Rozpoznávání a klasifikace

14. – 21. 3. 2007

Rozpoznávání a klasifikace

- předmětů
- jevů
- situací

. . .

Obecně hovoříme o rozpoznávání a klasifikaci **objektů** .

Reprezentace objektů

- prostřednictvím signálů
  - jednorozměrných (zvuky, hudba, řeč, …)
  - dvourozměrných (snímky scén, …)
  - vícerozměrných (obecné)

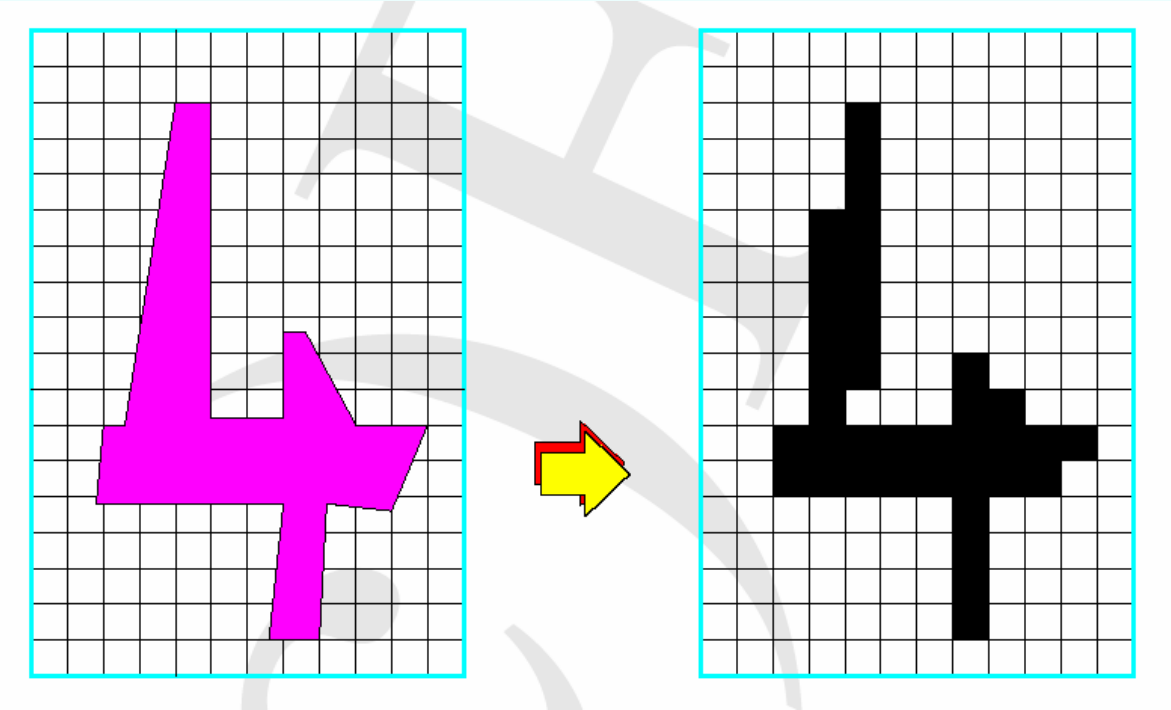
Reprezentace objektů

- prostřednictvím signálů
  - jednorozměrných (zvuky, hudba, řeč, …)
  - dvourozměrných (snímky scén, …)
  - vícerozměrných (obecné)

Podle charakteru reprezentace objektů signály rozlišujeme objekty (a jejich rozpoznávání)

- vizuální (reálné, upravené, symbolické, stylizované aj.)
- akustické (zvuky, hluky, hudba, řeč,…, vždy jednorozm.)
- taktilní (obecně reprezentované *n*-rozměrnými signály)

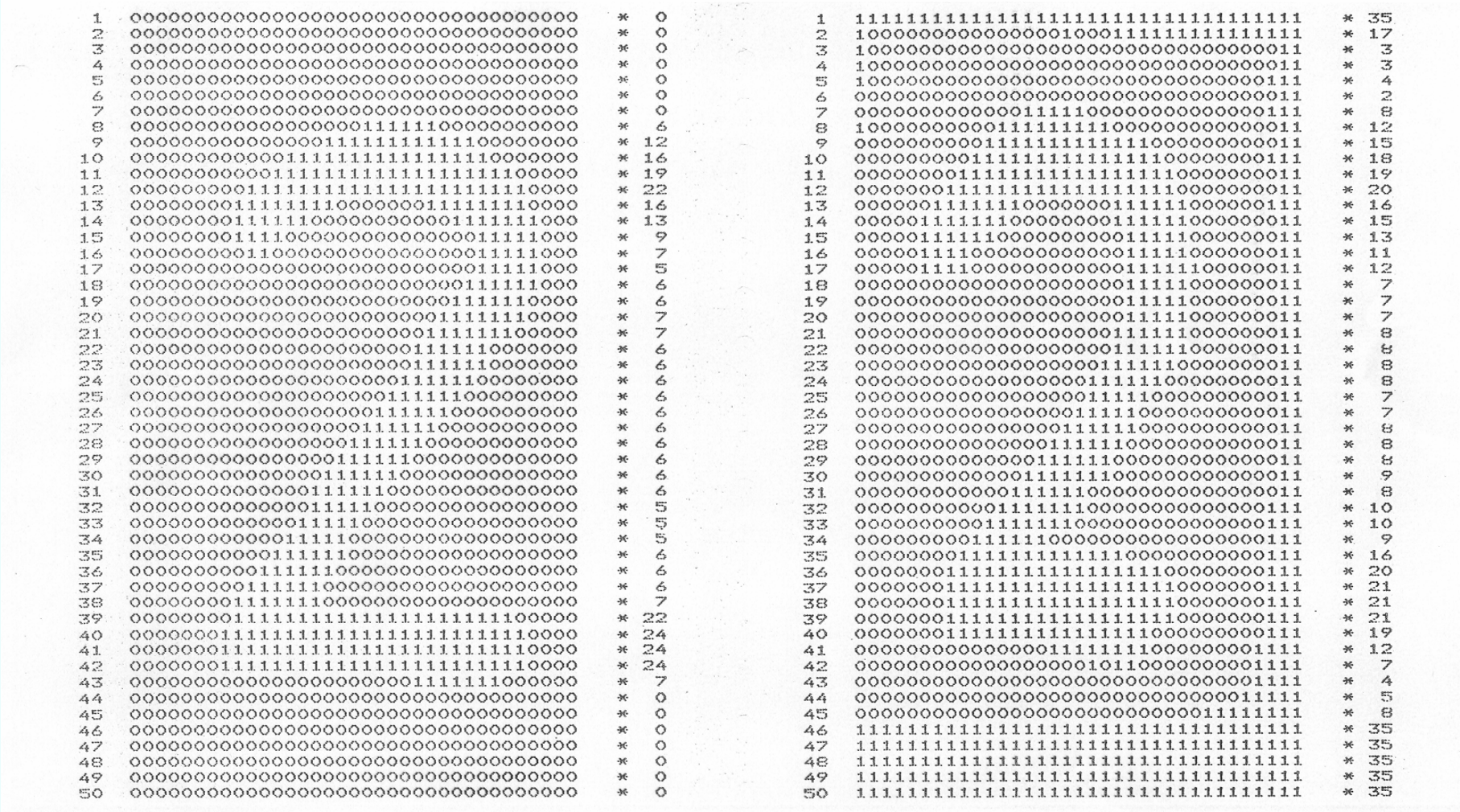
Příklad reprezentace číslice ´4´ v pravoúhlém rastru:



Reprezentace tzv. jasovou funkcí:

$$f(x,y) = \left\{ \begin{array}{l} 0\,,\ reprezentuje{-}li\ obrazový\ bod\ bílou\ barvu\ podkladu \\ 1\,,\ reprezentuje{-}li\ obrazový\ bod\ černou\ barvu\ \ písma \end{array} \right\}$$

Př.: Praktická ukázka:



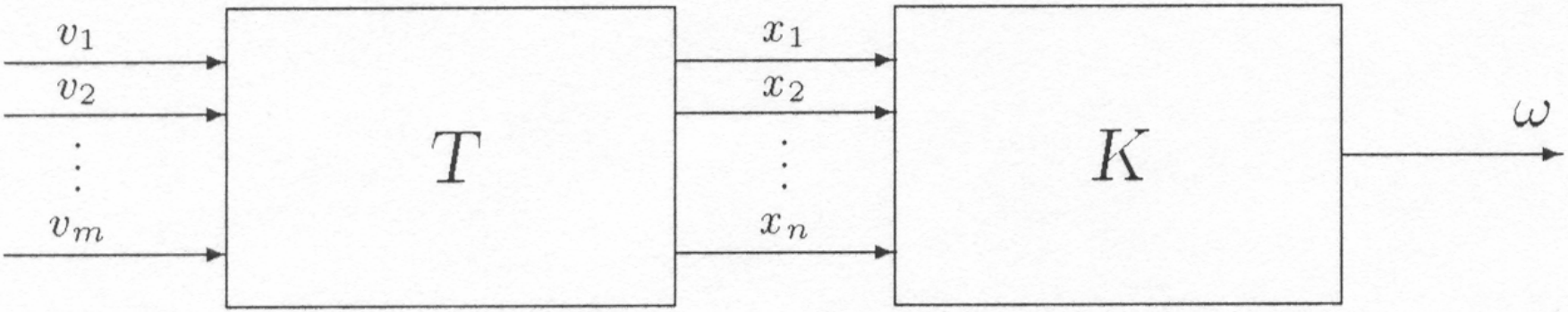
Rozpoznávání a klasifikace

Rozpoznávání chápeme jako úlohu, při které objekty zařazujeme do tříd podle jejich společných vlastností tak, že objekty vzájemně si podobné zařazujeme do stejné třídy.

Rozlišujeme:

- **klasifikaci – zařazujeme do předem známého, pevného počtu tříd (například rozpoznávání znaků)**
- **rozpoznávání – počet tříd není předem znám a třídy identifikujeme až během vlastního rozpoznávání (například rozpoznávání plynulé řeči)**

Obecná klasifikační úloha



T . . . transformace vstupních charakteristik – vytvoření obrazu
K . . . klasifikátor
$\mathbf{v}$ . . . vektor vstupních charakteristik
$\mathbf{x}$ . . . obraz (symbolický popis) objektu
ω . . . indikátor třídy

Rozhodovací pravidlo

Rozhodovací pravidlo, podle kterého klasifikátor přiřazuje obraz do klasifikační třídy, můžeme obecně definovat jako skalární funkci vektorového argumentu

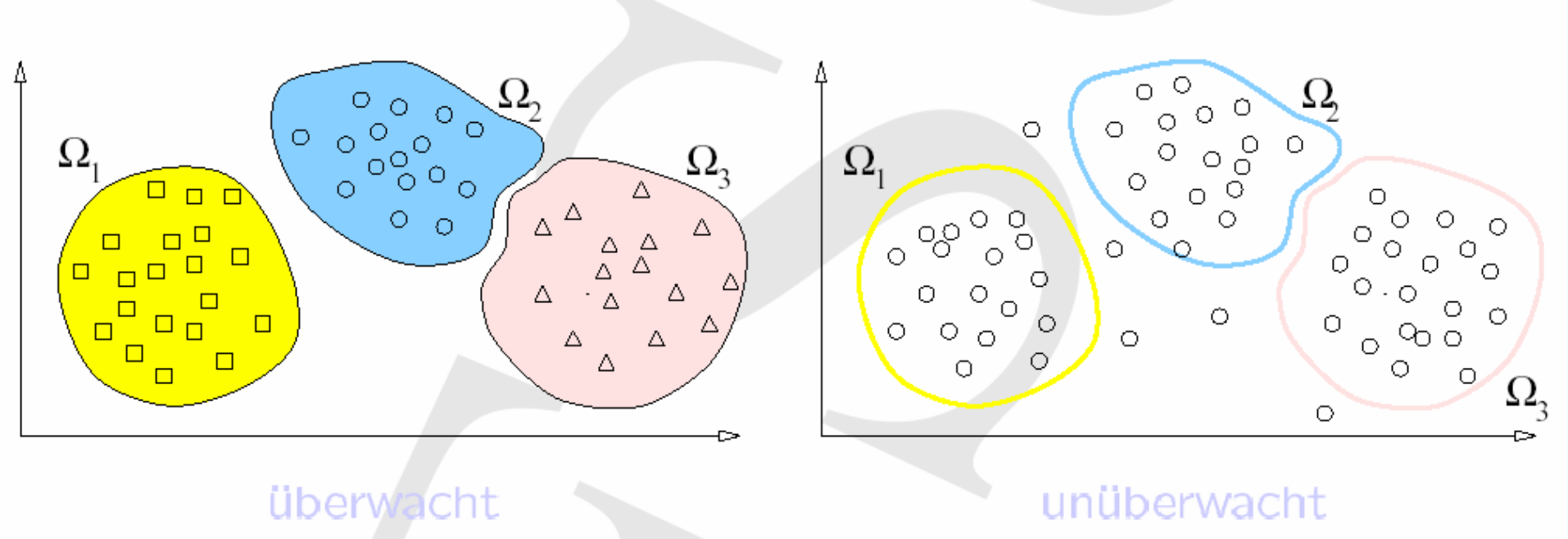
$$\omega = d(\mathbf{x}) .$$

Přesnější vyjádření rozhodovacího pravidla, které zohledňuje i tzv. nastavení klasifikátoru **q**, je

$$\omega = d(\mathbf{x}, \mathbf{q}) .$$

Nastavení klasifikátoru se provádí trénováním neboli učením. Rozlišujeme učení s učitelem, kdy klasifikátoru předkládáme obrazy, u nichž známe jejich příslušnost k třídě, a učení bez učitele, kdy správné zařazení do klasifikačních tříd neznáme.

Učení (trénování) klasifikátoru



s učitelem

bez učitele

Rozpoznávání a klasifikace objektů

- na základě vlastností objektů

Vlastnosti objektů

- měřitelné a kvantifikovatelné
- strukturní

Podle typu vlastností objektů hovoříme o

- vytváření příznakového nebo strukturního popisu objektů
- příznakových nebo strukturních metodách rozpoznávání

Příznakový nebo strukturní popis rozpoznávaných (klasifikovaných) objektů – obecně hovoříme o vytváření symbolických popisů objektů – nazýváme popisem objektů jejich **obrazy**.

Tedy: **Obraz objektu = symbolický popis objektu**

Stejně jako metody rozlišujeme i obrazy (symbolické popisy objektů)

- tvořené příznaky (vektory příznaků)
- strukturní

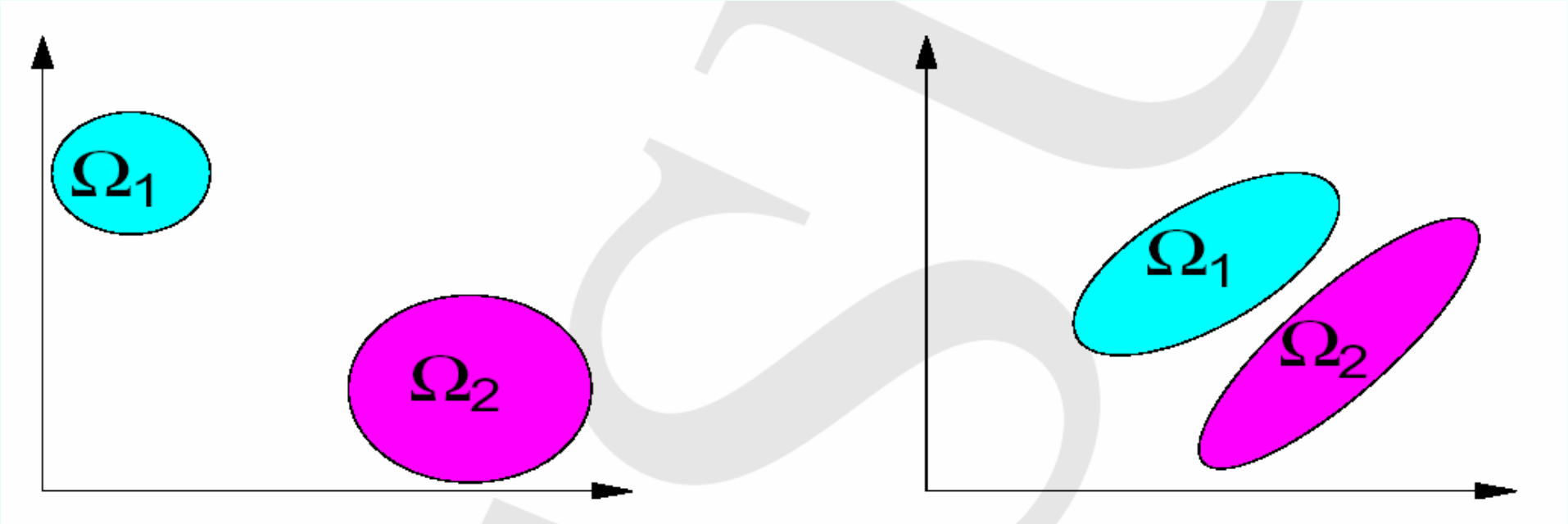
Obsahují-li obrazy objektů *n* složek, pak všechny rozpoznávané (klasifikované) obrazy tvoří obrazový prostor úlohy.

Jednotlivé klasifikační třídy objektů získáme rozkladem obrazového prostoru úlohy na *R* klasifikačních tříd.

Příklady rozkladu obrazového prostoru na třídy

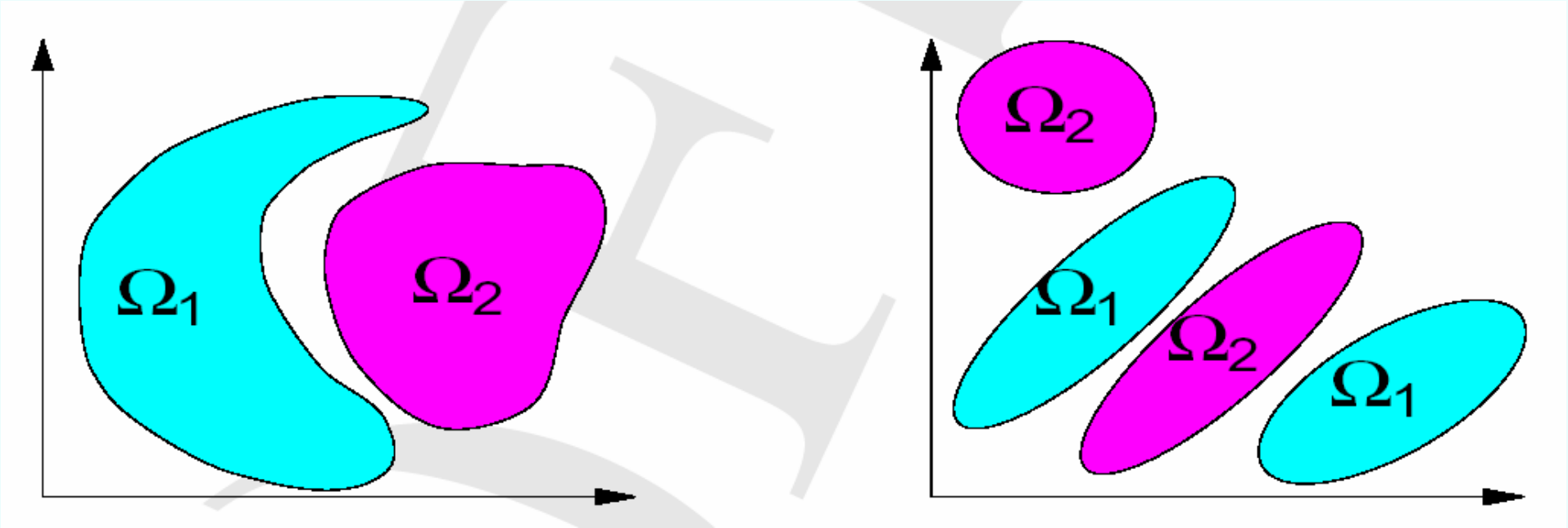
a) třídy disjunktní, s velkou mezi-třídní vzdáleností – ideální případ

b) třídy disjunktní, kompaktní, geometricky „blízké“

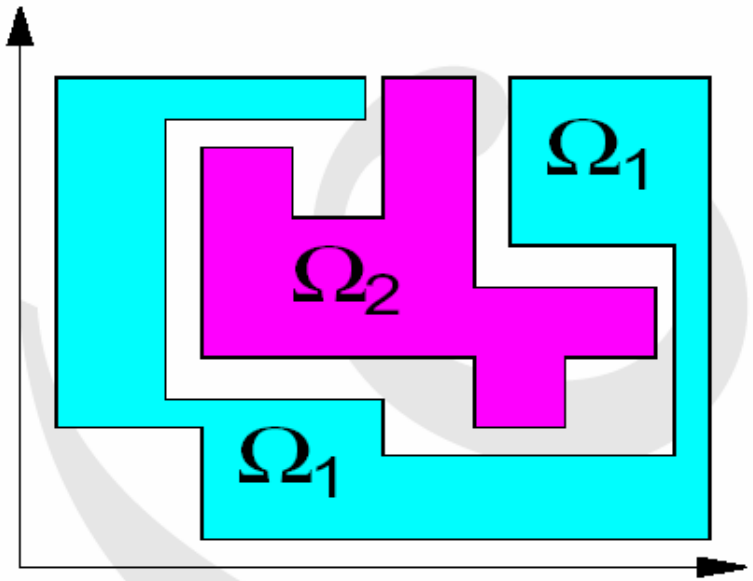


c) třídy disjunktní, kompaktní, s nelineární oddělující nadplochou

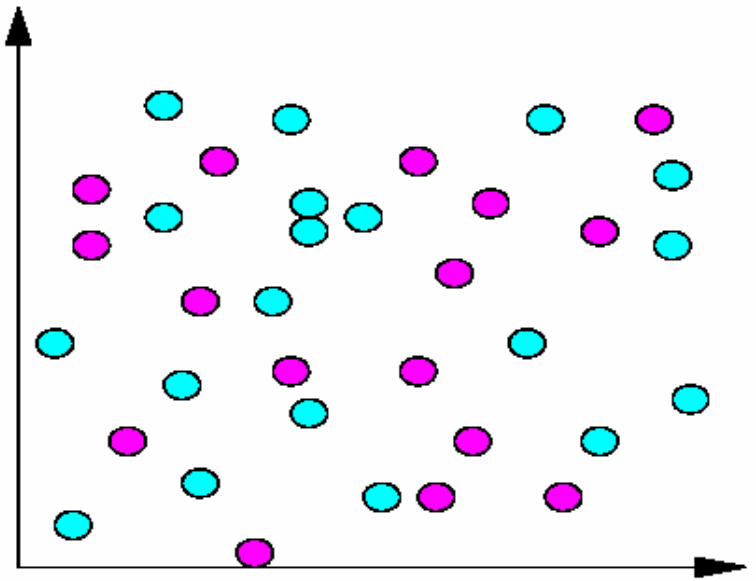
d) třídy disjunktní, nekompaktní, avšak lineárně oddělitelné



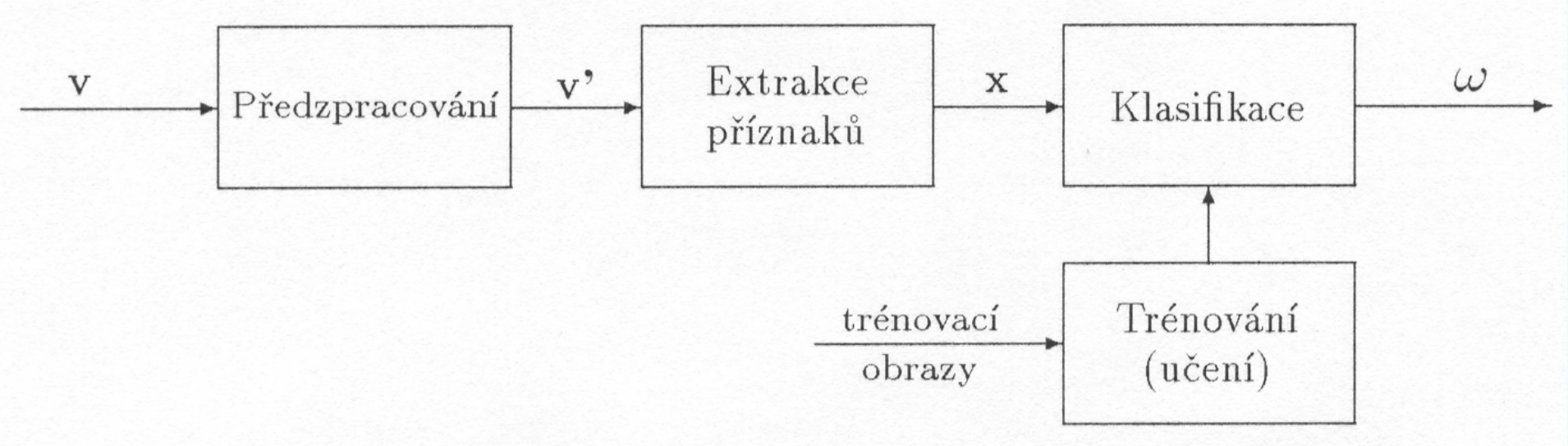
e) třídy disjunktní, kompaktní, vnořené, komplikovaně oddělitelné (s rozdělující nadplochou složitého tvaru)



f) nekompaktní, prolínající se třídy s velmi obtížně určitelnou rozdělující nadplochou



Příznakové metody rozpoznávání



Obrazy objektů reprezentovány vektory příznaků x , zařazování obrazů do tříd – klasifikace deterministickým nebo stochastickým rozhodovacím pravidlem ve tvaru

$$\omega = d(x) \text{ , resp. } \omega = d(x, q) \text{ ,}$$

kde x je klasifikovaný obraz objektu,
q je vektor nastavení klasifikátoru.

Jako obraz (vektor příznaků) lze použít např.:

1. naměřené veličiny přímo,
2. spektrální koeficienty,
3. logaritmus kvadrátu spektra posloupnosti vzorků, např. $10 * \log |\mathbf{V}|^2$, kde **V** je Four. obraz vektoru **v**,
4. histogram jednoho řádku snímku,
5. úhlové projekce (např. po 45 stupních) průsečíků průvodiče se sejmutým znakem,
6. řetězový kód středové osy podpisu,
7. ortogonální nebo biortogonální transformace signálů aj.

Klasifikátor s diskriminační funkcí

Pro každou třídu definujeme takovou diskriminační funkci, aby pro všechny obrazy patřící do *r*-té klasifikační třídy platilo

$$g_r(x) > g_s(x) \quad | \quad r \in \langle 1, R \rangle , \; s = 1,\ldots, R, \; r \neq s .$$

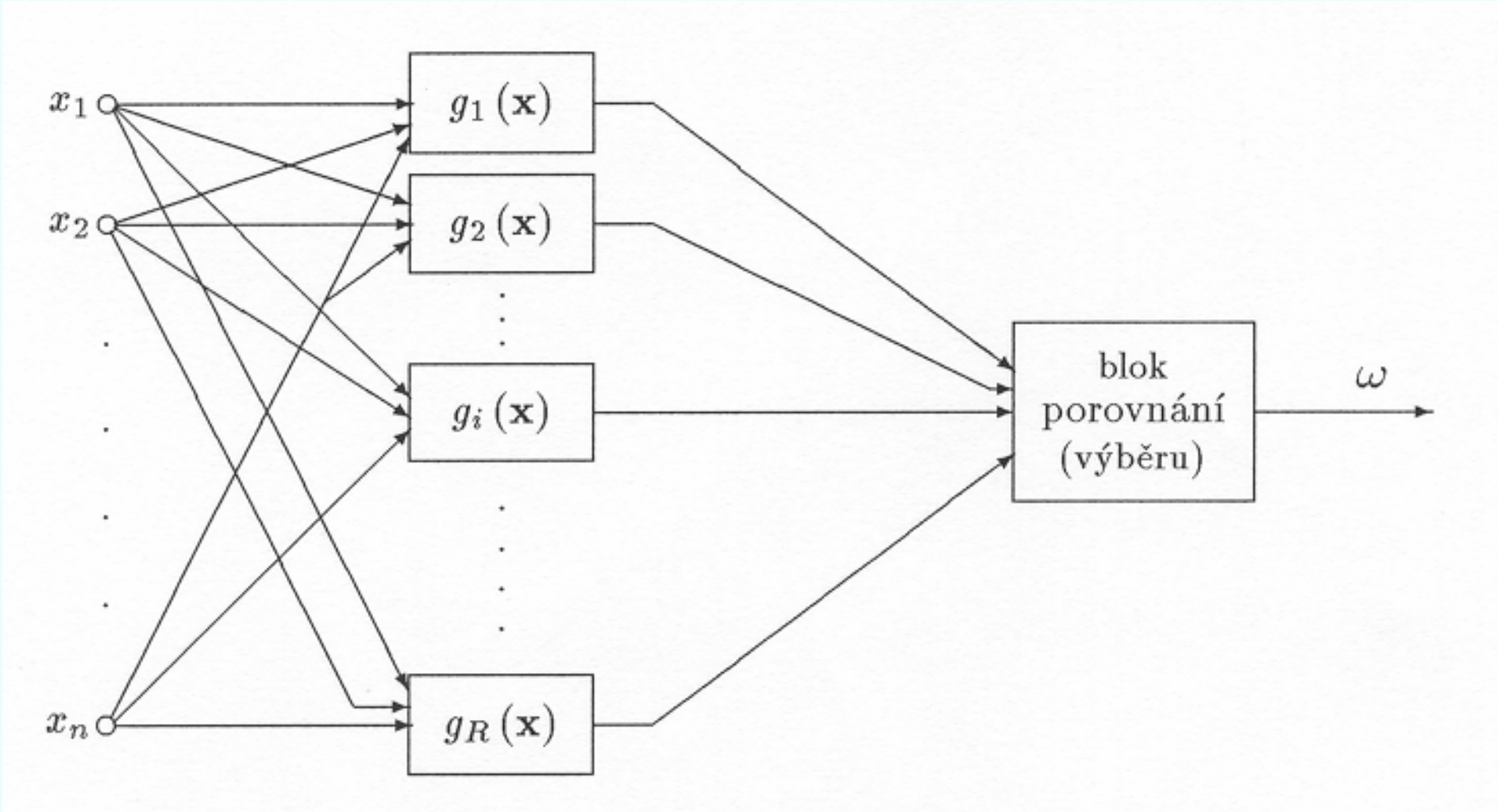
Rozhodovací pravidlo pak nabývá tvaru

$$\omega_r = max\,(g_s(x)) , \quad s = 1,\ldots, R$$

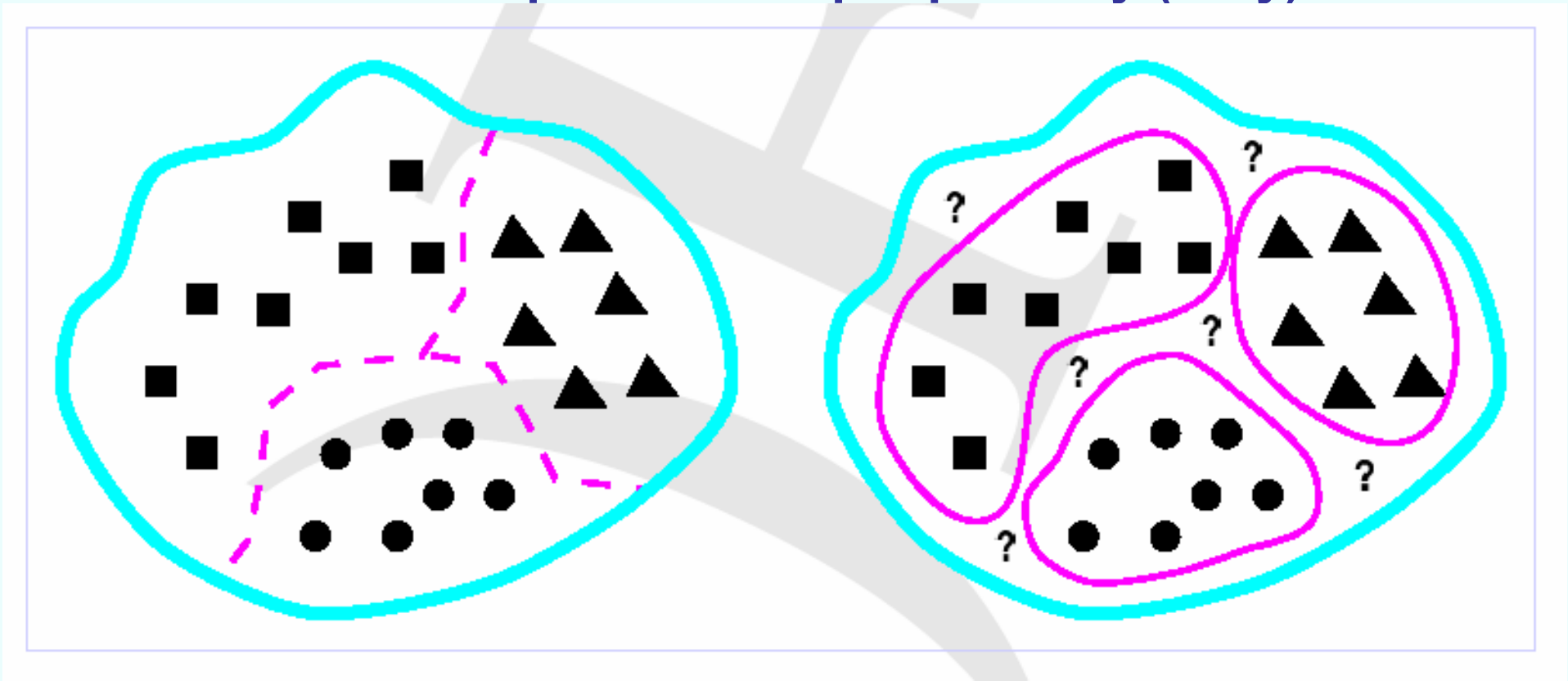
a rovnice rozdělujících nadploch určíme řešením soustavy rovnic

$$g_r(x) - g_s(x) \quad | \quad r , s = 1,\ldots, R, \; r \neq s .$$

Struktura klasifikátoru

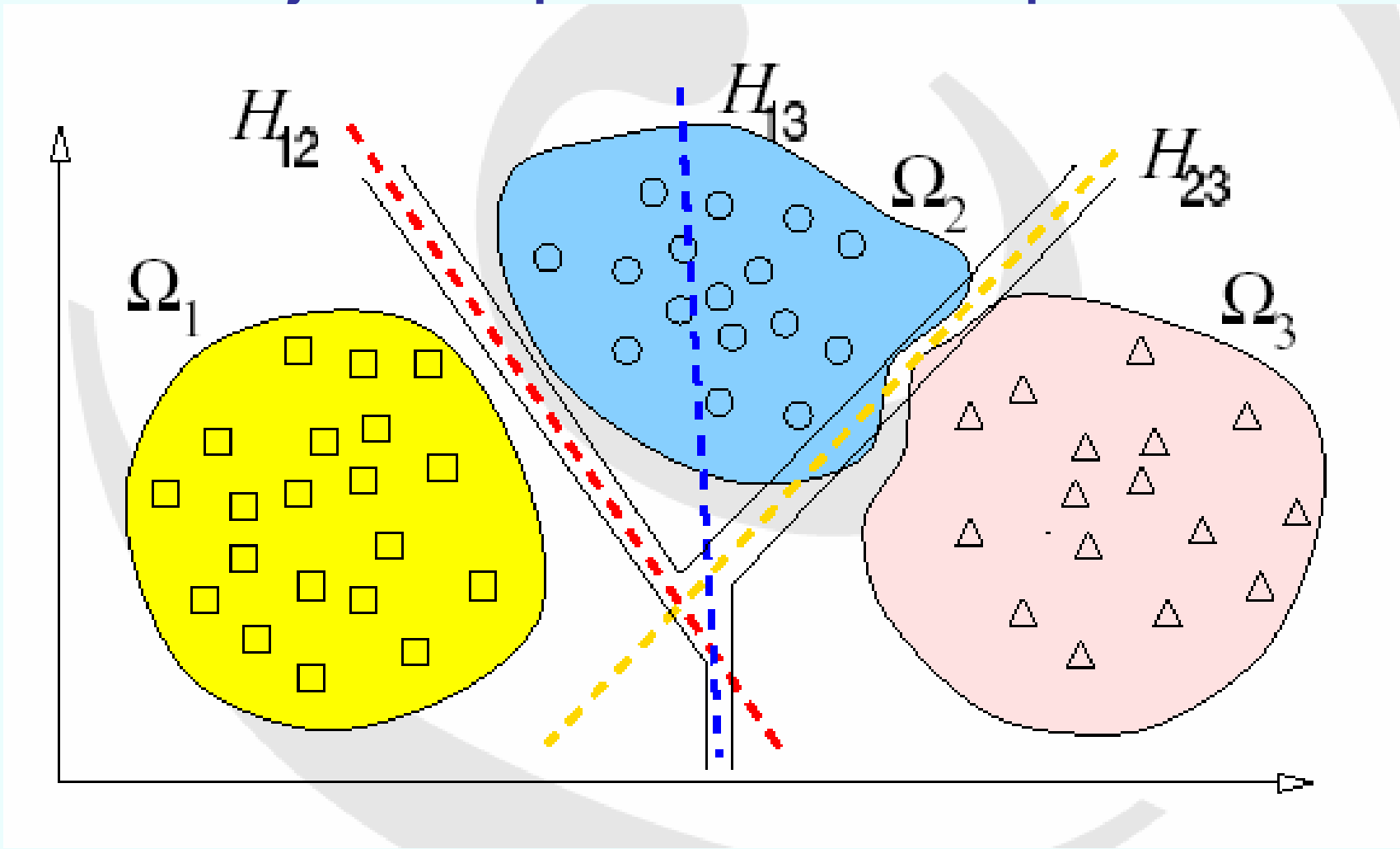


Rozklad obrazového prostoru na podprostory (třídy)



a) s rozdělujícími nadplochami **b)** s obalovými třídami

Určení rozdělujících nadploch v obrazovém prostoru



Klasifikátor na principu kritéria minimální vzdálenosti

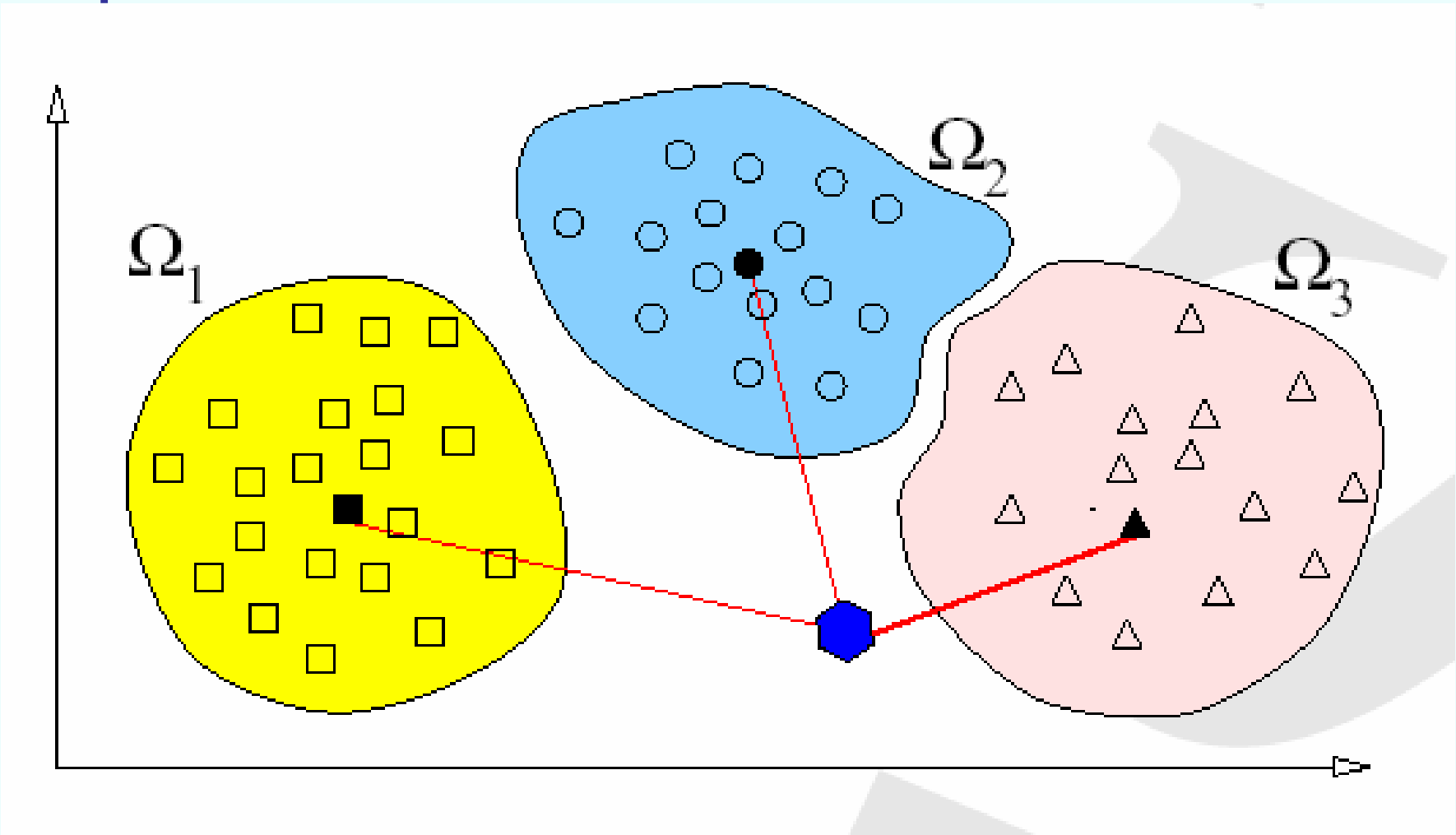
(porovnávání klasifikovaných obrazů se vzorovými obrazy – tzv. etalony)

Lze použít v úlohách s oddělitelnými množinami obrazů (třídami) – ve fázi trénování se vytvoří vzorové obrazy (etalony) jednotlivých tříd – označme je $\mathbf{e}_s$, $s = 1,\ldots, R$.

Neznámé obrazy pak klasifikujeme pravidlem

$$\omega_r = \| \mathbf{e}_r - \mathbf{x} \| = \min \| \mathbf{e}_s - \mathbf{x} \|, \quad s = 1,\ldots, R$$

Princip:



Kritérium minimální chyby (Bayesův klasifikátor)

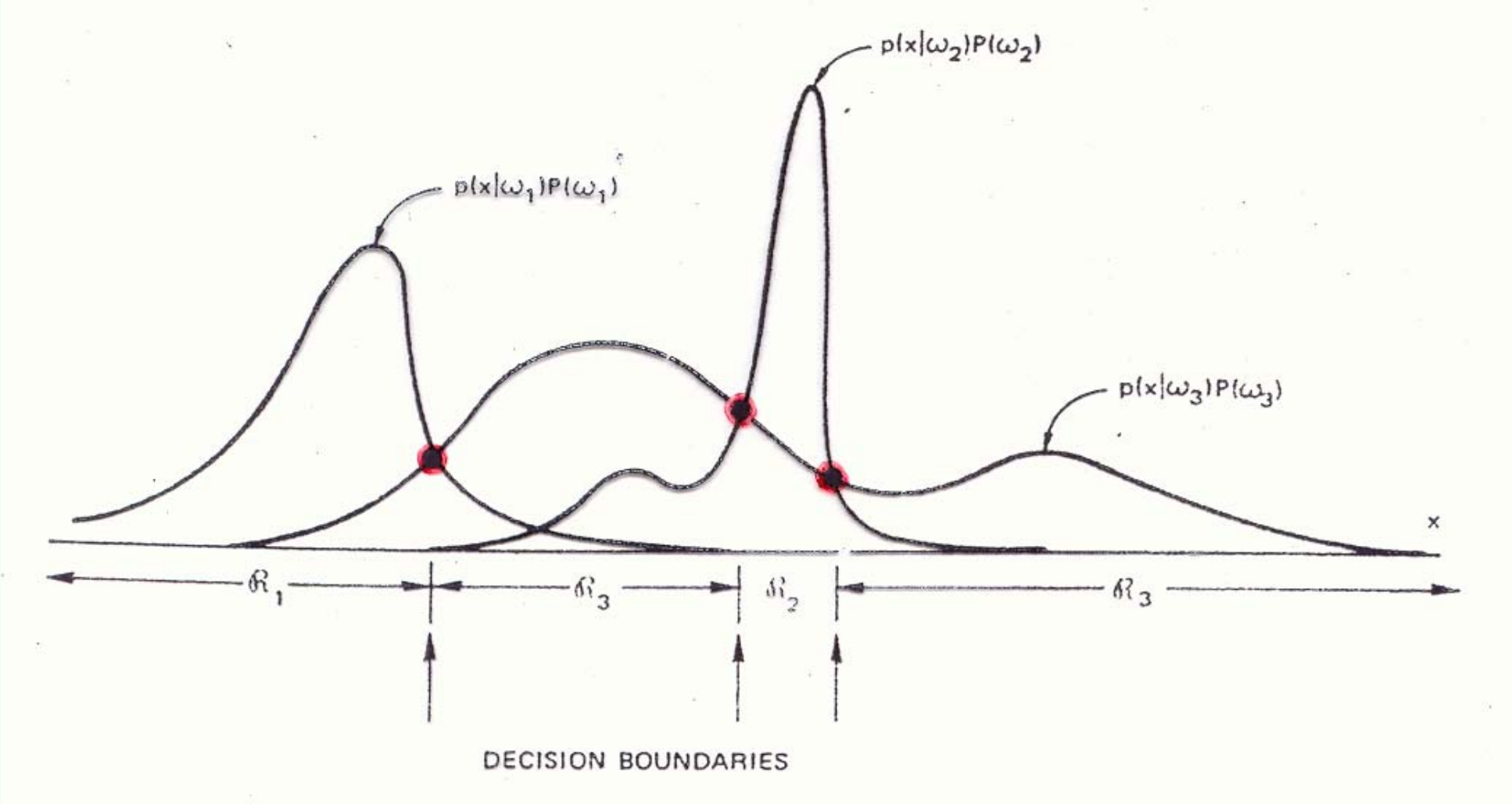
Pro úlohy s neoddělitelnými (prolínajícími se) třídami obrazů – příslušnost k třídě lze určit jen s určitou pravděpodobností:

Kritérium minimální chyby využívá k popisu úlohy pravděpodobnostního zápisu. Potom hodnotu ω pokládáme za náhodnou proměnnou. Pokud bychom znali pouze hodnoty apriorních pravděpodobností $P(\omega_i)$, zařazovali bychom všechny obrazy do té třídy ω_r, která má nejvyšší apriorní pravděpodobnost $P(\omega_r)$. To by však bylo chybné. Pravděpodobnost této chyby můžeme vyjádřit jako $1 - P(\omega_r)$.

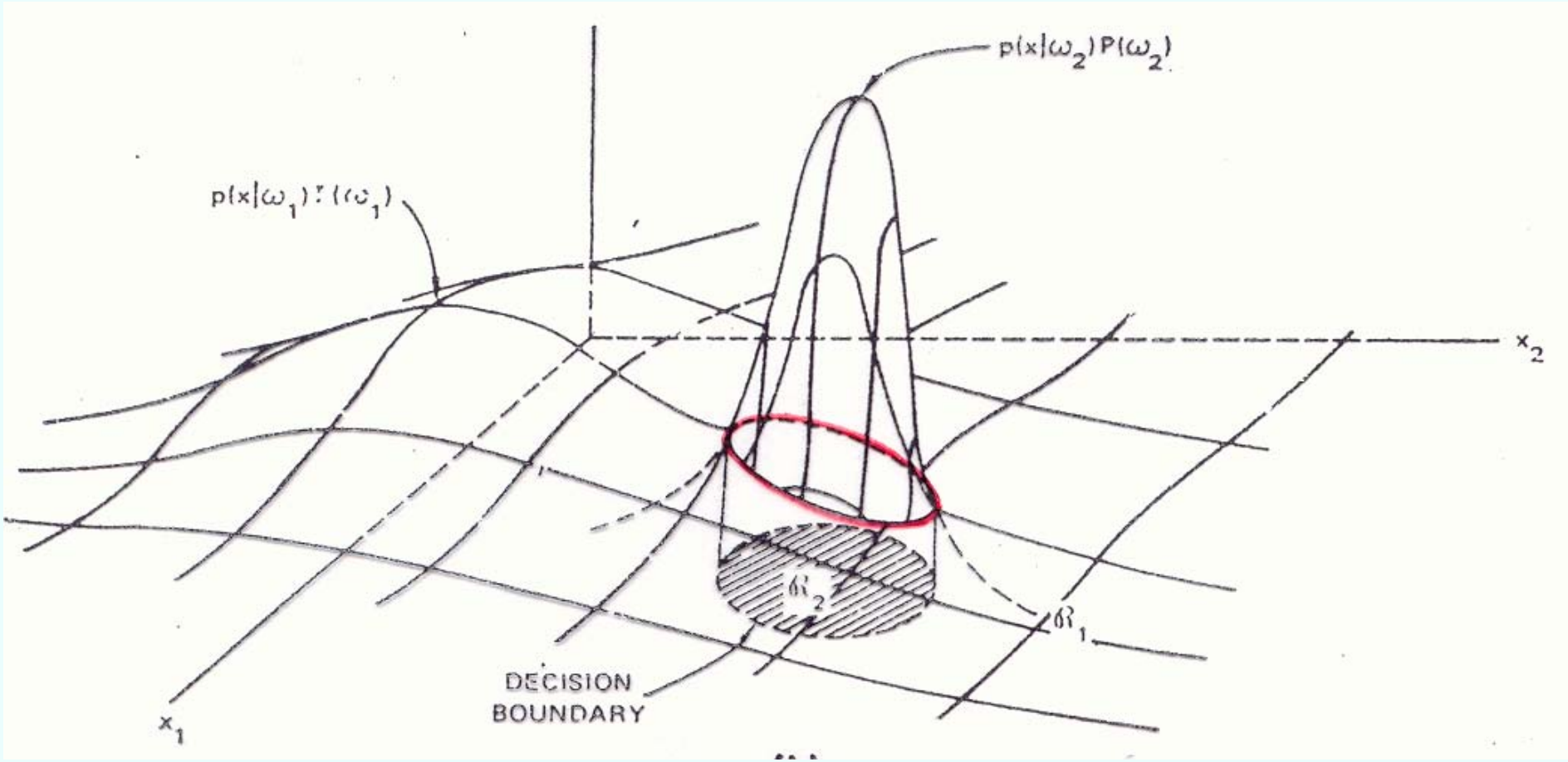
Většinou se nemusíme rozhodovat jen podle apriorních pravděpodobností tříd $P(\omega_i)$, ale máme k dispozici i hodnotu vektoru příznaků $\boldsymbol{x}$ a všechny podmíněné hustoty rozdělení pravděpodobnosti $p(\boldsymbol{x} \mid \omega_r)$.

Tyto podmíněné hustoty pravděpodobnosti vyjadřují rozložení hodnot $\boldsymbol{x}$ uvnitř jednotlivých tříd; získáme je trénováním klasifikátoru. Histogram četností výskytu jednotlivých obrazů uvnitř jedné třídy je dobrým odhadem hustoty pravděpodobnosti.

Příklad – pro jednorozměrný případ (vektor příznaků o jednom prvku):



Dvourozměrný případ:



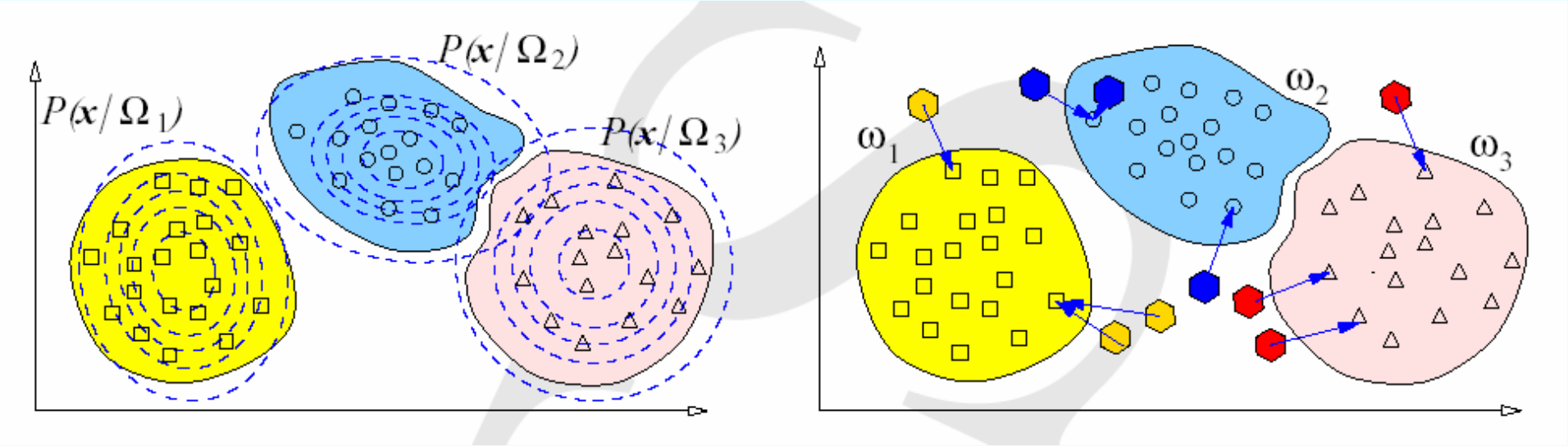
Pravděpodobnost (tzv. aposteriorní pravděpodobnost), že obraz ***x*** přísluší do třídy s identifikátorem ω_r , je dána vztahem

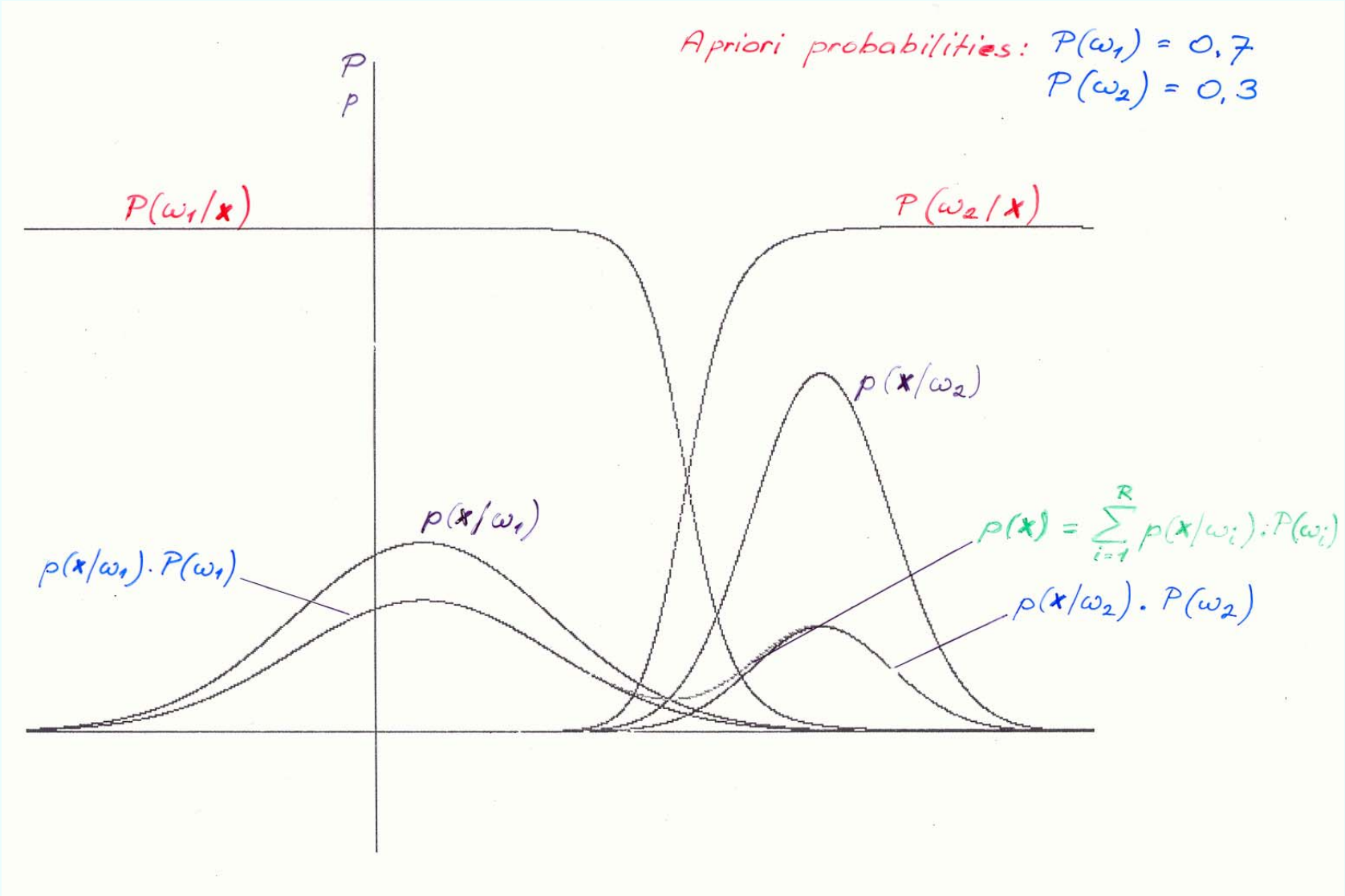
$$P(\omega_s/x) = \frac{p(x/\omega_s).P(\omega_s)}{p(x)} \quad , \quad p(x) = \sum_{i=1}^{R} p(x/\omega_i).P(\omega_i) \ .$$

Rozhodovací pravidlo pro zařazení objektu do *r*-té třídy pak má tvar:

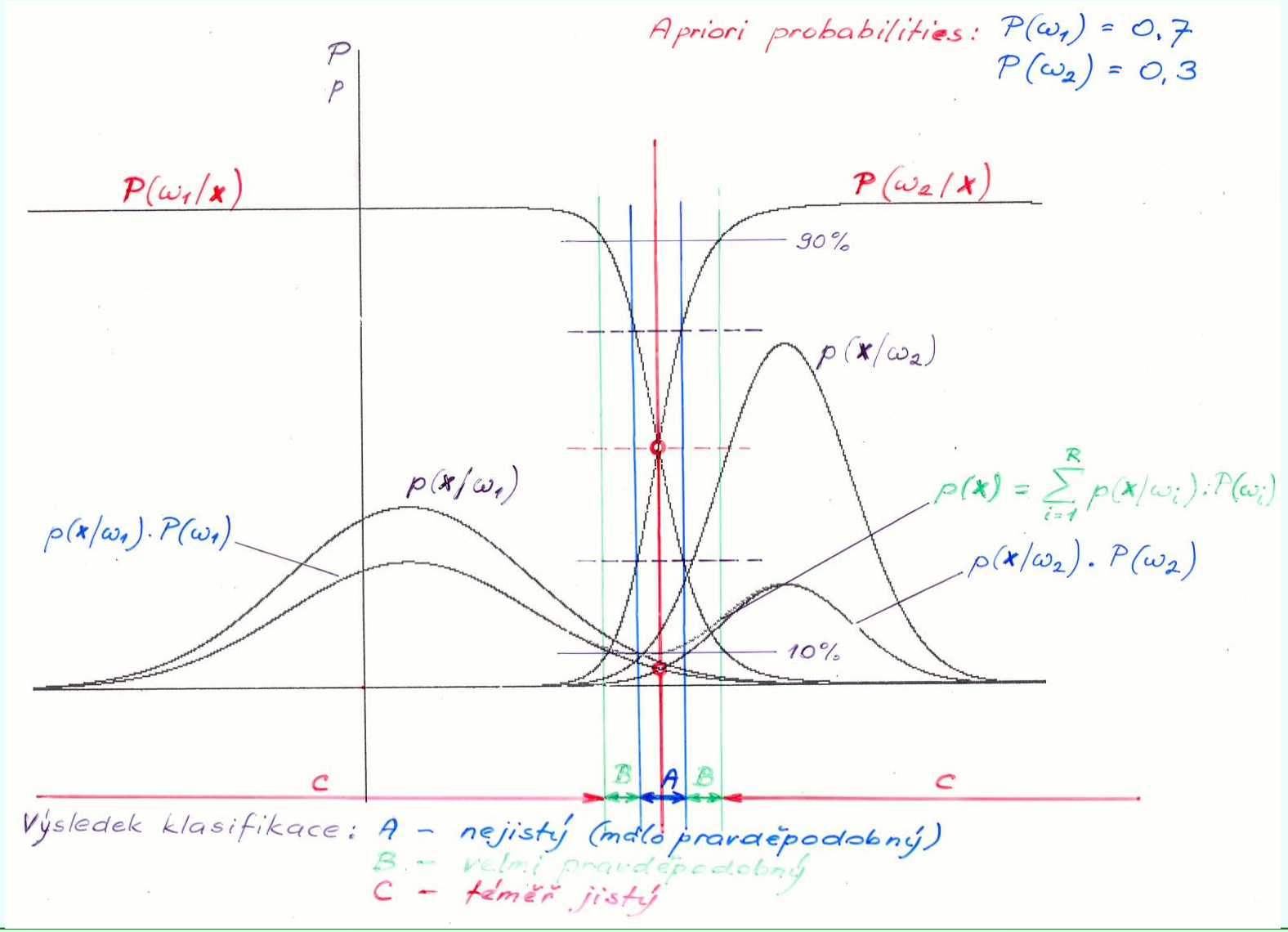
$$\omega_r = \max_{\forall s} \left(P(\omega_s/x)\right) , \ \ s = 1,...,R \ .$$

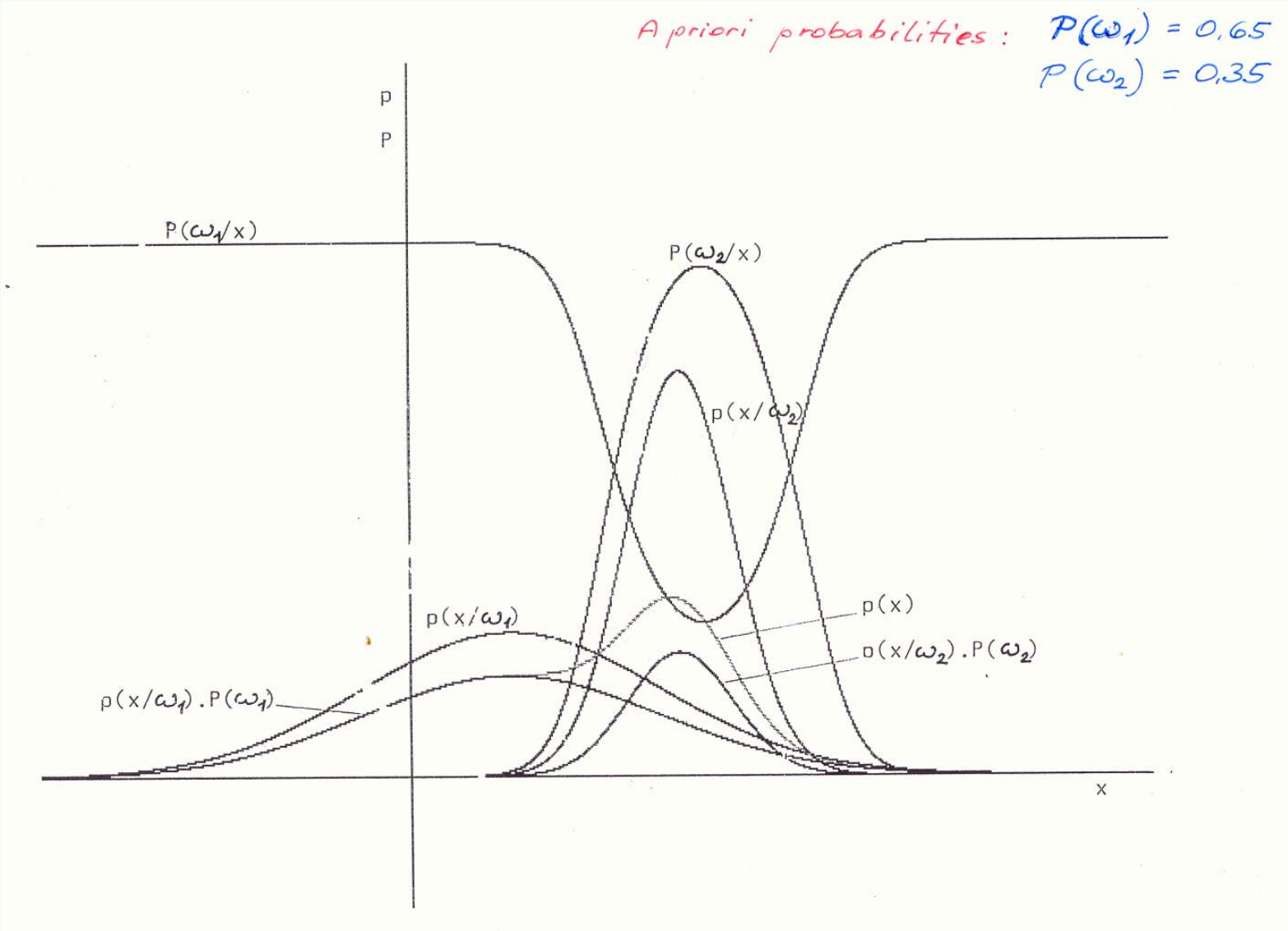
Ilustrace:




Apriori probabilities: $P(\omega_1) = 0,7$
$P(\omega_2) = 0,3$
P
p
$P(\omega_1/x)$
$P(\omega_2/x)$
$p(x/\omega_2)$
$p(x/\omega_1)$
$p(x/\omega_1) \cdot P(\omega_1)$
$p(x) = \sum_{i=1}^{R} p(x/\omega_i) \cdot P(\omega_i)$
$p(x/\omega_2) \cdot P(\omega_2)$

4. Rozpoznávání a klasifikace




A priori probabilities: P(ω1) = 0,65
P(ω2) = 0,35
p
P
P(ω1/x)
P(ω2/x)
p(x/ω2)
p(x/ω1)
p(x)
p(x/ω2).P(ω2)
p(x/ω1).P(ω1)
x

Kritérium minimální chyby

The decision rule

$$\omega = d(\mathbf{x}, \mathbf{q}^*)$$

that gives the minimum mean loss $J(\mathbf{q})$ is called the optimum decision rule, and $\mathbf{q}^*$ is called the vector of optimal parameters

$$J(\mathbf{q}^*) = \min_{\mathbf{q}} J(\mathbf{q}), \qquad d(\mathbf{x}, \mathbf{q}) \in D$$

The **minimum error criterion** (Bayes criterion, maximum likelihood) uses loss functions of the form $\lambda(\omega_r|\omega_s)$, where $\lambda(.)$ is the number that describes quantitatively the loss incurred if a pattern $\mathbf{x}$ which should be classified into the class ω_s is incorrectly classified into the class ω_r

$$\omega_r = d(\mathbf{x}, \mathbf{q})$$

The mean loss is

$$J(\mathbf{q}) = \int_X \sum_{s=1}^{R} \lambda(d(\mathbf{x}, \mathbf{q})|\omega_s) p(\mathbf{x}|\omega_s) P(\omega_s) d\mathbf{x}$$

where $P(\omega_s)$, $s = 1, \ldots, R$ are the a priori probabilities of classes, and $p(\mathbf{x}|\omega_s)$, $s = 1, \ldots, R$ are the conditional probability densities of objects $\mathbf{x}$ in the class ω_s.

$$\begin{aligned} \lambda(\omega_r|\omega_s) &= 0 \text{ for } r = s \\ &= 1 \text{ for } r \neq s \end{aligned}$$

and the discrimination functions are

$$g_r(\mathbf{x}) = p(\mathbf{x}|\omega_r) P(\omega_r), \qquad r = 1, \ldots, R$$

where $g_r(\mathbf{x})$ corresponds (up to a multiplicative constant) to the value of the a posteriori probability $P(\omega_r|\mathbf{x})$.

Výpočet aposteriorních pravděpodobností

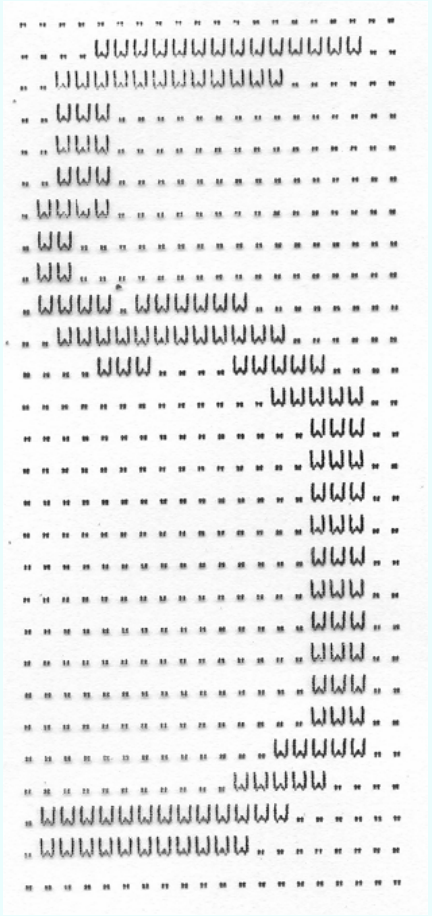
$$P(\omega_r|\mathbf{x}) = \max_{s=1,\ldots,R} P(\omega_s|\mathbf{x})$$

A posteriori probability may be computed from a priori probabilities using the Bayes formula

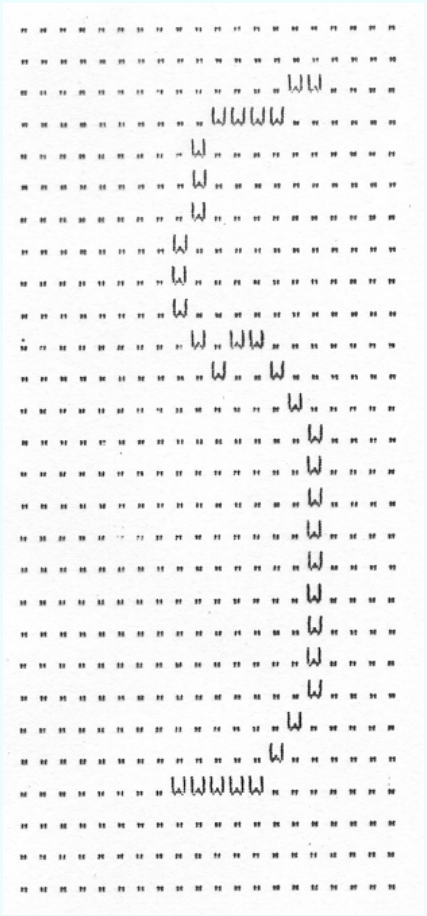
$$P(\omega_s|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_s)P(\omega_s)}{p(\mathbf{x})}$$

where $p(\mathbf{x})$ is the mixture density.

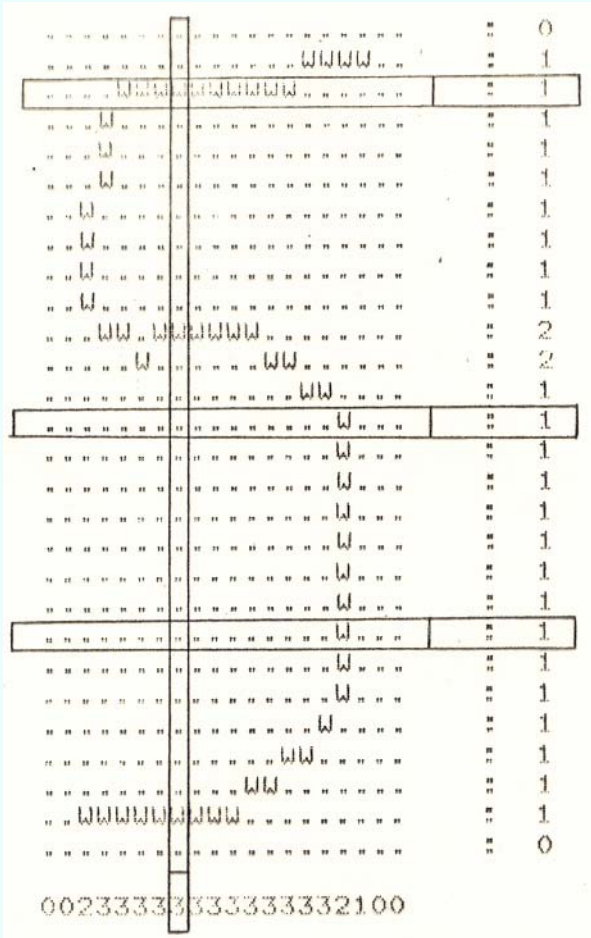
Příklad příznakového rozpoznávání číslic:



sejmutý znak

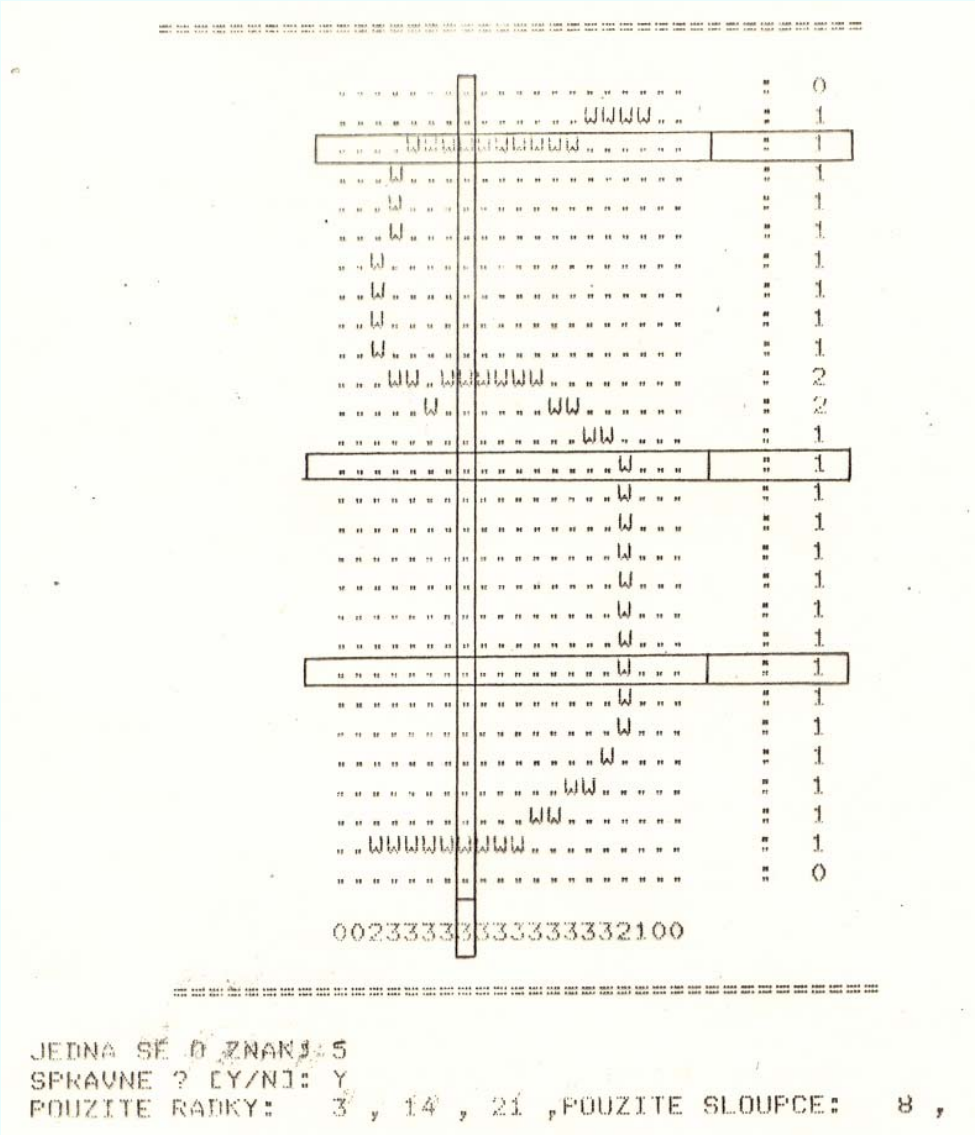


skelet znaku



extrakce příznaků

Výsledek rozpoznávání:



Strukturní metody rozpoznávání

Strukturní popisy rozpoznávaných objektů

- primitiva
- vlastnosti primitiv
- relace mezi primitivy
  - prostorové
  - časové
  - funkční

Vytvořený symbolický popis – **obraz** vystihuje (popisuje) **strukturní vlastnosti objektu .**

Jako strukturní popisy (obrazy) lze použít:

- řetězec symbolů označujících primitiva
- relační struktura
- graf (obecný, speciální, …)

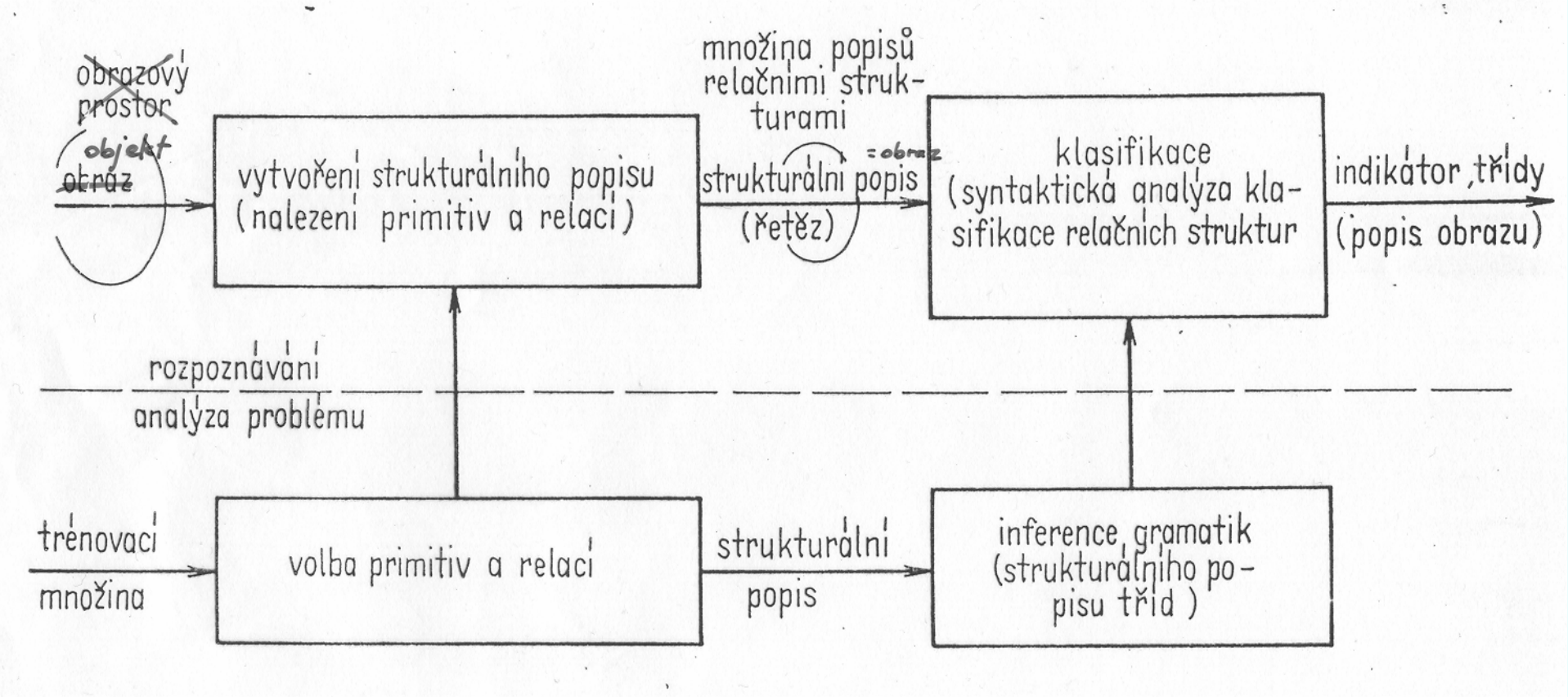
Strukturní popisy objektů (= slova) příslušejících do téže třídy tvoří **jazyk** třídy.

Rozpoznávání strukturně popsaných objektů = rozhodnutí, zda vytvořený strukturní popis objektu (slovo) **je slovem** (frází) **jazyka příslušné třídy**.

Výhody strukturních metod rozpoznávání:

- invariantní na pozici a natočení obrazu,
- méně složité popisy u složitých objektů

Úloha strukturního rozpoznávání



Slovní formulace:

Algoritmus 7.3 - Syntaktické rozpoznávání

Fáze učení

(1) Na základě analýzy úlohy definujte primitiva a jejich vzájemné relace.

(2) Analýzou syntaktických popisů předmětů jednotlivých tříd případně odvozováním (inferencí) (odstavec 7.2.3) sestrojte gramatiku reprezentující každou třídu.

Fáze rozpoznávání

(3) Při rozpoznávání, do které třídy předmět náleží, extrahujte nejprve jeho primitiva, určete jejich druh a vzájemné relace. Sestrojte slovo reprezentující popisovaný předmět.

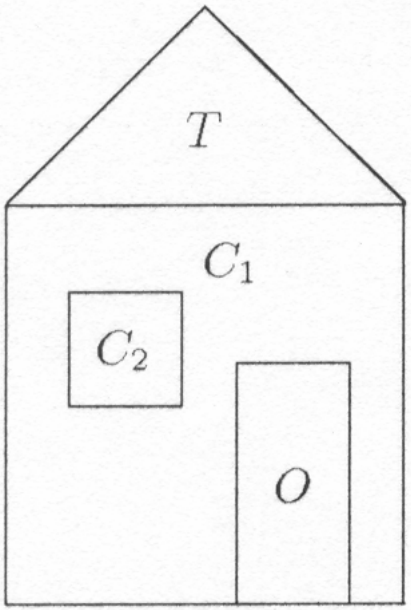
(4) Na základě výsledku syntaktické analýzy zařaďte předmět do té třídy, jejíž gramatika (sestrojená v kroku (2)) dané slovo generuje.

Jak postupovat při vytváření strukturních popisů:

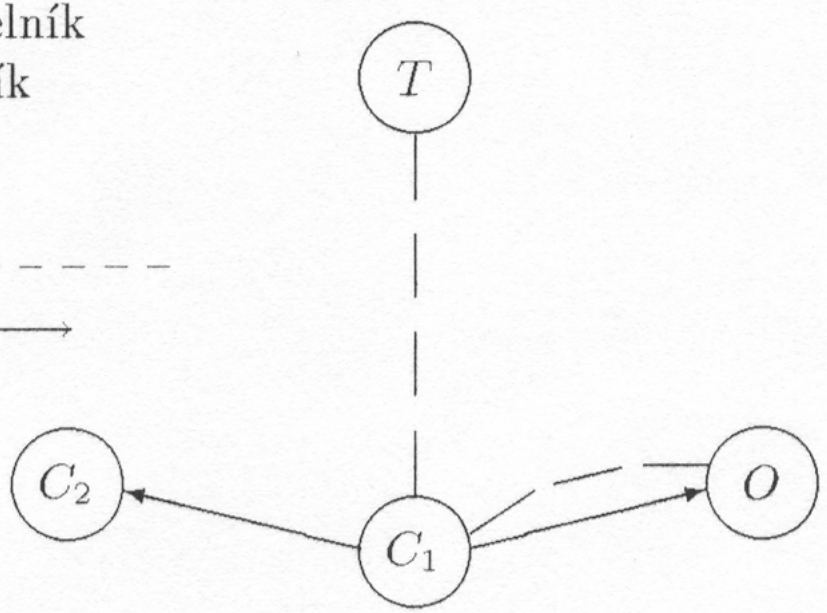
1. nalézt všechna **primitiva** a přiřadit jim prvky nosiče struktury;
2. každému prvku struktury přiřadit **vlastnost** (unární relaci) označenou jménem (symbolem) odpovídajícího primitiva;
3. určit **vztahy mezi primitivy** (binární relace), čímž vytvoříme **relační strukturu**;
4. doplnit případnou informaci číselné povahy (vytvoříme **sémantickou informaci**, popř. sémantický vektor).

Vytváření popisné relační struktury (strukturního obrazu)

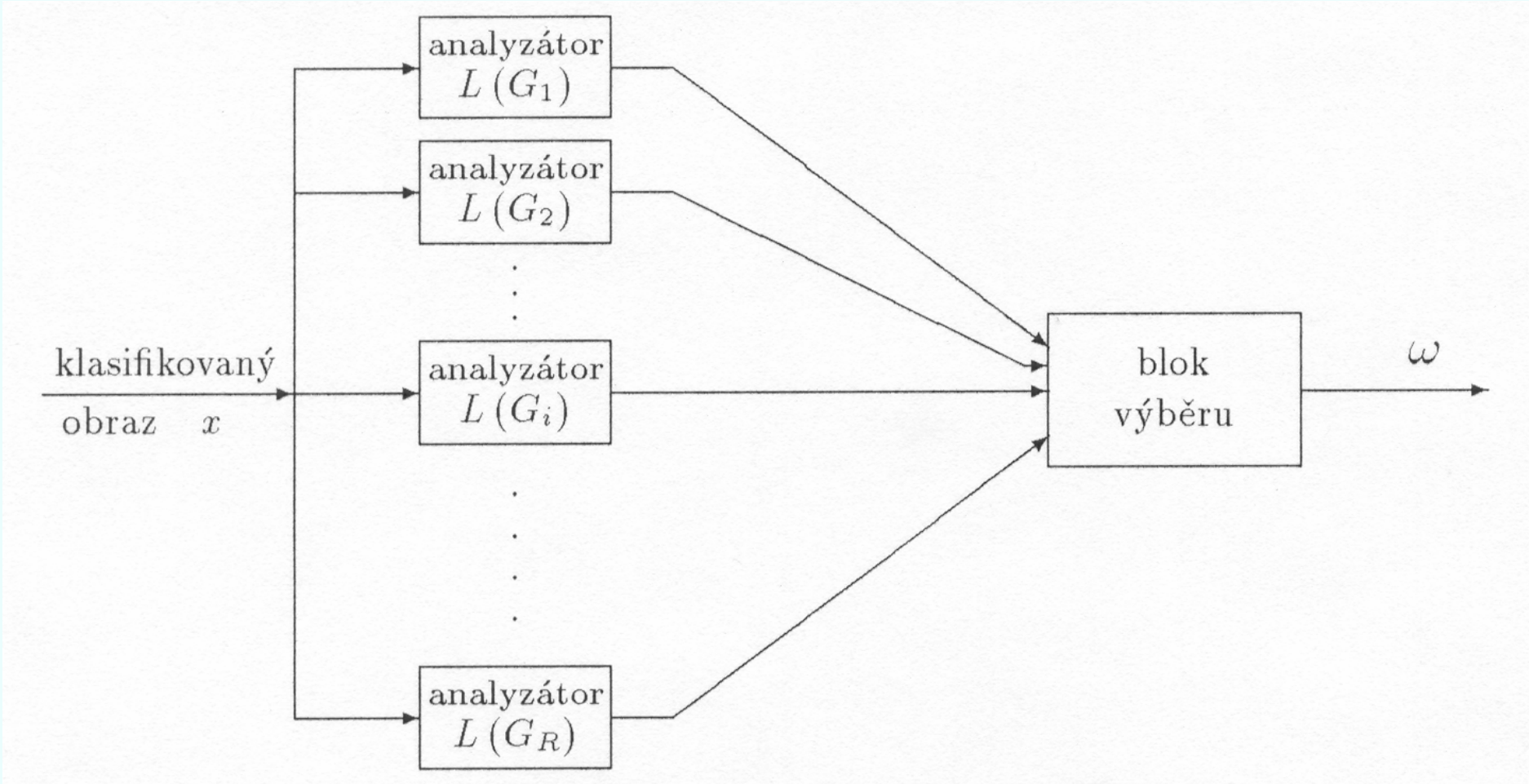
$\mathcal{M} \equiv \{ T, O, C_1, C_2 \}$



unární relace:
T ... je trojúhelník
O ... je obdélník
C ... je čtverec
binární relace:
dotýká se ... - - - - -
je uvnitř ... ⟶



Struktura klasifikátoru



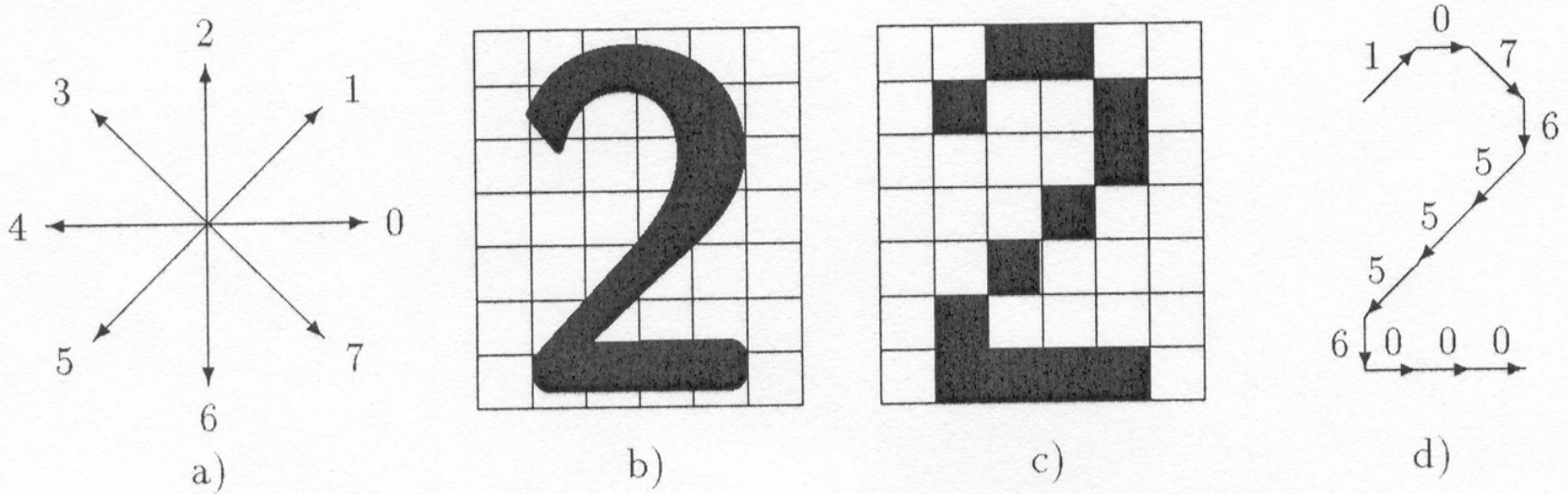
Extrakce primitiv, vytváření strukturních obrazů

1. **Počet** typů (druhů) primitiv i relací (vztahů) mezi nimi by měl být **co nejmenší .**

2. **Primitiva** by měla **odpovídat** základním (přirozeným) **strukturním elementům** objektu, jimiž lze objekt vyčerpávajícím způsobem popsat; přitom primitiva musejí být snadno **extrahovatelná** a **klasifikovatelná** (nejčastěji některou příznakovou metodou).

3. **Nalezení** (určení) **primitiv** a **relací** mezi nimi (způsob nalezení, algoritmus, …) by mělo být **algoritmicky co nejjednodušší.**

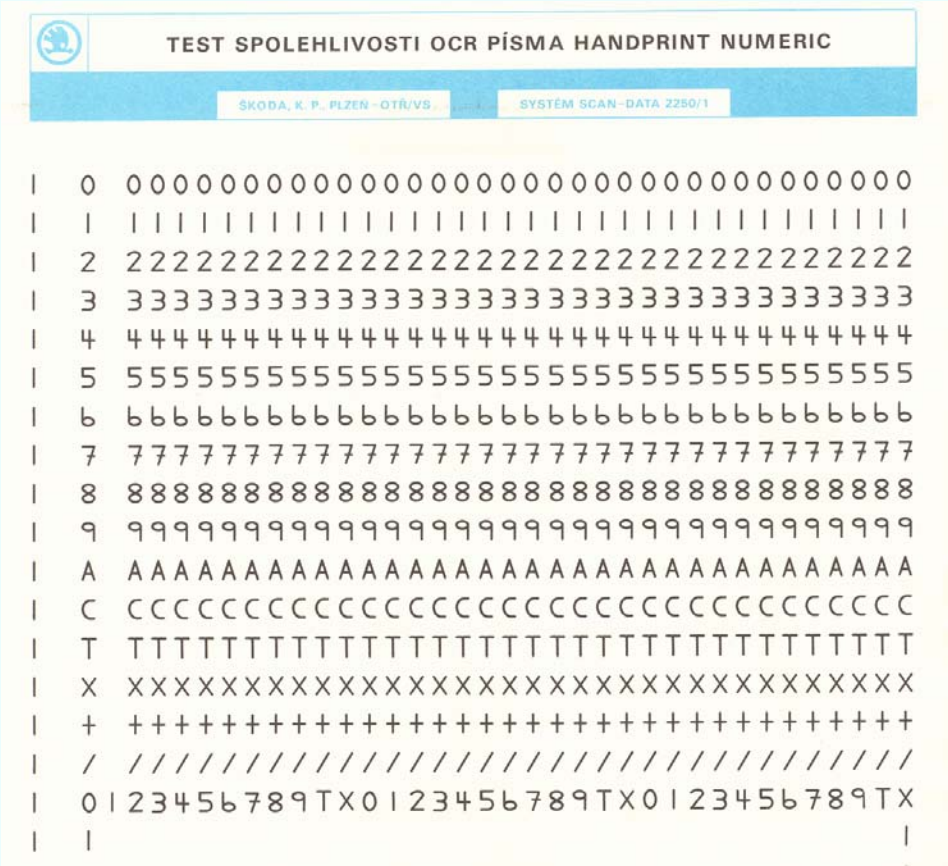
Příklad extrakce primitiv Freemanovým kódem



Výsledná reprezentace tvaru číslice **2**: 10765556000

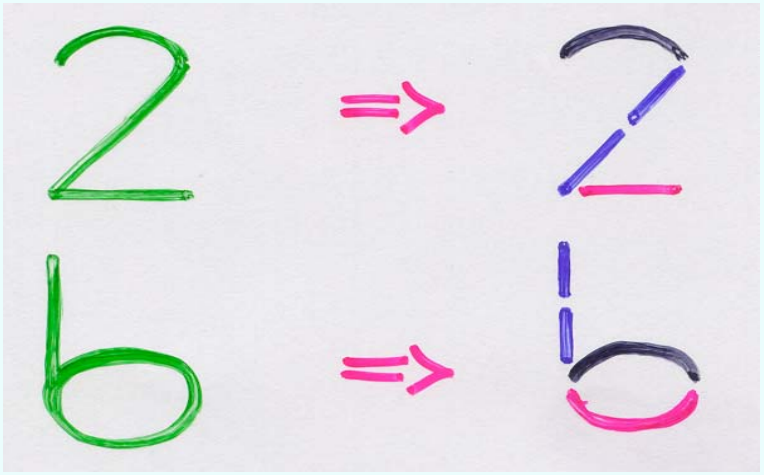
Obr. 5.7: Příklad popisu číslice **2** Freemanovým řetězovým kódem
a) směrová růžice Freemanova kódu
b) "překrytí" kódovaného znaku kódovacím rastrem
c) přibližná reprezentace znaku v kódovacím rastru
d) posloupnost směrů v kódovacím rastru reprezentující znak **2**

Př.: Strukturní rozpoznávání (klasifikace !) ručně psaných číslic





Vytvoření strukturního popisu



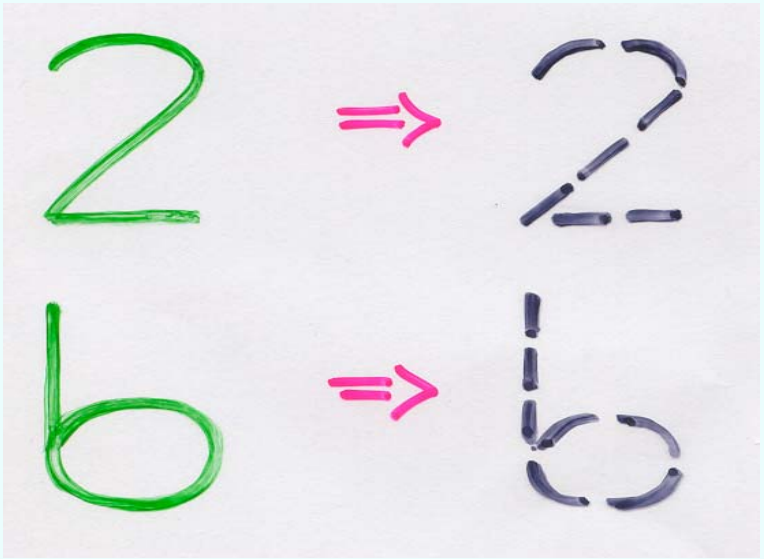
Definujeme primitiva

primitivum	/	|	–	\	(	)	⌒	⌣
označení	A	B	C	D	E	F	G	H

Strukturní popisy:

2: GAAC

6: BBHG (BBGH)



Definujeme primitiva

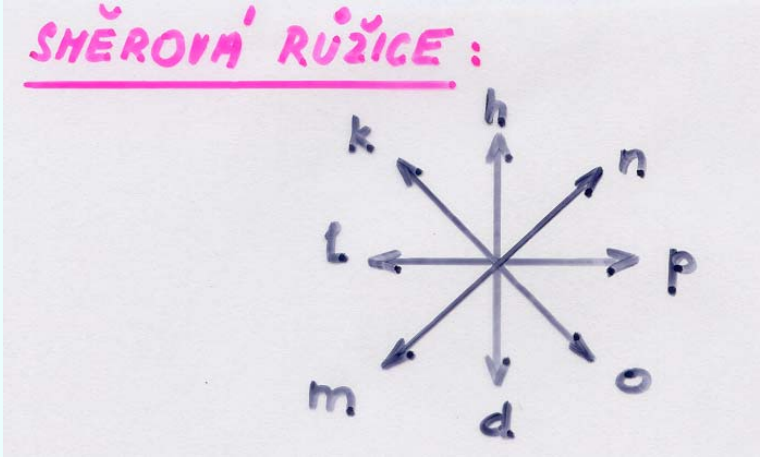
primitivum	/	|	–	\	⌜	⌝	⌞	⌟
označení	A	B	C	D	E	F	G	H

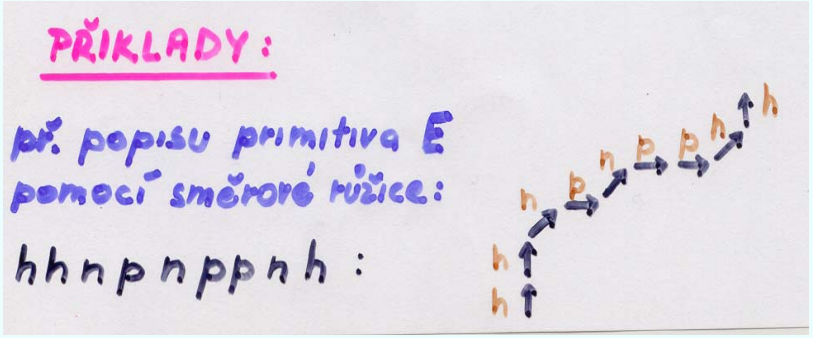
Strukturní popisy:

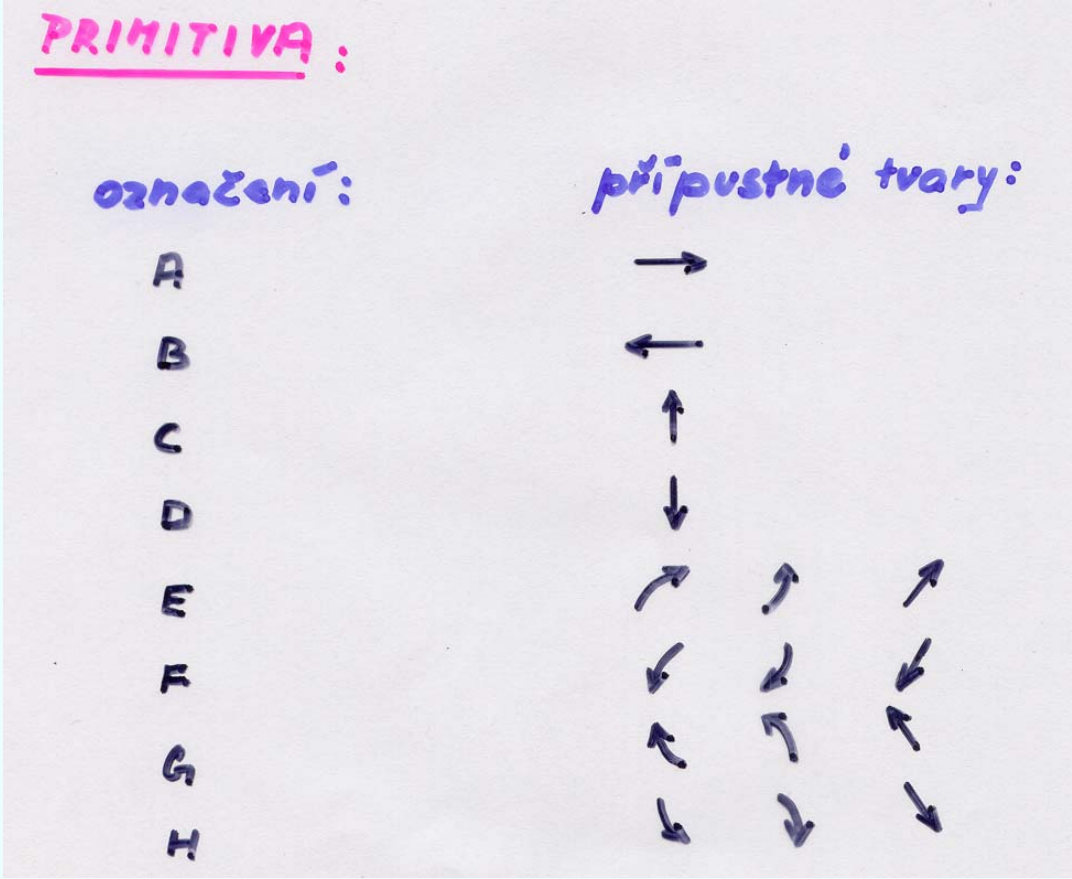
2: EFAAACC

6: BBBGHFE (BBBEFHG)

Praktický příklad extrakce primitiv



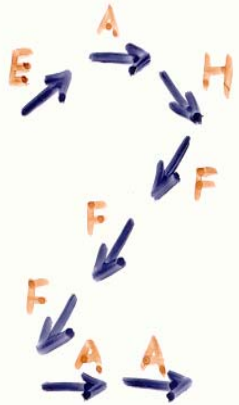




Strukturní popisy (obrazy) číslic **2** a **6** :

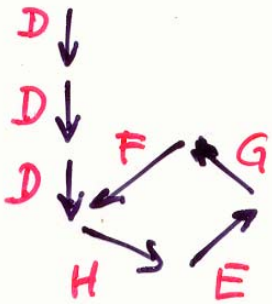
př. popisu číslice 2 pomocí primitiv :

EAHFFFAA:



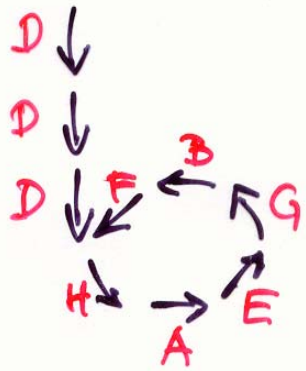
Číslice 6 :

DDDHEGF

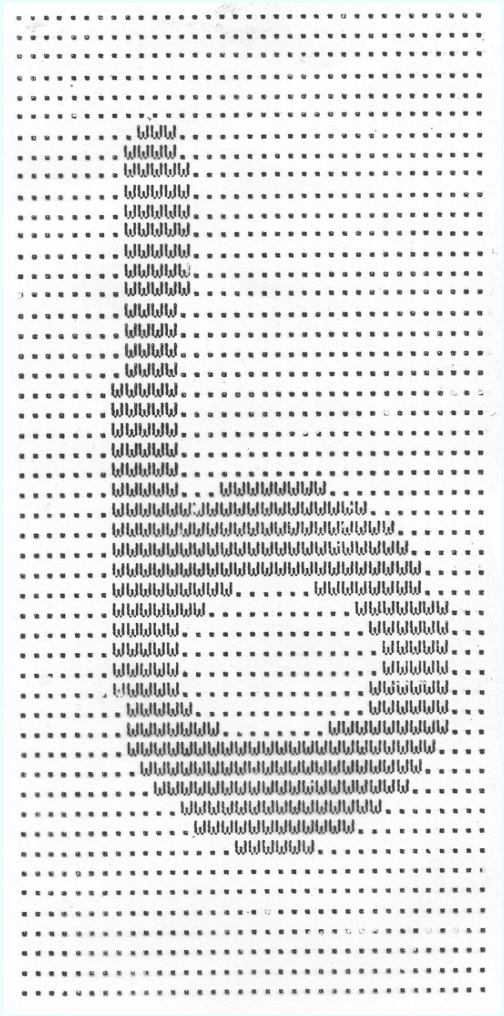


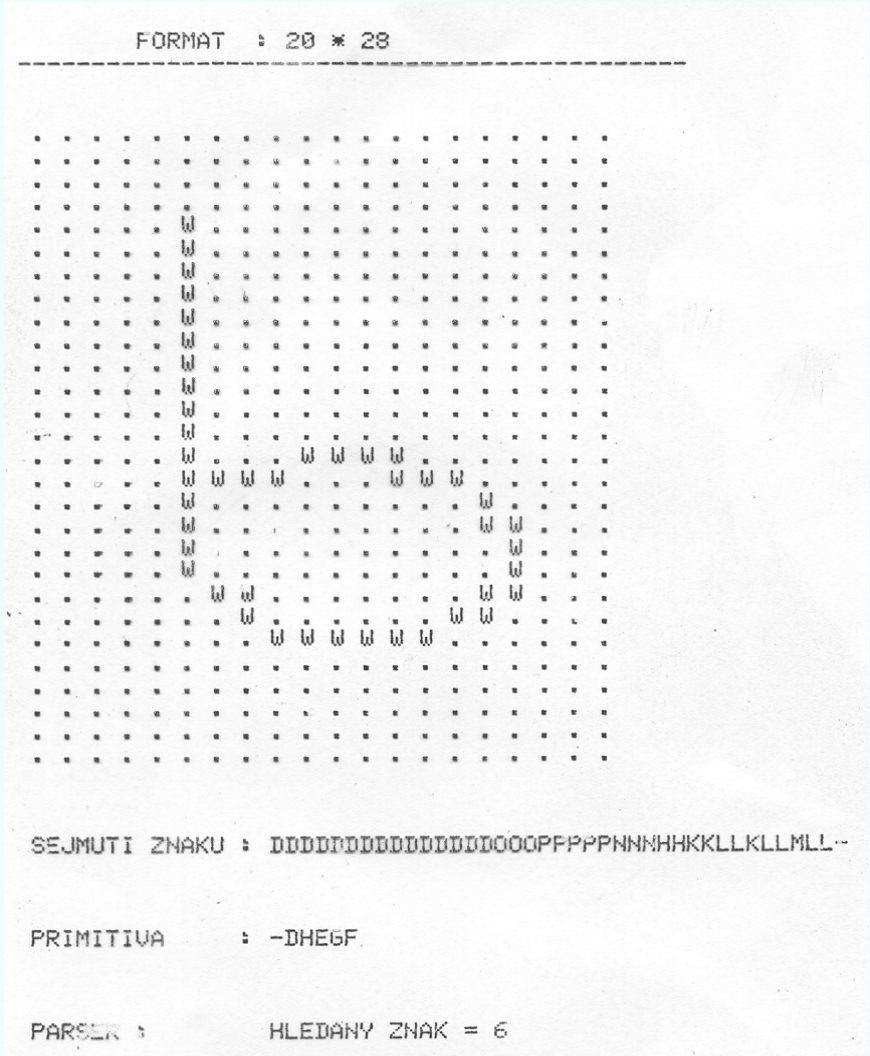
nebo

DDDHAEGBF

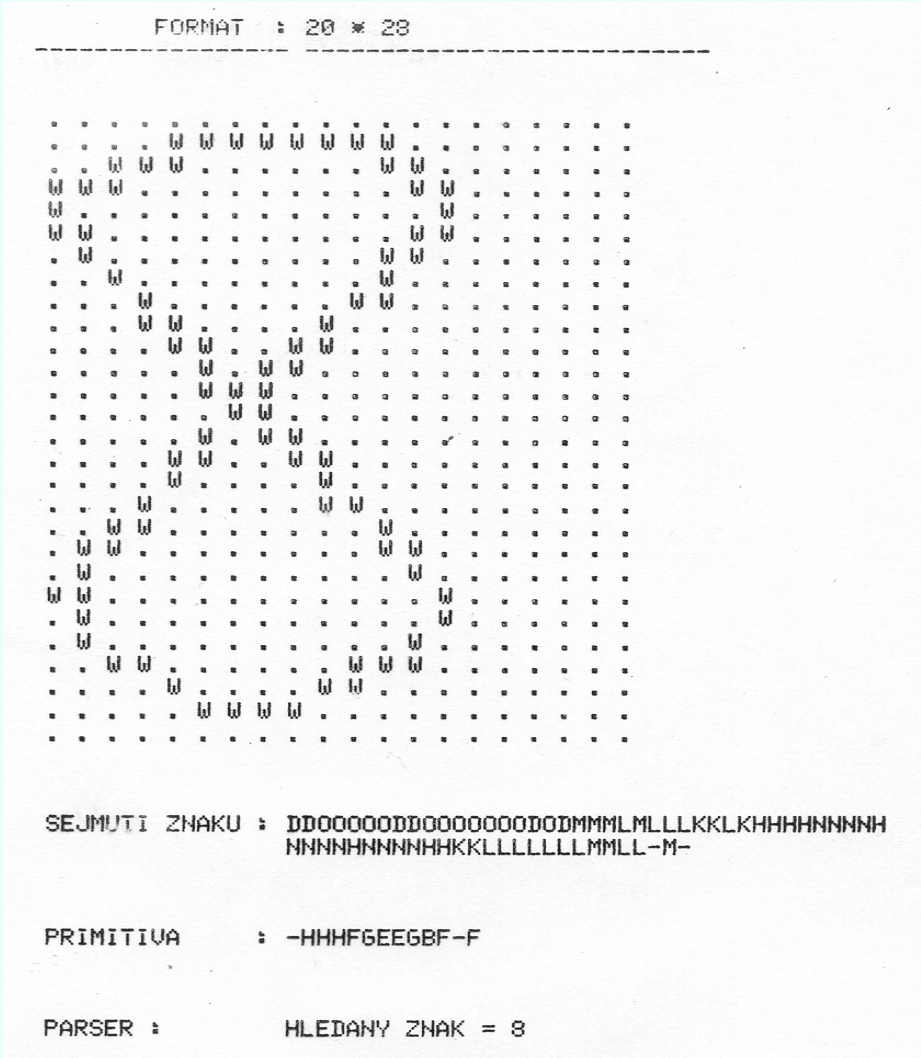


Zpracování a klasifikace číslice „**6**“

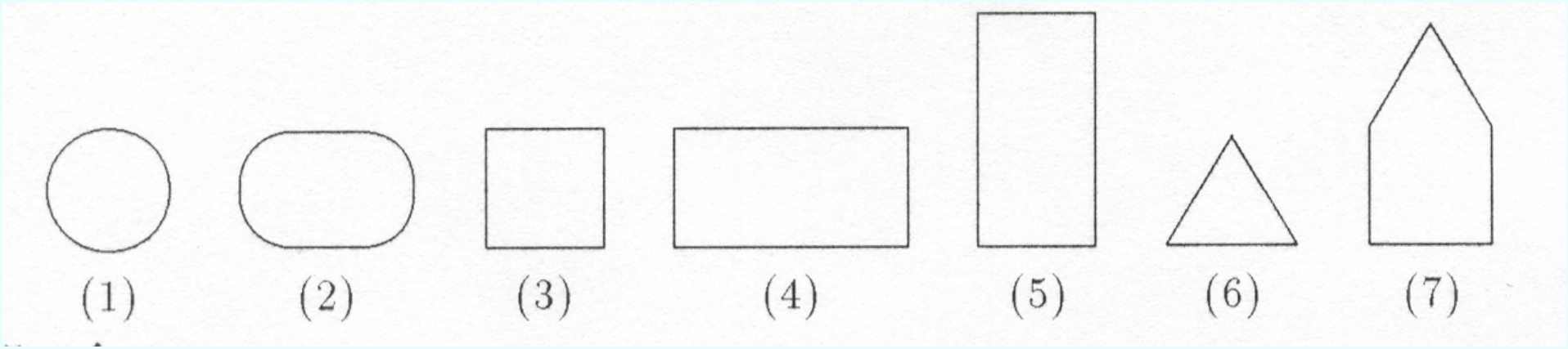




Totéž pro číslici „**8**“:



Strukturní rozpoznávání geometrických objektů



primitivum	—	|	/	\	(	)
symb. označení	a	b	c	d	e	f

Gramatiky a strukturní popisy objektů

$G_1 = < V_{N1}, V_{T1}, S_1, R_1 >$
$V_{N1} \equiv \{S_1, X_1, Y_1\}$
$V_{T1} \equiv \{\mathsf{e}, \mathsf{f}\}$
$R_1 \equiv \{S_1 \rightarrow X_1 Y_1,\ X_1 \rightarrow \mathsf{e},\ Y_1 \rightarrow \mathsf{f}\}$

$G_2 = < V_{N2}, V_{T2}, S_2, R_2 >$
$V_{N2} \equiv \{S_2, X_2, Y_2, Z\}$
$V_{T2} \equiv \{\mathsf{a}, \mathsf{e}, \mathsf{f}\}$
$R_2 \equiv \{S_2 \rightarrow X_2 Y_2,\ X_2 \rightarrow \mathsf{e}Z,\ Y_2 \rightarrow \mathsf{f}Z,\ Z \rightarrow \mathsf{a}Z \mid \mathsf{a}\}$

$G_3 = < V_{N3}, V_{T3}, S_3, R_3 >$
$V_{N3} \equiv \{S_3, X_3, Y_3\}$
$V_{T3} \equiv \{\mathsf{a}, \mathsf{b}\}$
$R_3 \equiv \{S_3 \rightarrow (X_3 Y_3)^2,\ X_3 \rightarrow \mathsf{b},\ Y_3 \rightarrow \mathsf{a}\}$

$G_4 = < V_{N4}, V_{T4}, S_4, R_4 >$
$V_{N4} \equiv \{S_4, X_4, Y_4\}$
$V_{T4} \equiv \{\mathsf{a}, \mathsf{b}\}$
$R_4 \equiv \{S_4 \rightarrow (X_4 Y_4)^2,\ X_4 \rightarrow \mathsf{b},\ Y_4 \rightarrow \mathsf{a}Y_4 \mid \mathsf{a}\}$

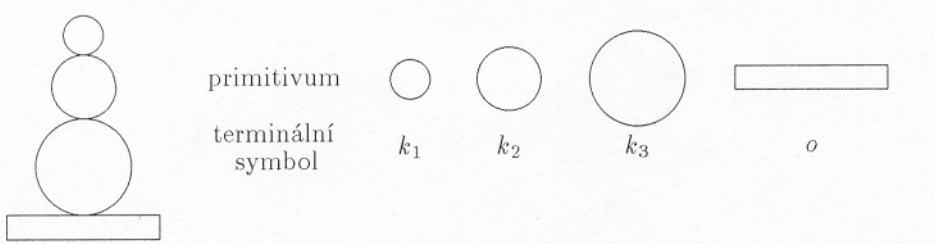
$G_5 = < V_{N5}, V_{T5}, S_5, R_5 >$
$V_{N5} \equiv \{S_5, X_5, Y_5\}$
$V_{T5} \equiv \{\mathsf{a}, \mathsf{b}\}$
$R_5 \equiv \{S_5 \rightarrow (X_5 Y_5)^2,\ X_5 \rightarrow \mathsf{b}X_4 \mid \mathsf{b},\ Y_5 \rightarrow \mathsf{a}\}$

$G_6 = < V_{N6}, V_{T6}, S_6, R_6 >$
$V_{N6} \equiv \{S_6, X_6\}$
$V_{T6} \equiv \{\mathsf{a}, \mathsf{c}, \mathsf{d}\}$
$R_6 \equiv \{S_6 \rightarrow \mathsf{c}X_6\mathsf{d},\ X_6 \rightarrow \mathsf{a}\}$

$G_7 = < V_{N7}, V_{T7}, S_7, R_7 >$
$V_{N7} \equiv \{S_7, X_7, Y_7\}$
$V_{T7} \equiv \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}\}$
$R_7 \equiv \{S_7 \rightarrow \mathsf{c}X_1\mathsf{d},\ X_7 \rightarrow Y_7\mathsf{a}Y_7,\ Y_7 \rightarrow \mathsf{b}Y_7 \mid \mathsf{b}\}$

(1) ef
(2) eafa, resp. eaafaa ... $\mathsf{e}(\mathsf{a})^n\mathsf{f}(\mathsf{a})^n$
(3) baba
(4) baabaa, resp. baaabaaa ... $\mathsf{b}(\mathsf{a})^n\mathsf{b}(\mathsf{a})^n$
(5) bbabba, resp. bbbabbba ... $(\mathsf{b})^n\mathsf{a}(\mathsf{b})^n\mathsf{a}$
(6) cad
(7) cbabd, resp. cbbabbd ... $\mathsf{c}(\mathsf{b})^n\mathsf{a}(\mathsf{b})^n\mathsf{d}$

Strukturní rozpoznávání s využitím sémantické informace



Obr. 5.8: Obrazec typu "sněhulák" a volba primitiv

Strukturní popis:

$$k_1\, k_2\, k_3\, o$$

„klasická" regulární gramatika:

$$G_{sněh} = < V_{N\,sněh}, V_{T\,sněh}, S, R_{sněh} > ,$$

kde $V_{N\,sněh} \equiv \{S, A, B, C\}$,
$V_{T\,sněh} \equiv \{k_1, k_2, k_3, o\}$,
$R_{sněh} \equiv \{S \rightarrow k_1 A,\ A \rightarrow k_2 B,\ B \rightarrow k_3 C,\ C \rightarrow o\}$.

Využijeme-li atributovou regulární gramatiku, dostaneme:

$$k_{(6)}\ k_{(8)}\ k_{(10)}\ o \quad .$$

Třídu "sněhuláků" lze pak reprezentovat atributovou regulární gramatikou

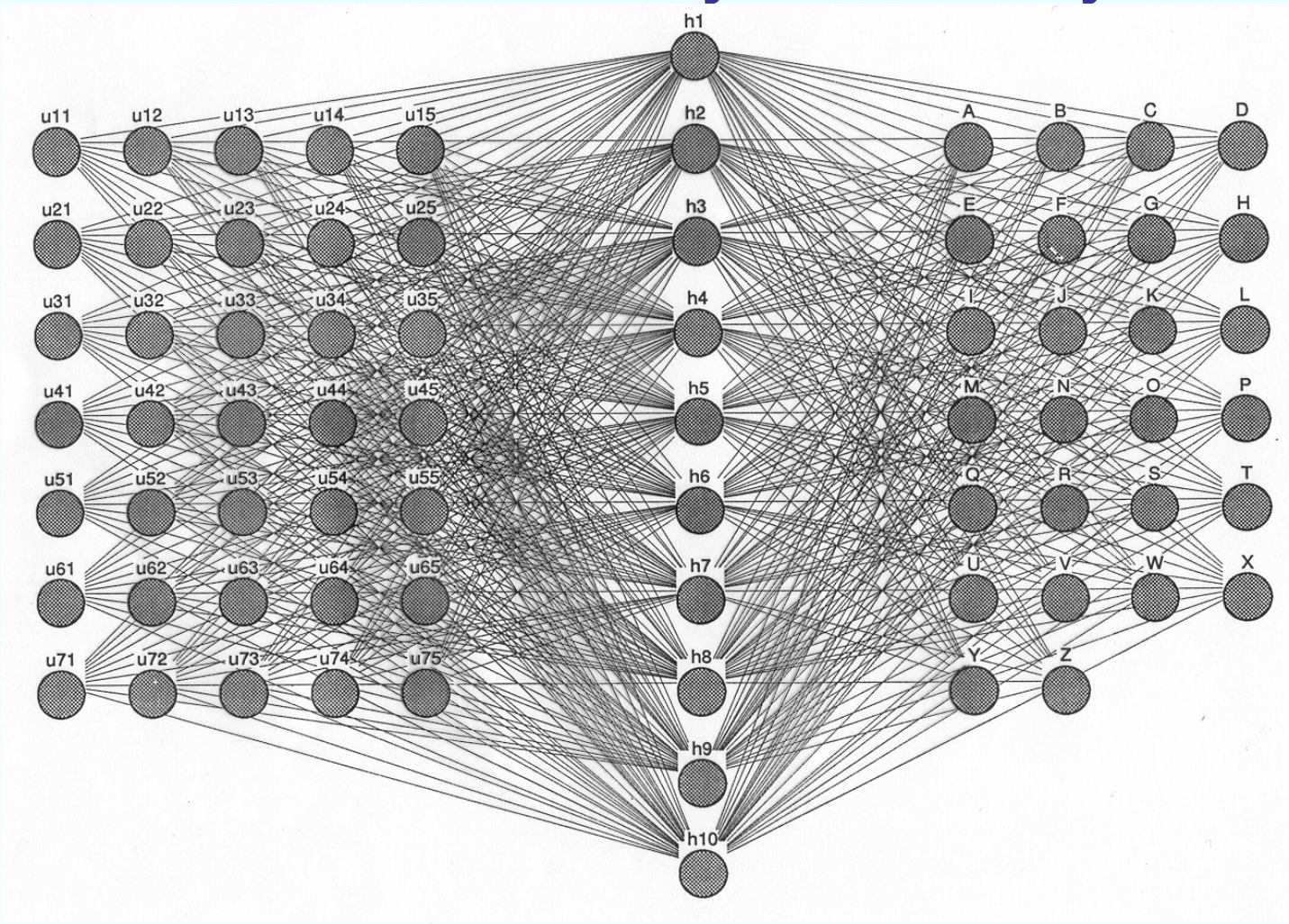
$$G_{a_sněh} = < V_{N\,a_sněh}, V_{T\,a_sněh}, S, R_{a_sněh}, \mathcal{C}_{a_sněh}, \mathcal{P}_{a_sněh}, \mathbf{v} > ,$$

kde $V_{N\,a_sněh} \equiv \{S, A\}$, $V_{T\,a_sněh} \equiv \{k, o\}$,
vektor $\mathbf{v}$ obsahuje jediný parametr v_1 označující poloměr koule a množina přepisovacích pravidel bude ve tvaru

$R_{a_sněh}$:

r_i	$\mathcal{C}_{a_sněh}$	$\mathcal{P}_{a_sněh}$
$S \rightarrow k\,A$		$v_1 = 0$
$A \rightarrow k\,A$	$par \geq v_1$	$v_1 = par$
$A \rightarrow o$		

Příklad klasifikace znaků umělými neuronovými sítěmi



Klasifikace znaku M:

