

## 6 Rozpoznávanie obrazu

### Rozpoznávanie objektov, rozpoznávanie obrazcov

- Rozpoznávanie obrazcov sa používa na **klasifikáciu** oblastí a objektov a predstavuje dôležitú súčasť zložitejších systémov počítačového videnia.
- Žiadne rozpoznávanie nie je možné bez znalostí. Vyžadujú sa špecifické tak o objektoch, ktoré sa rozpoznávajú ako aj všeobecnejšie znalosti o triedach objektov.

### Reprezentácia znalostí

- Reprezentácia je súhrn syntaktických a sémantických pravidiel, ktorú umožňujú popisovať veci.
- Využívajú sa techniky umelej inteligencie:
  - o Popisy a príznaky
  - o Gramatiky a jazyky
  - o Predikátová logika
  - o Produkčné pravidlá
  - o Fuzzy logika
  - o Sémantické siete
  - o Rámce, skripty
- **Popisy** reprezentujú nejaké skalárne vlastnosti objektov a nazývajú sa **príznaky**. Jeden príznak na objekt nestačí, preto sa príznaky organizujú do tzv. príznakových vektorov.

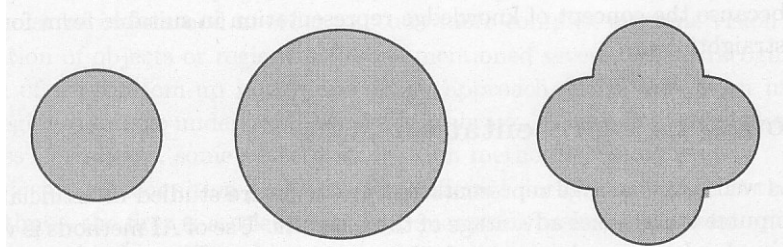


Figure 7.2: Feature description of simple objects.

- Na popis objektov na obrázku stačí dvojrozmerný vektor príznakov  $x = (\text{veľkosť}, \text{kompaktnosť})$ , čo umožní ich zaradenie do tried objektov: veľký, malý, okrúhly, neokrúhly apod.
- **Gramatiky a jazyky**. Ak sa má popísať štruktúra objektu, príznakový popis nestačí. Primitívum je najmenšia, ďalej nedeliteľná časť obrazu a štruktúralne popisy sa vytvárajú pomocou relácií medzi nimi. Môžu sa využiť reťazcové kódy, postupnosť segmentov, stromy, vo všeobecnosti grafy symbolov. Gramatiky predstavujú pravidlá ako sa z abecedy vytvárajú slová jazyka.

- **Predikátová logika.** Prináša do reprezentácie znalostí matematický formalizmus a umožňuje odvodzovať nové znalosti pomocou dôkazov a pravidiel inferencie ako sú modus ponens a rezolvenca.
- **Produkčné pravidlá.** Sú založené na dvojici podmienka - akcia. Opisujú sa všeobecnou schémou
 

*ak podmienka X platí potom je vhodná akcia Y*

pričom informácia o tom, ktorá akcia je vhodná, predstavuje znalosť. Produkčný charakter uchovávanía znalostí zaručuje, že nemusia byť zapísané všetky znalosti, pretože sa dajú odvodiť.
- **Fuzzy logika** slúži na vyjadrenie neostrých vlastností, napr. namiesto produkčného pravidla
 

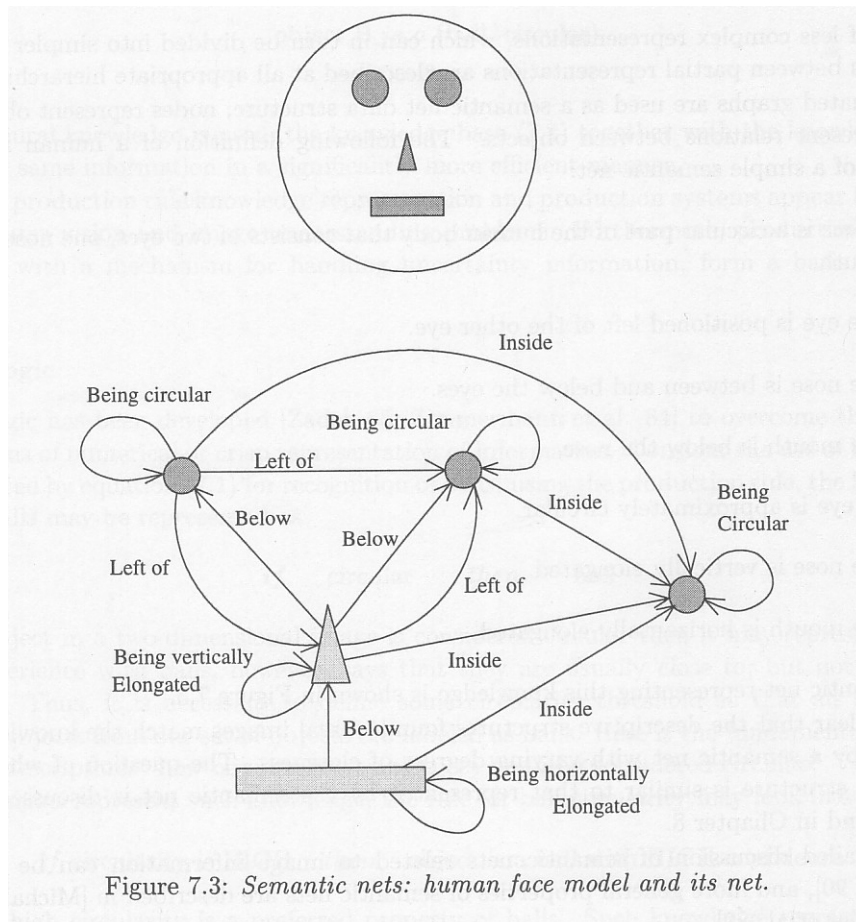
*ak okrúhly potom lopta*

sa vo fuzzy logike povie

*ak okrúhlosť je VYSOKÁ  
potom objekt je lopta s VYSOKOU určitosťou.*

Vo všeobecnosti fuzzy pravidlá majú formu

*ak X je A potom Y je B.*
- **Sémantické siete** predstavujú špeciálny prípad relačných štruktúr. Obsahujú informáciu o objektoch a o ich vzťahoch, obvykle majú hierarchickú štruktúru.



- **Rámce alebo skripty** predstavujú veľmi všeobecnú metódu na reprezentáciu znalostí, ktorá môže zahrňovať všetky uvedené princípy.

### Štatistické rozpoznávanie obrazcov

- Objekt je fyzická jednotka, obvykle reprezentovaná oblasťou v segmentovanom obraze. Množinu všetkých objektov možno rozdeliť na podmnožiny so spoločnými príznakmi, ktoré sa nazývajú triedy.
- **Rozpoznanie (klasifikácia) objektu** je založené na priradení triedy k neznámemu objektu a zariadenie, ktoré vykonáva toto priradenie sa nazýva klasifikátor. Počet tried sa obyčajne vie dopredu a obvykle sa dá odvodiť zo špecifikácie problému.
- Klasifikátor nerozhoduje o zaradení do triedy na základe objektu priamo, ale na základe vlastností objektu, ktoré sa nazývajú obrazec.

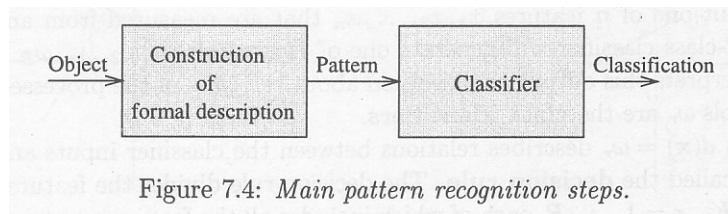


Figure 7.4: Main pattern recognition steps.

- Pre štatistické rozpoznávanie obrazcov je charakteristický **kvantitatívny** popis objektov, obyčajne sa používajú numerické príznaky. **Obrazec** je potom príznakový vektor  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ . Množina všetkých možných obrazcov sa nazýva priestor obrazcov alebo **príznakový priestor**. Triedy obrazcov vytvárajú v príznakovom priestore zhluky, ktoré je možné rozdeliť diskriminačnými (oddeľujúcimi) nadplochami.

## Neurónové siete

- Viaceré **neurónové prístupy** sú založené na kombinácii elementárnych procesorov (neurónov), z ktorých každý má viacero vstupov a jeden výstup. Každému vstupu je priradená váha a výstup je sumou váhovaných vstupov. Rozpoznávanie obrazcov je jednou z mnohých oblastí použitia neurónových sietí.

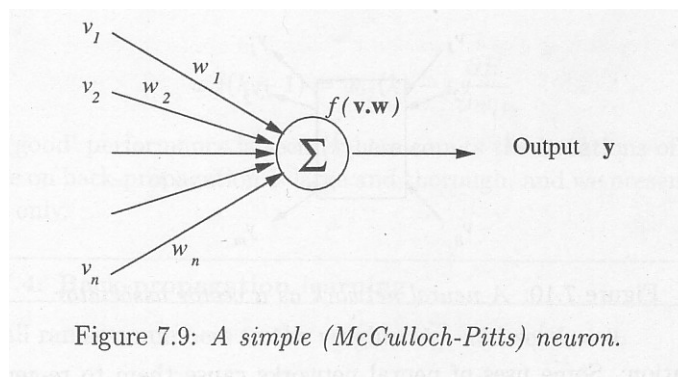


Figure 7.9: A simple (McCulloch-Pitts) neuron.

- Vstup do neurónu je definovaný ako  $x = \sum_{i=1}^n v_i w_i - \theta$ , kde  $\theta$  je nejaký prah. Výstup neurónu je spojený s transfer funkciou  $f(x)$ , napríklad

$$f(x) = 0 \text{ ak } x \leq 0 \text{ a } f(x) = 1 \text{ inak.}$$

$$f(x) = \frac{1}{1 + e^{-x}}.$$

- Príkladov využitia neurónovej siete môže byť viacero:  
 klasifikácia: ak výstupný  $m$ -rozmerný vektor je binárny a obsahuje iba jednu jednotku, tak tá reprezentuje klasifikáciu vstupu do jednej z  $m$  kategórií.

všeobecná asociácia: vstupný vektor  $\mathbf{v}$  a výstupný vektor  $\mathbf{y}$  reprezentujú obrazce v rôznych oblastiach a neuronová sieť vytvára vzťah medzi nimi. Napr. čítanie textu.

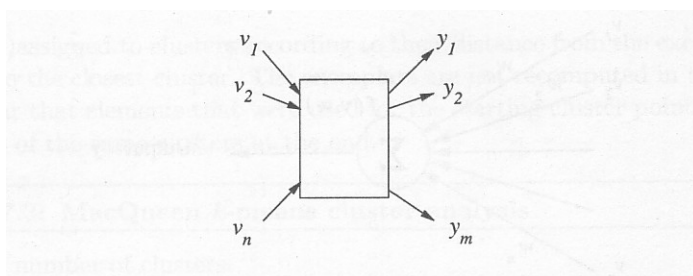


Figure 7.10: A neural network as a vector associator.

- Najprv sa používali siete bez skrytej vrstvy, ale ukázali sa ako veľmi obmedzujúce. Dá sa ukázať, že reálne stačia dve skryté vrstvy.

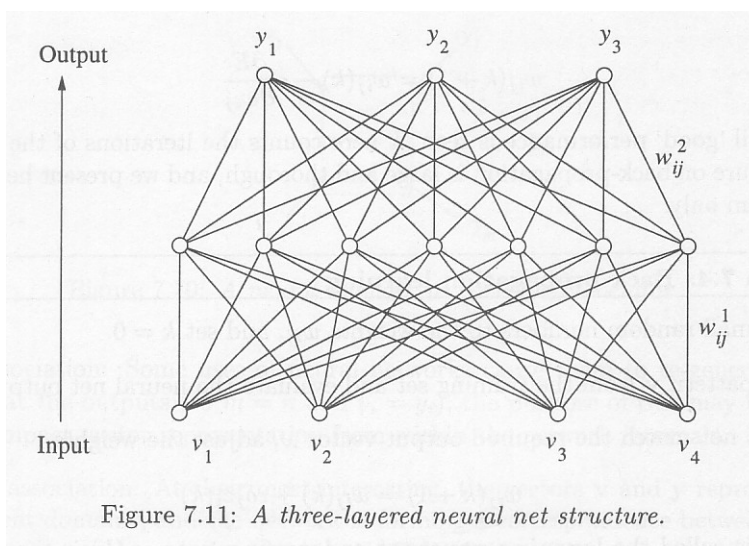


Figure 7.11: A three-layered neural net structure.

- V problémoch rozpoznávania obrazcov sa bežne používajú **dopredné siete**. Ich učenie používa trénovaciu množinu príkladov a je často založené na algoritme spätného posunu. Vyhodnocuje sa výstup siete a očakávaný výstup a počíta sa chyba založená na štvorci rozdielov

$$E = \sum_i \sum_j (y_j^i - \omega_j^i)^2 ,$$

kde  $y^i$  je aktuálny výstup a  $\omega^i$  je želaný výstup. Potom algoritmus urobí spätné

korekcie

$$w_{ij}(k+1) = w_{ij}(k) - \varepsilon \frac{\partial E}{\partial w_{ij}}$$

kde  $k$  zodpovedá počtu iterácií opráv koeficientov.

- Samoorganizujúce sa siete si nevyžadujú trénovaciu množinu na vytvorenie zhlukov z rozpoznávaných obrazcov. Existujú rôzne typy takýchto sietí, najznámejšie sú Kohonenove mapy príznakov. Táto sieť má vrstvu neurónov, z ktorých každý je

spojený so všetkými  $n$  zložkami vektora príznakov. Ten neurón, ktorý má najväčší vstup, je považovaný za víťaza a jeho váhy reprezentujú výstup siete. Pre každý analyzovaný vektor  $\mathbf{v}$  sa nájde neurón s najväčším vstupom:

$$j^* = \max_j \sum_i w_{ij} v_i .$$

Potom pre všetky neuróny  $n_j$  v okolí  $r$  neurónu  $n_{j^*}$  sa vypočítajú nové váhy s krokom  $\alpha > 0$ , tak že  $w_{ij} = w_{ij} + \alpha(v_i - w_{ij})$  s postupnou redukciou  $r$  a  $\alpha$ .

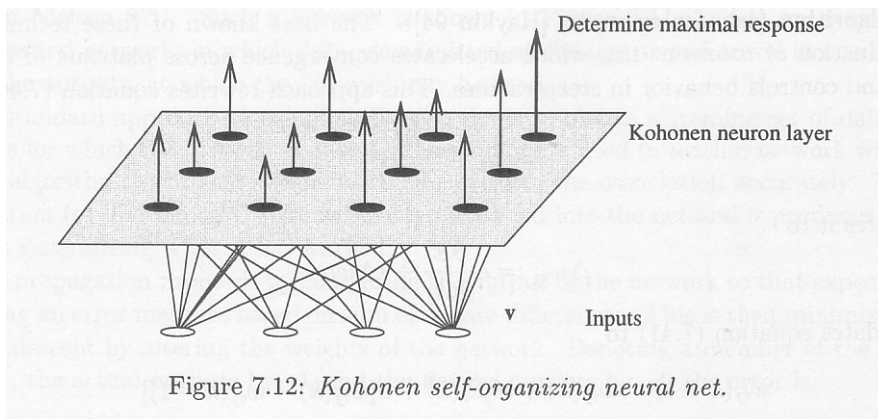


Figure 7.12: Kohonen self-organizing neural net.

- Hopfieldove neurónové siete nemajú priradené vstupy a výstupy, ale aktuálna konfigurácia siete predstavuje jej stav. Hopfieldova sieť pracuje ako asociatívna pamäť, v ktorej sa ukladajú príklady.

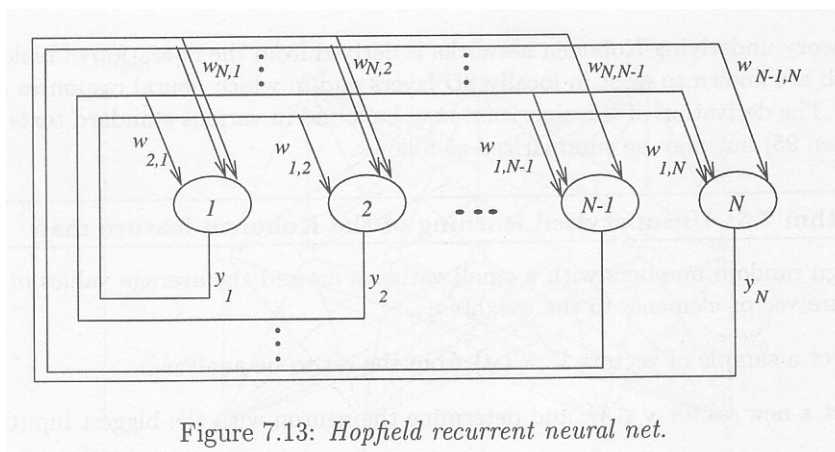


Figure 7.13: Hopfield recurrent neural net.

Váhy sa nezískavajú učením, ale sa vypočítajú pri inicializácii siete na základe známej množiny etalónov tak, že

$$w_{ij} = \sum_r (v_i^r v_j^r) \quad (i \neq j),$$

kde  $w_{ij}$  je váha spojenia medzi vrcholmi  $i$  a  $j$  a  $v_i^r$  je  $i$ -ty prvok  $r$ -tého exemplára. Potom sa neznámy vektor  $\mathbf{x}$  aplikuje ako  $\mathbf{y}(0)$  a iteruje sa až do konverencie podľa vzorca

$$y_j(k+1) = f \left\{ \sum_{i=1}^N [w_{ij} y_i(k)] \right\}$$

### Syntaktické rozpoznávanie obrazcov

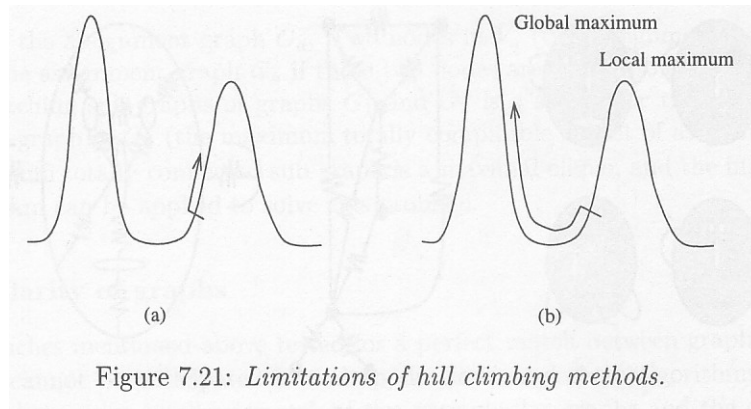
- Pre syntaktické rozpoznávanie obrazcov je charakteristický kvalitatívny popis objektov. Syntaktický popis treba použiť vtedy, ak príznakový popis nie je schopný popísať zložitosť objektu alebo keď sa dá objekt zapísať ako hierarchická štruktúra pozostávajúca z jednoduchších častí. Elementárnymi vlastnosťami syntakticky popísaných objektov sú primitíva. Na popis vzťahov medzi primitívami objektu sa používajú relačné štruktúry. Neexistuje postup na vytvorenie primitív a ich vzťahov.
- Syntaktické rozpoznávanie pozostáva z nasledovných krokov:
  - a. Definuj primitíva a vzťahy medzi nimi
  - b. Zostroj gramatiku pre každú triedu objektov.
  - c. Pre každý objekt vytiahni primitíva rozpoznaj ich a vzťahy medzi nimi a zostroj slovo reprezentujúce objekt.
  - d. Na základe syntaktickej analýzy zaraď objekt do tej triedy, ktorej gramatika ho generuje.

### Rozpoznávanie ako porovnávanie zhody dvoch grafov

- Porovnávanie zhody (matching) modelu a objektu popísaných grafom sa dá využiť na rozpoznávanie.

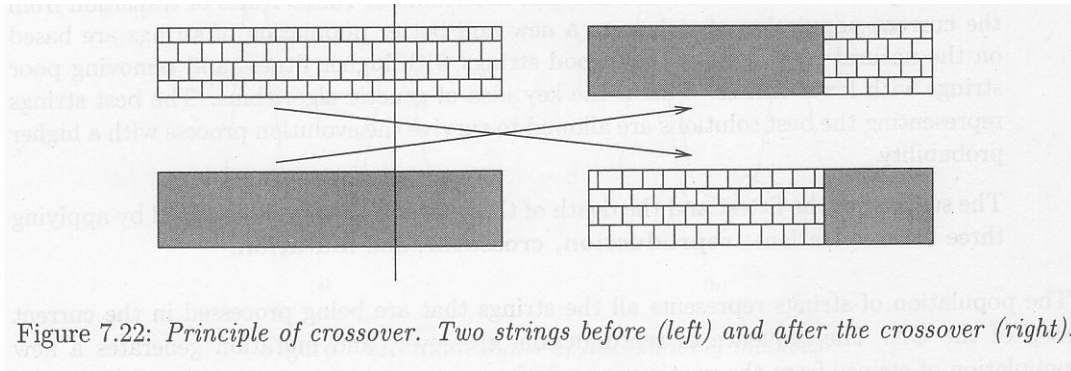
### Optimalizačné techniky v rozpoznávaní

- Optimalizačné problémy hľadajú minimalizáciu alebo maximalizáciu nejakej objektívnej funkcie. Návrh tejto objektívnej funkcie je kľúčovým faktorom fungovania optimalizačných algoritmov. Chceme, aby interpretácia obrazu bola *najlepšia*.
- Najbežnejšie prístupy k optimalizácii sú horolezecké metódy, pri ktorých môže hľadanie ľahko skončiť v lokálnom maxime a globálne maximum algoritmus nenájde.

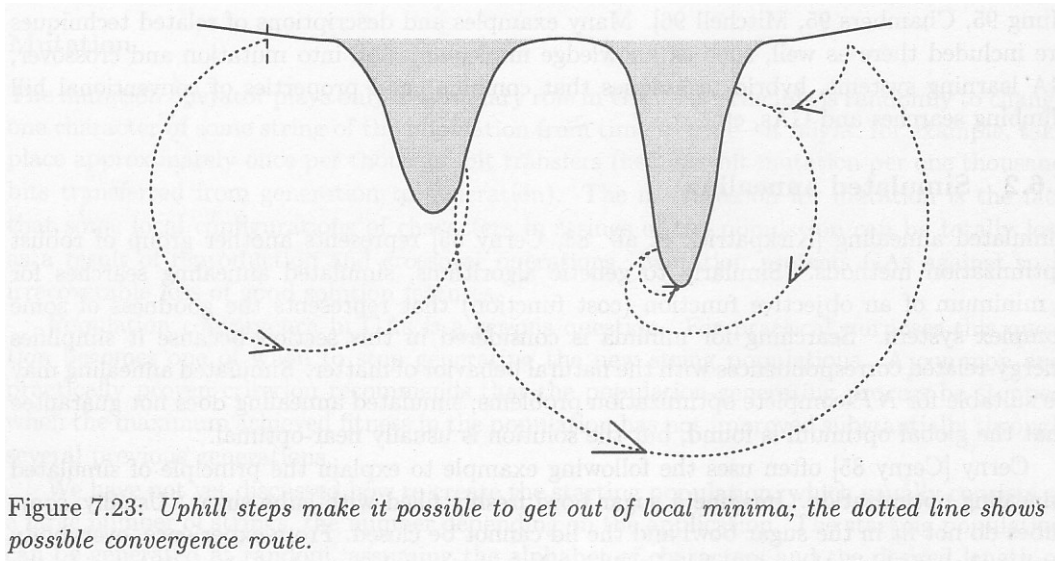


- Existuje viacero metód na nájdenie globálneho maxima, napr. začať hľadať horolezeckým algoritmom vo viacerých bodoch prehľadávacieho priestoru, použiť dynamické programovanie, alebo použiť genetické algoritmy či simulované žihanie.
- **Genetické algoritmy** používajú evolučné mechanizmy prežitia najschopnejšieho na hľadanie maxima objektívnej funkcie. Nezaručujú nájdenie globálneho optima, ale empirické príklady ukazujú, že nájdené riešenie býva veľmi blízko.
- Možné riešenia sú reprezentované reťazcami, často sa používa binárna abeceda. Genetické algoritmy hľadajú v populácii možných riešení, nie jednotlivé riešenie, to dáva väčšiu šancu nájsť globálne optimum. Kritečná funkcia vyjadruje kvalitu konkrétneho reťazca, tá sa nazýva fitness (spôsobilosť). Transformácie jednej populácie na druhú sú pravdepodobnostné, nie deterministické. Hlavná myšlienka je založená na podpore dobrých reťazcov s vysokou spôsobilosťou a odstraňovaní zlých reťazcov s nízkou spôsobilosťou. Postupnosť **reprodukcie, kríženia a mutácie** vytvára novú populáciu reťazcov z predchádzajúcej populácie. Najschopnejší (najsposobilejší) reťazec reprezentuje konečné riešenie.
- Operátor **reprodukcie** je zodpovedný za prežitie najschopnejších a smrť ostatných na pravdepodobnostnom princípe. Kopíruje reťazce s najvyšším ohodnotením do ďalšej populácie. Môžu sa kopírovať aj viacnásobne. Celkový počet reťazcov sa nemení, preto každá ďalšia generácia má vyššie priemerné ohodnotenie.
- Hlavná myšlienka operátora **kríženia** je náhodné spojenie častí dvoch reťazcov na vytvorenie dvoch nových reťazcov. Nie všetky reprodukované reťazce podliehajú kríženiu, opäť sa to deje na pravdepodobnostnom princípe, ktorý odhaduje, koľko reťazcov prejde krížením.





- Operátor **mutácie** hrá sekundárnu úlohu. Jeho princípom je z času na čas náhodne zmeniť jeden znak v nejakej reťazci. Môže to byť napr. jeden z tisíca spracovaných bitov.
- Konvergencia v genetických algoritmoch je vážny problém. Súvisí s otázkou, kedy zastaviť genetický algoritmus: vtedy, keď sa priemerná hodnota ohodnotenia nezlepšila niekoľko po sebe idúcich generácií. Ako zostrojiť prvotnú populáciu: vo všeobecnosti náhodne, ale je výhodné využiť v nej apriórnu znalosť, ak je k dispozícii.
- **Simulované žihanie** spája dva základné optimalizačné princípy, rozdeľuj a panuj a iteratívne zlepšovanie (horolezecký algoritmus). Táto kombinácia sa vyhýba uviaznutiu v lokálnom optime.
- Nezaručuje nájdenie minima, ale riešenie je blízko optimálneho. Vychádza z inšpirácie z termodynamiky a chladnutia telies a vylepšil ho slovenský fyzik stále pôsobiaci na FMFI UK Doc. Dr. Vladimír Černý. Pri vysokej teplote sa môže ľubovoľný stav stať nasledujúcim, pri nízkej teplote iba ten, ktorý má nízku energiu. Teplota sa kontrolovane znižuje (nemôže sa ani príliš rýchlo, ani príliš pomaly, lebo sa nedosiahne optimum). Iteračné kroky smerom dolu sa kombinujú s krokmi smerom hore, čo umožňuje sa vyhnúť sa lokálnemu minimu. Nevie sa presne, aký treba zvoliť krok, ako znižovať teplotu, resp. ako dlho pokračovať, vie sa iba, že žihanie musí pokračovať dosť dlho.
- Využitie simulovaného žihania je pri detekcii hranice, textúrnej segmentácii, hľadaní hrán.



## Fuzzy systémy

- Fuzzy systémy sú schopné reprezentovať rôznorodé, nepresné, neurčité a neostré znalosti alebo informácie. Používajú kvalifikátory, ktoré sú veľmi podobné ľudskému vyjadrovaniu znalostí, *svetlý, veľký, okrúhly*.
- **Fuzzy množiny** reprezentujú vlastnosti fuzzy priestorov. Funkcie príslušnosti vyjadrujú neostrosť popisu a určujú stupeň istoty o príslušnosti určitého prvku do konkrétnej fuzzy množiny. Tvar funkcií príslušnosti sa dá modifikovať množinou fuzzy posunov (hedges). Fuzzy posuny a ich fuzzy množina tvoria jednu sémantickú jednotku, ktorá sa nazýva lingvistická premenná.
- Fuzzy množina  $S$  vo fuzzy priestore  $X$  je množina usporiadaných dvojíc

$$S = \{(x, \mu_S(x)) \mid x \in X\},$$

kde  $\mu_S(x)$  reprezentuje stupeň príslušnosti  $x$  do  $S$ .

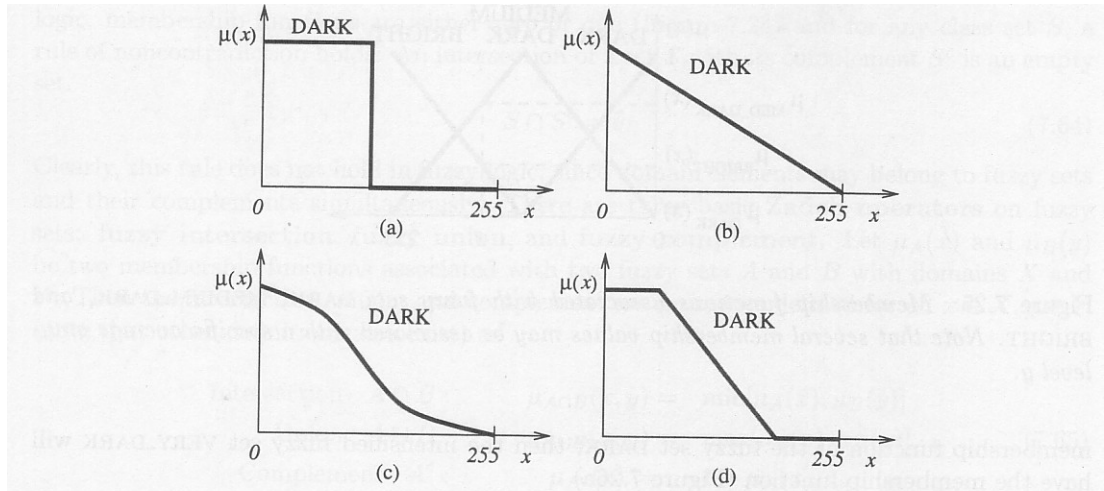


Figure 7.24: Crisp and fuzzy sets representing cloud regions of the same size and roundness, varying average gray-level  $g$ : (a) crisp set showing the Boolean nature of the DARK set; (b) fuzzy set DARK; (c) another possible membership function associated with the fuzzy set DARK; (d) yet another possible membership function.

- Jedna hodnota úrovně šedej môže patriť do viacerých fuzzy množín.

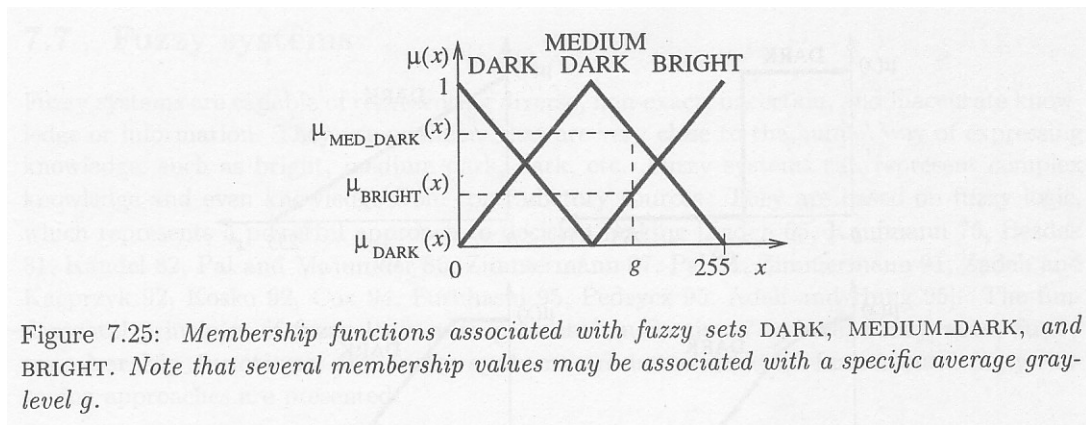


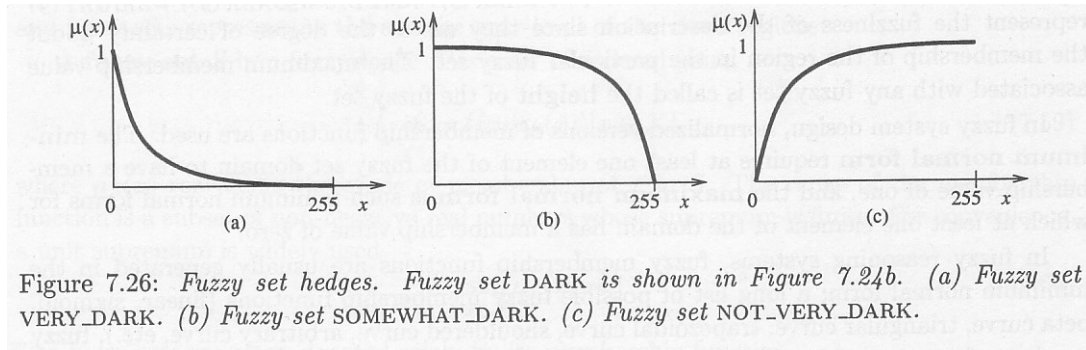
Figure 7.25: Membership functions associated with fuzzy sets DARK, MEDIUM\_DARK, and BRIGHT. Note that several membership values may be associated with a specific average gray-level  $g$ .

- Fuzzy posuny môžu modifikovať, zintenzívňovať, zmiernovať, vytvárať doplnok apod. Príklady sú nasledovné:

$$\mu_{VELMI\_TMAVY}(x) = \mu_{TMAVY}^2(x)$$

$$\mu_{TROCHU\_TMAVY}(x) = \sqrt{\mu_{TMAVY}(x)}$$

$$\mu_{VELMI\_VELMI\_TMAVY}(x) = \mu_{TMAVY}^2(x) \cdot \mu_{TMAVY}^2(x) = \mu_{TMAVY}^4(x)$$



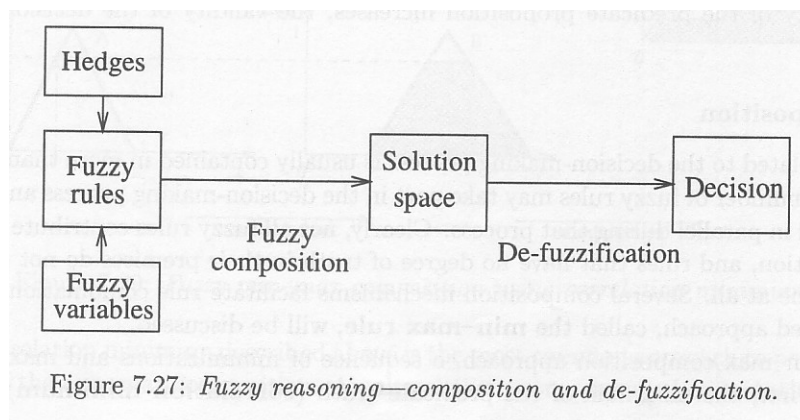
- Fuzzy operátory sú prienik, zjednotenie a komplement

prienik  $A \cap B$ :  $\mu_{A \cap B}(x, y) = \min [\mu_A(x), \mu_B(x)]$

zjednotenie  $A \cup B$ :  $\mu_{A \cup B}(x, y) = \max [\mu_A(x), \mu_B(x)]$

komplement  $A^c$ :  $\mu_{A^c}(x) = 1 - \mu_A(x)$

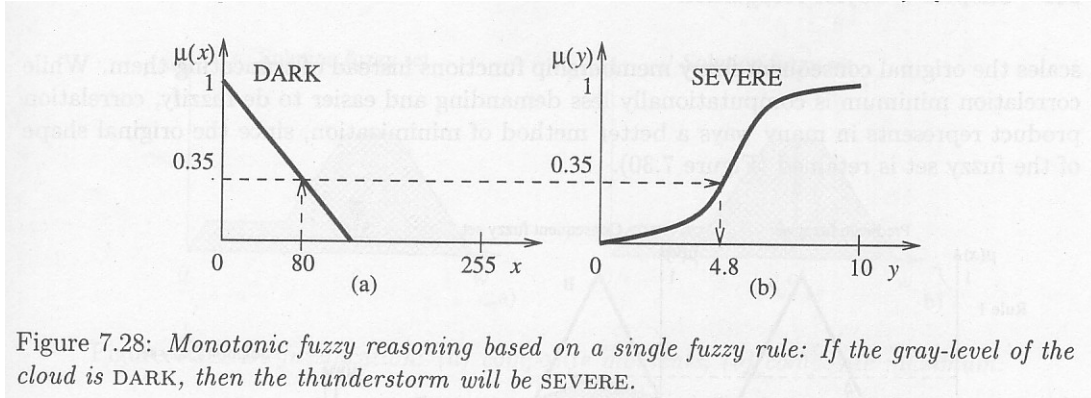
- **Fuzzy ak-tak pravidlá** reprezentujú fuzzy asociatívnu pamäť, v ktorej sa ukladajú znalosti. Nepodmienené fuzzy pravidlo  $x$  je  $A$ . Podmienené fuzzy pravidlo má tvar *ak*  $x$  je  $A$  potom  $w$  je  $B$ , kde  $A$  a  $B$  sú lingvistické premenné a  $x$  a  $w$  sú skaláry z ich definičného oboru.
- Vo fuzzy uvažovaní sa informácie obsiahnuté v jednotlivých fuzzy množinách kombinujú a vytvára sa rozhodnutie. Funkčné vzťahy určujúce stupeň príslušnosti do príslušných fuzzy oblastí sa nazýva metódou kompozície a jeho výstupom je definícia fuzzy priestoru riešenia. Na získanie riešenia sa vykonáva de-fuzzifikácia. Proces dekompozície a defuzzifikácie sú základom fuzzy uvažovania.



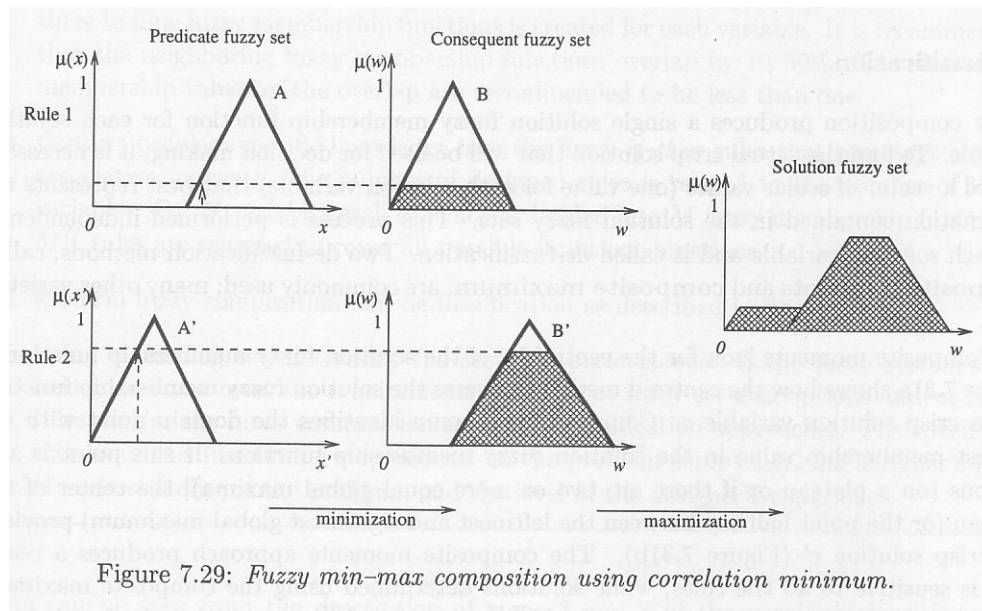
- Ak uvažujeme podmienené fuzzy pravidlá, je niekoľko spôsobov, ako dospieť k rozhodnutiu. Jedna možnosť je monotónne fuzzy uvažovanie a druhá je min-max pravidlo.
- **Monotónne fuzzy uvažovanie** ukážeme na príklade pravidla, ktoré sa vzťahuje na oblaky a ich vzťah ku búrke

ak  $x$  je *TMAVY* potom  $w$  je *SILNY*.

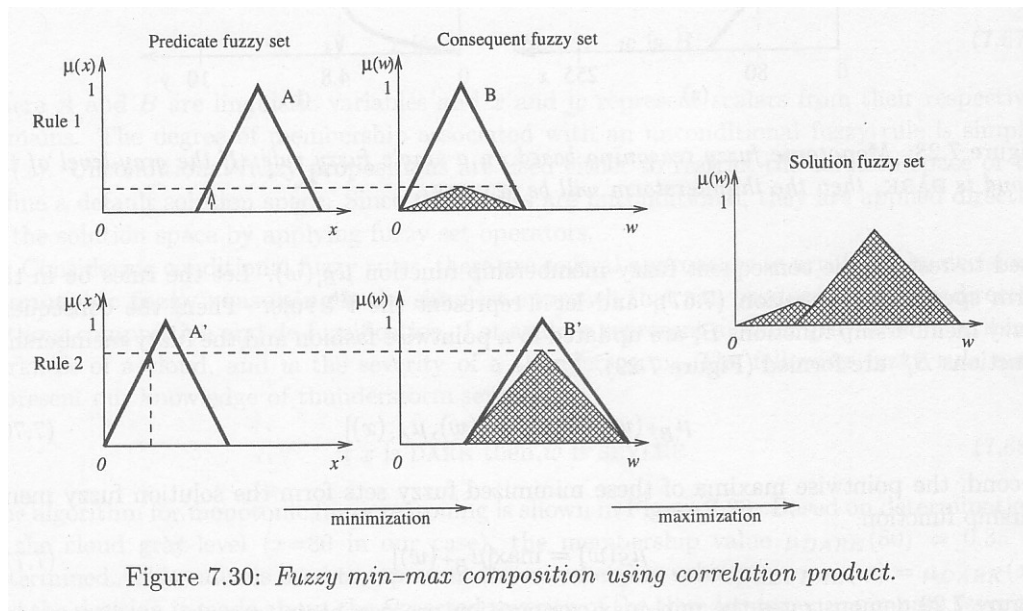
Rovnaký prístup je možné použiť pre komplexnejšie predikáty, ktoré sú vytvorené na základe prieniku, zjednotenia a komplementu.



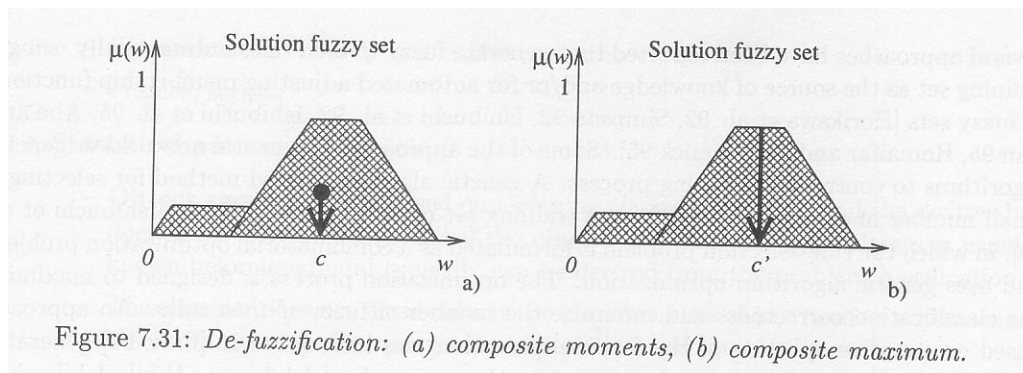
- Fuzzy kompozícia je spôsob rozhodovania v prípade viacerých pravidiel. Kompozícia pravidiel sa robí tzv. min-maxovým pravidlom. Najprv sa použije minimum pravdy predpokladu (korelačné minimum)  $\mu_{A_i}(x)$ , ktorým sa ohraničí funkcia príslušnosti dôsledku  $\mu_{B_i}(w)$ . Potom  $\mu_{B_i^+}(w) = \min[\mu_{B_i}(w), \mu_{A_i}(x)]$  následne sa robí  $\mu_S(x) = \max_i[\mu_{B_i^+}(w)]$ .



- Druhá možnosť je použiť korelačný súčin, ktorý viac zachováva tvar pôvodnej funkcie.

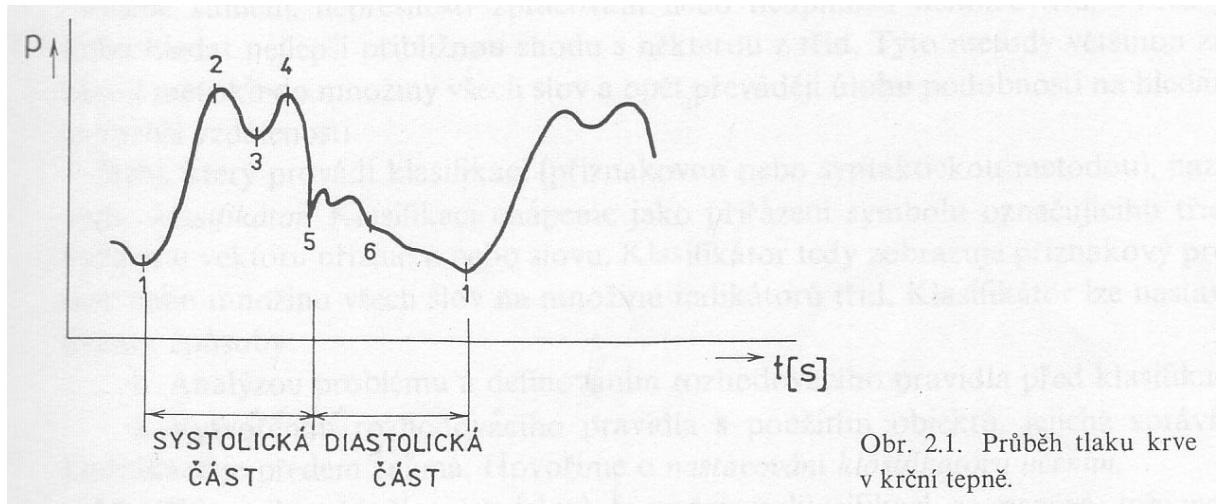


- De-fuzzifikácia je proces, pomocou ktorého získame z funkcie príslušnosti jedno riešenie pre každú premennú. Na nájdenie jednej hodnoty sa používa buď princíp kompozitného momentu (vyráta sa ťažisko a učí sa príslušná hodnota premennej) alebo kompozitné maximum (ak je viac hodnôt, berie sa stred).



- Pri nasadení v rozpoznávaní obrazcov vstupmi sú obrázky a výstupmi rozhodnutia.

- motivácia k príznakovým a syntaktickým metódam – klasifikácia na pacientov, ktorí trpia určitou chorobou srdca a tých ostatných. Priebeh tlakovej vlny v krčnej tepne je na nasledujúcom obrázku (systolická časť zodpovedá sťahu srdcového svalu, diastolická tej fáze, keď sa srdce plní krvou).



- čo by mohli byť príznaky, ktoré budú tvoriť príznakový vektor?
  - ako by mohol vyzerat' vhodný štruktúrálny popis?
- 
- príznaky – časová stredná hodnota priebehu, maximálne prevýšenie vzhľadom na túto hodnotu, plocha pod kvadrátom krivky apod.
  - syntaktický popis – využíva pulz (úseky 2 a 4), zákmity (úseky 5 a 6) apod.
  - klasifikátor možno nastaviť dvoma spôsobmi:
    - 1) analýzou problému a definovaním rozhodovacieho pravidla pred klasifikáciou
    - 2) vytvorením rozhodovacieho pravidla učením s použitím objektov, ktorých správna klasifikácia je známa
  - trénovacia množina môže nekonečná alebo konečná
  - potrebu nekonečných trénovacích množín nahradzame pravdepodobnostným prístupom