & \text{ch}_4^* = 29 & \text{ch}_6^* = 29 \end{array}$$

i następujących wartościach funkcji przystosowania:

$$\begin{array}{lll} F(\text{ch}_1) = 579 & F(\text{ch}_3) = 883 & F(\text{ch}_5) = 1683 \\ F(\text{ch}_2) = 1059 & F(\text{ch}_4) = 1683 & F(\text{ch}_6) = 1683 \end{array}$$

Jeśli w kolejnej iteracji zostałyby wylosowane pary chromosomów

ch_4 ch_2 ch_5 ch_2 ch_6 ch_2
z punktem krzyżowania 2 lub 3, otrzymany zostanie chromosom o fenotypie 31.

8. Wyprowadzenie najlepszego chromosomu

Jeśli spełnione są warunki zatrzymania (jeśli dalsze działanie nie poprawia już uzyskanej najlepszej wartości), należy wyprowadzić wynik działania, tj. podać rozwiązanie problemu.

Najlepszym rozwiązaniem jest chromosom o największej wartości funkcji przystosowania