

# ROZPOZNÁVÁNÍ

## Úvod, vymezení hřiště

Václav Hlaváč

České vysoké učení technické v Praze  
Český institut informatiky, robotiky a kybernetiky  
160 00 Praha 6, Jugoslávských partyzánů 1580/3

<http://people.ciirc.cvut.cz/hlavac>, [vaclav.hlavac@cvut.cz](mailto:vaclav.hlavac@cvut.cz)

*Poděkování: M.I. Schlesinger, V. Franc*

### **Osnova přednášky:**

- ◆ Nadhled, gnozeologie
- ◆ Modelování a teorie systémů.
- ◆ Rozpoznávání a role učení.
- ◆ Statistické × strukturní rozpoznávání.
- ◆ Bayesovská formulace úlohy.
- ◆ Co se umí v rozpoznávání?



## Co je rozpoznávání / strojové učení ?

- ◆ **Epistemologie**, česky též gnozeologie, je část filozofie zabývající se původem, podstatou, metodami a možnostmi poznání/znalosti. Rozpoznávání je jednou z metod.
  - ◆ **Rozpoznávání / strojové učení** (téměř synonyma) je vědecká disciplína vytvářející a studující algoritmy, které se učí vytvářením statistických modelů z dat a používají se pro rozhodování a předvídání.
  - ◆ “**Rozpoznávání** přiřazuje skutečný objekt nebo událost do jedné nebo více předem stanovených tříd” – kniha Duda & Hart 1977, 2001”.
- 
- ◆ Satoshi Watanabe, teoretický fyzik (\*1910 Tokyo, †1993 Tokyo, činný v Německu, Dánsku a USA): “*Vzor je protikladem k chaosu; Je to nejistě definovaná entita, které lze dát jméno.*” (1981)
  - ◆ Isaiah Berlin, philosopher (\*1909 Riga, †1997 Oxford): “*Porozumět znamená vnímat vzory. Pochopení spočívá v odhalení základního vzoru.*”

## Pojmy

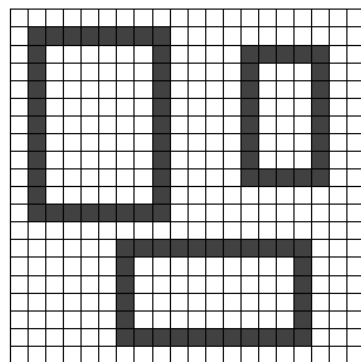
- ◆ **Vzor** je objekt, proces nebo událost, které lze pojmenovat.
- ◆ **Třída vzorů** (nebo kategorie) je množina  $M \subset X$  vzorů, jejíž prvky sdílejí podobné rysy, tj. konečné rozpoznatelné vlastnosti (popsané příznaky).
- ◆ **Klasifikace** (nebo rozpoznávání) přiřazuje daný objekt do předem daných tříd.
- ◆ **Klasifikátor** je stroj (program), který klasifikaci realizuje.
- ◆ A **vzorek** (příklad, objekt, situace) je jedna instance vzoru, který míníme klasifikovat.
- ◆ **Trénovací multimnožina** (trénovací data) je soubor příkladů, který se užívá k učení statistického modelu (klasifikátoru).
- ◆ **Testovací multimnožina** (testovací data) je soubor příkladů, které jsou jiné než v trénovací multimnožině.
- ◆ **Příznaky** jsou vlastnosti určující vlastnosti příkladů podstatných pro uvažovanou klasifikaci.

## Poznámka k názvu v češtině

- ◆ Význam českého názvu **rozpoznávání** chápou jako ekvivalent disciplíny anglicky nazývané **pattern recognition**.
- ◆ Zejména v dřívějších českých publikacích z šedesátých až sedmdesátých let zazníval ve stejném smyslu i chybně zavedený pojem rozpoznávání obrazců. Do češtiny asi přešel z původního anglického názvu přes ruský překlad – raspoznavanie obrazcov.
- ◆ V ruštině obrazec odpovídá českému **vzor** či **anglickému pattern**. Obrazec není totéž co obraz.

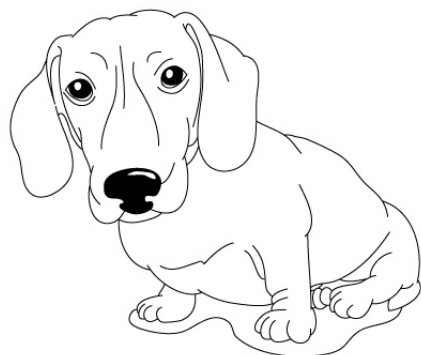
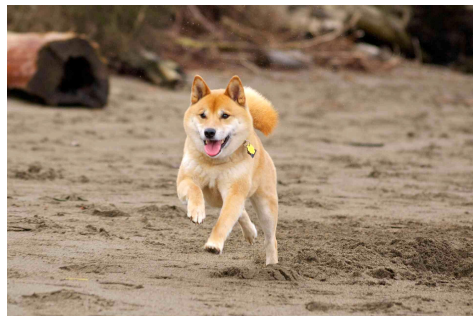
## Třída vzorů, příklady (1)

- ◆ Třída syntakticky správných aritmetických výrazů, např.  $2x(a + 3b) - 6y + (x - y)/7$ .
- ◆  $M$  je podmnožinou množiny  $X$  všech konečných řetězů nad nějakou abecedou.  $M$  lze popsat bezkontextovou gramatikou.
- ◆ Třída binárních obrazů obsahující nepřekrývající a sebe se nedotýkající obdélníkové rámečky s jednopixelovou tloušťkou.  $M$  je podmnožinou množiny  $X$  všech pravoúhlých binárních obrazů.



## Třída vzorů, příklady (2)

- ◆ Množina všech psů v obrazech.



## Základní pojmy, ilustrace

- ◆ Studovaný **vzor** se analyzuje (například brambora, viz obrázek).
- ◆ **Vektor příznaků**  $\mathbf{x} \in X$  je vektor tvořený jednotlivými pozorováními (měřeními). Vektor  $\mathbf{x}$  odpovídá jednomu bodu v prostoru příznaků  $X$ .
- ◆ **Skrytý stav** (ve obvyklém případě přímo značka třídy)  $y \in Y$  není přímo pozorovatelný. Vzory se stejnými skrytými stavy vytvářejí jednu třídu.
- ◆ **Úkolem je navrhnout klasifikátor** (rozhodovací pravidlo)  $q: X \rightarrow Y$ , které přiřazuje pozorované instance vzoru ke skrytému stavu.

skrytý stav (nebo značka třídy)  $y$



studovaný vzor

vektor příznaků  $\mathbf{x}$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

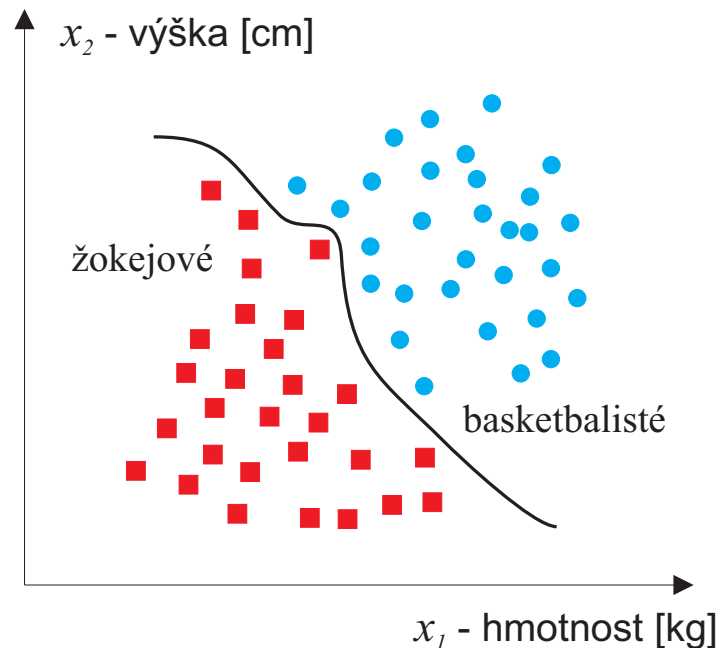
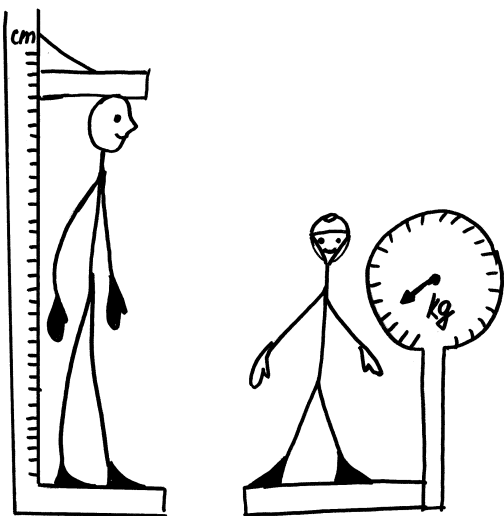
# Rozpoznávání, motivační příklad

**Objekt (situace)** se popisuje dvěma parametry:

$x$  – pozorovatelný *příznak* (též pozorování).

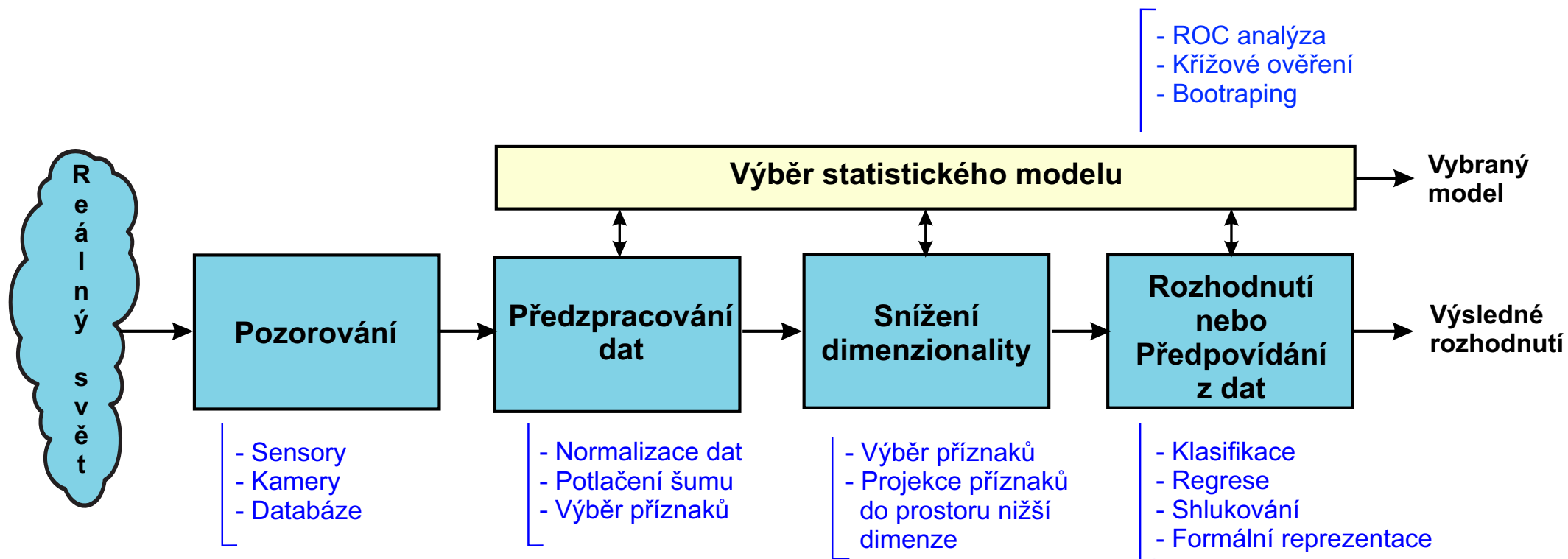
$y$  – *skrytý parametr* (stav, speciální případ—klasifikační třída).

**Příklad** statistické rozpoznávání: *žokejové a basketbalisté*.



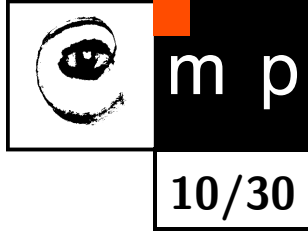


# Celkový pohled, části



- ◆ Input: Data, trénovací (multi)-množina.
- ◆ Statistické modely a jejich parametry se empiricky učí z trénovacích dat.
- ◆ Výstupy: různorodá rozhodnutí, viz obrázek.

# Dávná vědecká úloha, gnozeologie



- ◆ Podstata klasifikace a rozhodování je hlavním tématem části filozofie, **gnozeologie** (teorii poznání), která studuje podstatu znalosti.
- ◆ Základy rozpoznávání lze tedy odkázat až k **Platónovi** (Πλατων, 427 př.n.l. – 347 př.n.l.) a jeho žákovi **Aristotelovi** (Αριστοτελης, 384 př.n.l. – 322 př.n.l.).
- ◆ Oba rozlišovali mezi:
  - *základními vlastnostmi* sdílenými všemi příslušníky třídy;
  - *nahodilými vlastnostmi*, kterými se mohou jednotliví příslušníci jedné třídy lišit.

# Klasifikace/kategorizace (nebo popis podle toho, k čemu objekt slouží)





## Typy možných rozhodnutí / predikčných úloh

12/30

**Klasifikace** – přiřazuje pozorování do jedné z malé množiny tříd. Výstupem je identifikátor třídy, její značka. Např. značka označující kvalitu jablka jako A, B, C a odmítnutí (zmetek).

**Regrese** – předpovídá hodnotu z pozorování. Zobecňuje klasifikaci. Např. výstupem může být reálné číslo odhadující příští hodnotu akcie na burze podle předchozích hodnot a dalších indikátorů chování akciového trhu.

**Učení bez učitele** (shlukování) uspořádává pozorování do smysluplných tříd podle jejich vzájemných podobností. Např. v genetice hledá skupiny genů s podobnými vzory exprese.

---

**Reprezentace strukturních vztahů** se opírá o primitiva, např. vyjádřuje člověka pozorovaného dohlížecí kamerou pomocí předem poloh těla a s nimi spojených prototypů aktivit.

## Další obory sdílející podobné hlavní myšlenky

**Statistické modelování** – hledá (generativní) model popisující objekt zájmu, např. pravděp. rozdělení a ohodnocuje kvalitu modelu statistickými metodami.

**Strojové učení** (což je dnes módnější název pro rozpoznávání) – při dané trénovací množině se má rozhodovací pravidlo naučit automaticky. Člověkem zadaná (subjektivní) pravidla nejsou použita. Různé úlohy potřebují různé trénovací množiny.

**Dolování v datech** – hledání explicitních, předem neznámých a potenciálně užitečných znalostí v datech.

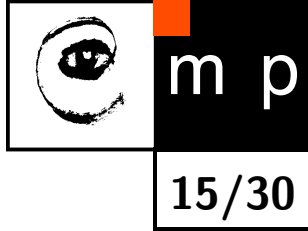
**Vizualizace ve vědě** – vysokedimenzionální úloha se člověku zobrazuje v pro něj přirozeném 2D obrázku nebo 3D scéně. My lidé více dimenzí nevidíme.

**Neuronové sítě** – jeden z matematických formalismů řešící rozhodovací bez nutnosti vytvářet generativní model skutečného biologického systému.

## Biologická motivace

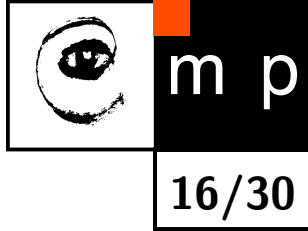
- ◆ Člověk je na špičce pomyslné pyramidy živočichů i proto, že je schopen přemýšlet o postupech, jakými sám uvažuje.
- ◆ Panuje všeobecný zájem o strojové napodobení biologického vnímání s cílem napodobit inteligentní chování v nepříliš známém prostředí.
- ◆ Základním atributem inteligentního chování je schopnost **učit se** na základě vnímání okolního prostředí.
- ◆ Klíčová je **otázka reprezentace znalosti**. Přirozený jazyk je nejdokonalejší nástroj lidí pro vyjádření pozorování, pro popis jevu, formulaci úloh, jejich řešení a pro související otázky učení.

# Složité jevy a systémové myšlení



- ◆ Potřeba porozumět složitým jevům například v biologii, technice nebo sociálních vědách vede k nutnosti zkoumat jevy komplexně v mnoha souvislostech.
- ◆ Přístup je nazýván **systemovým myšlením**, aby se odlišil od newtonovské snahy zredukovat každý jev na vztahy mezi základními prvky a jejich vlastnostmi.

# Pojmy z teorie systémů



- ◆ Při zkoumání složitého jevu se omezujeme na část, která nás zajímá, a říkáme jí **objekt** (nebo systém).
- ◆ Vše ostatní, co nám z daného pohledu připadá nezajímavé, nazýváme **pozadí**.
- ◆ Objekty většinou nezkoumáme v celé jejich složitosti. Při jednom zkoumání pozorujeme nebo měříme jen určité vlastnosti, které nám právě připadají zajímavé. Teorie systémů zde používá pojem **rozlišovací úroveň**.
- ◆ Popis a chápání objektu se přirozeně může vyvíjet s měnící se rozlišovací úrovní. Jde o metapohled hledající kvalitativní změnu v popisu objektu.





# Generativní × diskriminativní reprezenatace objektu

Snaha o exaktní popis objektů (složitých dějů) matematickými nástroji vyústila (zhruba řečeno) ve dva možné přístupy:

## 1. Generativní $\Rightarrow$ modelování.

- ◆ Snaží se o porozumění fyzikálním / jiných principům a jejich vyjádření modelem.
- ◆ Tento model umí generovat data dosti se podobající empirickým pozorováním.
- ◆ Příkladem je matematické modelování fyzikálního / technického děje (v newtonovském smyslu).

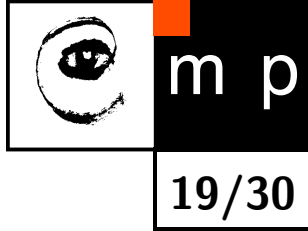
## 2. Diskriminativní $\Rightarrow$ klasifikace.

- ◆ Snaží se porozumět vnějšímu chování bez detailní znalosti dílčích principů (což se u složitých objektů / dějů ani neumí).
- ◆ Výstupem jsou rozhodnutí / předvídání ve smyslu regrese.
- ◆ Příkladem je rozpoznávání, např. stanovení diagnózy lékařem / počítačovým programem.

## Matematické modelování

- ◆ Podstatné rysy objektu se napodobují formou matematických rovnic. Často se hledá relace mezi vstupem a výstupem.
  - ◆ Má blíže k newtonovskému pojetí. Snaha o co nejpodrobnější a deterministické vysvětlení.
- 
- ◆ *Příklad:* dobrý matematický model elektrárenského kotle v teorii řízení předpovídá téměř stejné odezvy na vstupní signály jako kotel sám.
  - ◆ *Protipříklad 1:* V mnoha případech nejsme schopni matematický model vůbec vytvořit (např. model fungování lidského těla).
  - ◆ *Protipříklad 2:* Počítačové vidění. Inverzní úloha k fyzikálnímu postupu vzniku obrazu je příliš složitá, a tudíž prakticky nepoužitelná.

# Alternativou k modelování je rozpoznávání



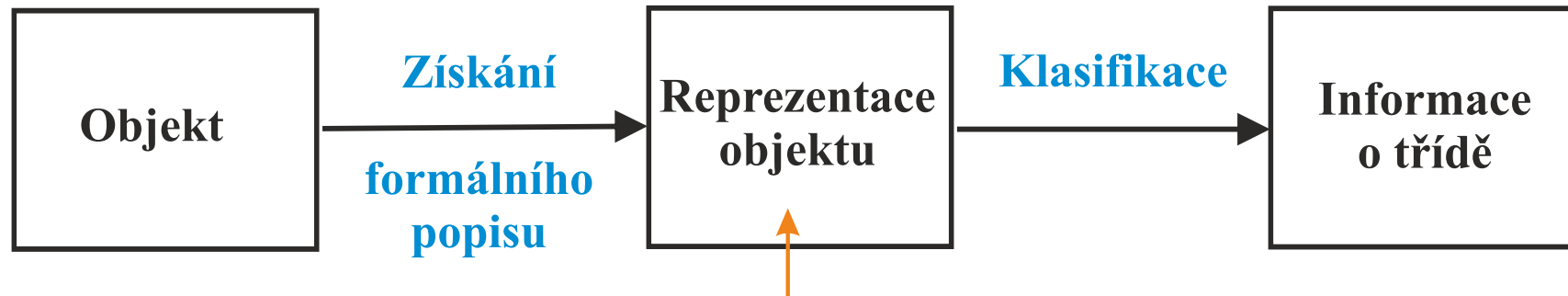
- ◆ Rozpoznávání **zařazuje** pozorování podle nějakého rozhodovacího pravidla **do** předem známých tříd.
- ◆ Třídy ekvivalence (relace ekvivalence: reflexivní, symetrická, tranzitivní).
- ◆ Uvnitř těchto tříd jsou si objekty podobnější než mezi třídami navzájem.
- ◆ V rozpoznávání bývá porozumění objektu méně podrobné než v modelování.

## Role učení v rozpoznávání

- ◆ Výhodou rozpoznávání je, že člověk vytvářející rozhodovací pravidlo (rozhodovací strategii, klasifikátor) nemusí rozumět složité podstatě objektu či jevu, o kterém se má rozhodovat.
- ◆ Rozhodovací pravidlo může být naučeno empiricky z mnoha pozorovaných příkladů.
- ◆ Paradox znalostního inženýrství: Pro člověka je snazší poskytnout příklady správné klasifikace než explicitně vyjádřit pravidlo, podle kterého rozhoduje.
- ◆ Tři hlavní přístupy k učení:
  - Učení s učitelem** na základě trénovací multimnožiny zahrnující pozorování a informaci o třídě, kterou přisoudil učitel (znalec).
  - Učení bez učitele** na základě hledání podobnosti mezi pozorováními, aniž by byla k dispozici znalcová klasifikace.
  - Podporované učení** (reinforcement learning) místo informace od učitele odhaduje odměny nebo pokuty z prostředí. Maximalizuje se kumulativní odměna.

# Metody rozpoznávání a aplikace

Velkou výhodou rozpoznávání je, že je lze oddělit od aplikačních disciplín.



**Příznaky ve statistickém rozpoznávání**  
**Primitiva ve strukturním rozpoznávání**

# Hlavní přístupy k rozpoznávání

## 1. Statistické (příznakové) rozpoznávání.

- ◆ Předpokládá se existence statistického modelu jednotlivých vzorů a tříd vzorů.
- ◆ Souřadné osy prostoru odpovídají jednotlivým číselně vyjádřeným pozorováním, tedy příznakům.
- ◆ Objekty jsou reprezentovány jako body ve vektorovém prostoru.

## 2. Strukturní rozpoznávání.

- ◆ Mezi pozorováními existuje struktura a ta je reprezentována.
- ◆ Nejrozvinutější a nejstarší je reprezentace struktury gramatikami.

3. **Umělé neuronové sítě.** Klasifikátor je realizován sítí navzájem propojených uzlů, které modelují neurony v lidském mozku (přístup konekcionisty, např. model dopředné neuronové sítě (McCulloch, Pitts, 1943).

# Bayesovské rozhodování

Bayesovská úloha statistického rozpoznávání (rozhodování) hledá

pro množiny  $X$  (pozorování),  $Y$  (skryté stavy) and  $D$  (rozhodnutí) sdruženou pravděpodobnost  $p_{XY}: X \times Y \rightarrow \mathbb{R}$  a pokutovou funkci  $W: Y \times D \rightarrow \mathbb{R}$

strategii  $q: X \rightarrow D$ , která **minimalizuje bayesovské riziko**

$$R(q) = \sum_{x \in X} \sum_{y \in Y} p_{XY}(x, y) W(y, q(x)) .$$

Řešením bayesovské úlohy je **bayesovská strategie**  $q$  minimalizující riziko.

- 
- ◆ Poznámky: deterministická strategie není nikdy horší než náhodná strategie; rozdění na konvexní podmnožiny.
  - ◆ **Klasifikace** je speciálním případem rozhodovací úlohy, při které jsou totožné množiny rozhodnutí  $D$  a množiny skrytých stavů  $Y$ .

# Obecnost bayesovské formulace rozhodování (1)



24/30

Motto: *“Nechť jsou množiny  $X$  (pozorování) a  $Y$  (skryté stavy) dvě konečné množiny.”*

- ◆ Úlohy statistického rozpoznávání jsou velmi obecné, protože množiny pozorování  $X$  a skrytých stavů  $Y$  nebyly nijak omezeny.
- ◆ Množiny  $X$  a  $Y$  mohou mít i při formálním vyjádření velmi různorodou (matematickou) strukturu.
- ◆ Přístup lze tudíž použít ve velmi různorodých aplikacích.



# Obecnost bayesovské formulace rozhodování (2)

Pozorování  $x$  může být číslo, symbol, funkce dvou proměnných (e.g., např. obrázek), graf, algebraická struktura, atd.

Aplikace	Pozorování	Rozhodnutí
hodnota mince v mincovním automatu rozpoznávání znaků (OCR) rozpoznávání registračních značek aut rozpoznávání otisků prstů rozpoznávání řeči analýza EEG, EKG signálů forfeit detection identifikace řečníka ověření řečníka	$\mathbf{x} \in \mathbb{R}^n$ 2D šedotónový obrázek 2D šedotónový obrázek 2D šedotónový obrázek signál z mikrofону $x(t)$ $\mathbf{x}(t)$ various signál z mikrofону $x(t)$ signál z mikrofону $x(t)$	hodnota znaky, slova písmena, číslice identita člověka slova diagnóza {ano, ne} identita člověka {ano, ne}

## Generativní × diskriminativní klasifikátor

- ◆ Viz obecnější rozlišení mezi generativními a diskriminativními modely na průsvitce č. 17 této přednášky.
- ◆ Chceme se naučit buď rozhodovací pravidlo  $q: X \rightarrow Y$  nebo a posteriorní pravděpodobnost  $P(Y|X)$ .

**Generativní klasifikátory**, např. naivní Bayesův klasifikátor, model gaussovské směsi, ...

- ◆ Předpokládají se funkce  $P(X|Y)$ ,  $P(Y)$ .
- ◆ Odhadují se  $P(X|Y)$ ,  $P(y)$  přímo z trénovacích dat.
- ◆ Použije se Bayesův vzorec k výpočtu  $P(Y|X = x_i)$
- ◆ “Generativní” znamená, že model poskytuje data získaná vzorkováním pravděpodobnostní distribuce.

**Diskriminativní klasifikátory**, např.. perceptron, SVM,  $k$ -NN, ...

- ◆ Předpokládá se funkce, kterou je a posteriorní pravděpodobnost  $P(Y|X)$ .
- ◆ Odhaduje se  $P(Y|X)$  z trénovacích dat.
- ◆ “Diskriminativní” znamená, že model umožňuje klasifikaci  $x$ , ale neumožňuje generování  $x$  splňujícího pravděpodobnostní model.

## Co se umí ve statistickém rozpoznávání?

- ◆ Vyřešit několik málo nebayesovských úloh, např. s třídou “nevím” (povolené odmítnutí rozhodnout), s minimaxním klasifikátorem nebo úlohy s nenáhodnými zásahy.
- ◆ Lineární klasifikátory a jejich učení. Např. SVM klasifikátor s podpůrnými vektory (Support Vector Machines).
- ◆ Odhad potřebné délky trénovací multimnožiny pro dosažení předepsané přesnosti a míry důvěry klasifikátoru, např. Vapnikova-Červnoněnkisova teorie učení.
- ◆ Řešení pro nelineární úlohy jejich zanořením do vícedimenzionálního vektorového prostoru, které umožní použít lineární klasifikátory (vyrovnání celého příznakového prostoru nebo lokálně působící jádrové metody).
- ◆ Učení bez učitele. Různé varianty EM algoritmu.

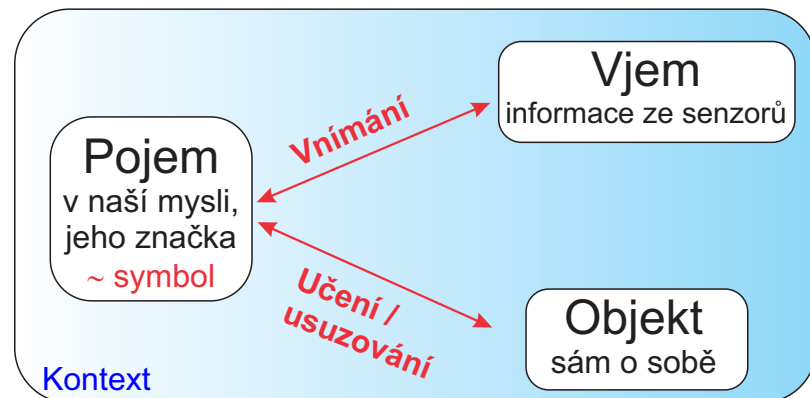
## O použití matematické statistiky

- ◆ Nejrozvinutější je statistika náhodných čísel.
  - ◆ Poskytuje doporučení opírající se o pojmy jako: střední hodnota (matematické očekávání), rozptyl, korelace, kovarianční matice, ...
  - ◆ Nástroje matematické statistiky slouží k řešení mnoha praktických úloh **za předpokladu, že náhodný objekt může být reprezentován číslem** (nebo vektorem čísel).
  - ◆ Statistické rozpoznávání slaví významné **úspěchy** pro objekty vyjádřené jako **vektory příznaků**.
- 
- ◆ **Selhání pro obrázky.** Viz příští průsvitka.

## Analýza obrazů & objekty

- ◆ Selhání pro obrázky  $f(x, y)$ , kde  $f$  je jas nebo barva pixelu a  $x, y$  jsou souřadnice pixelu.
- ◆ Snaha obrátit proces pořízení obrazu vede na špatně podmíněné úlohy, které je činí prakticky nepoužitelnými.
- ◆ Potřebujeme se opřít o pojem **objekt** a jeho sémantiku.
- ◆ Detekce objektů, jejich segmentace např. v obrázcích je příkladem úlohy: Co bylo dříve? Slepice nebo vejce. Hledá se vztah mezi vzhledem a sémantikou.

- 
- ◆ Znalost  $\sim$  pozorování + kontext + zkušenost.
  - ◆ Úloha ukotvení symbolů (angl. symbol grounding).



## Doporučené čtení

- ◆ Duda Richard O., Hart Peter E., Stork, David G.: Pattern Classification, John Wiley & Sons, New York, USA, 2001, 654 p.
- ◆ Schlesinger M.I., Hlaváč V.: Ten lectures on statistical and syntactic pattern recognition, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002, 521 p. (předchůdce v češtině, Vydavatelství ČVUT 1999).
- ◆ Bishop C.: Pattern Recognition and Machine Learning, Springer-Verlag New York 2006, 758 p.

