

# Co je to informace?

**Základní koncepty a definice**

# Pozdní příchod kamaráda

## Představte si situaci...

- Kamarád měl přijít v 15:00
- Je 15:30 a stále tu není
- Co všechno se mohlo stát?

*Pojďme společně sestavit seznam možností...*

# Možné důvody zpoždění

## První nejistota

- Dopravní problémy
- Zdravotní komplikace
- Zaspal
- Zapomněl na schůzku
- Špatně si zapsal čas
- Má jinou neodkladnou povinnost
- ...další možnosti?

# První zpráva

## "Uvízl v dopravě"

### Co jsme se dozvěděli?

- Víme, že je na cestě a  Víme, že používá dopravu
- A tady víme, že  Nezaspal  Není nemocný a  Nezapomněl

### Co stále nevíme?

- Kde přesně je?
- Jaký druh dopravy používá?
- Jak dlouho bude trvat zpoždění?

# Druhá zpráva

## "Stojí na D1"

### Další vyloučené možnosti:

- ✗ MHD
- ✗ Železniční doprava
- ✗ Lokální silnice

### Zbývající nejistoty:

- Přesná lokace
- Důvod zdržení
- Předpokládaný čas příjezdu

# Třetí zpráva

"Kvůli nehodě"

Nyní víme:

- Přesnou příčinu zpoždění
- Přibližnou lokaci
- Způsob dopravy

*S každou zprávou se zmenšuje počet možných scénářů*

# Co je tedy informace?

## Klíčová definice

### Informace je:

- Úbytek možností
- Snížení nejistoty
- Vyloučení alternativ

*Čím více možností vyloučíme, tím více informace jsme získali*

# Vlastnosti informace

## Důležité charakteristiky

### 1. Závisí na předchozí znalosti

- Stejná zpráva → různá informace pro různé příjemce

### 2. Je měřitelná

- Více vyloučených možností = více informace

### 3. Je objektivní

- Nezávisí na obsahu, ale na počtu vyloučených možností



# Shrnutí

## Klíčové poznatky

1. Informace snižuje nejistotu
2. Každá zpráva vylučuje některé možnosti
3. Více zpráv = méně nejistoty
4. Informace se skládají
5. Lze je měřit

# **Jak efektivně získávat informace?**

**Praktické ukázky a hry**

# Hra: Myslím si zvíře

## Pravidla

1. Jeden hráč si myslí zvíře
2. Ostatní kladou otázky
3. Odpovědi mohou být jen ANO/NE
4. Cíl: Uhodnout zvíře na co nejméně otázek

*Pojďme si to vyzkoušet!*

# Špatné vs. dobré otázky

## Příklady otázek

### ✗ Špatné otázky:

- "Je velké?" (subjektivní)
- "Je to kočka?" (jen jedna možnost)
- "Žije v Africe nebo Asii?" (složená)

### ✓ Dobré otázky:

- "Má více než 4 nohy?"
- "Žije ve vodě?"
- "Je to savec?"

# Proč jsou některé otázky lepší?

## Analýza strategií

### Efektivní otázka:

- Je jednoznačná
- Dělí množinu možností na podobné části
- Má jasnou odpověď ANO/NE

### Příklad:

"Má to 4 nohy?"

- ANO → vyloučí hady, ryby, ptáky, hmyz...
- NE → vyloučí psy, kočky, koně...

# Hra: Myslím si číslo

## Nová varianta hry

### Pravidla:

- Číslo je od 1 do 16
- Opět jen otázky ANO/NE
- Cíl: Minimum otázek

*Jaká bude nejlepší strategie?*

# Strategie hádání čísel

## Porovnání přístupů

### Přístup 1: Postupné tipy

- "Je to 1?" → "Ne"
- "Je to 2?" → "Ne"
- "Je to 3?" → "Ne"
- ...

*Může vyžadovat až 16 otázek!*

### Přístup 2: Půlení intervalu

- "Je větší než 8?" → Vyloučí polovinu možností
- *Vždy maximálně 4 otázky*

# Půlení intervalu

## Proč je efektivní?

### První otázka: "Je větší než 8?"

- ANO → zůstává 9-16
- NE → zůstává 1-8

### Druhá otázka:

- Pro 1-8: "Je větší než 4?"
- Pro 9-16: "Je větší než 12?"

*Každá otázka vyloučí polovinu zbývajících možností*



# Optimální strategie

## Klíčové principy

1. Dělit možnosti na poloviny
2. Používat objektivní kritéria
3. Klást jednoznačné otázky
4. Využívat předchozí odpovědi

*Tyto principy fungují univerzálně*

# Shrnutí a aplikace

## Co jsme se naučili?

1. Efektivní otázky dělí množinu možností na poloviny
2. Subjektivní kritéria vedou k nejasnostem
3. Postupné tipování je neefektivní
4. Počet otázek lze předem odhadnout

*V další části si ukážeme, jak tento proces vizualizovat pomocí rozhodovacích stromů...*

# Rozhodovací stromy

## Vizualizace rozhodovacího procesu

[podrobné vysvětlení](#)

# Co je rozhodovací strom?

## Základní prvky

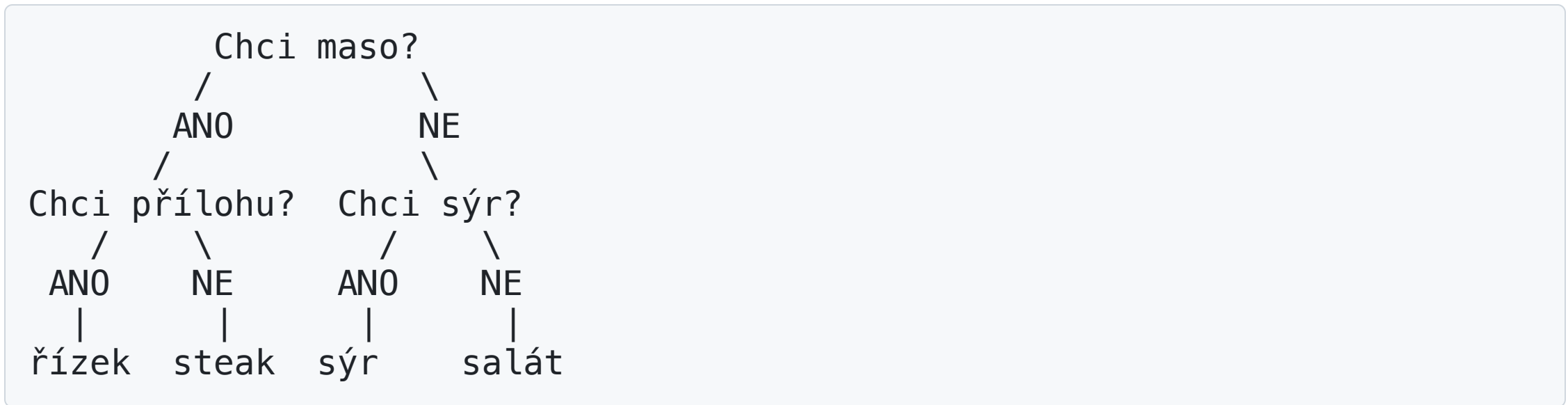
### Struktura stromu:

- **Kořen** = první otázka
- **Uzly** = další otázky
- **Hrany** = možné odpovědi
- **Listy** = konečné výsledky

*Strom vizualizuje celý proces rozhodování*

# Jednoduchý příklad

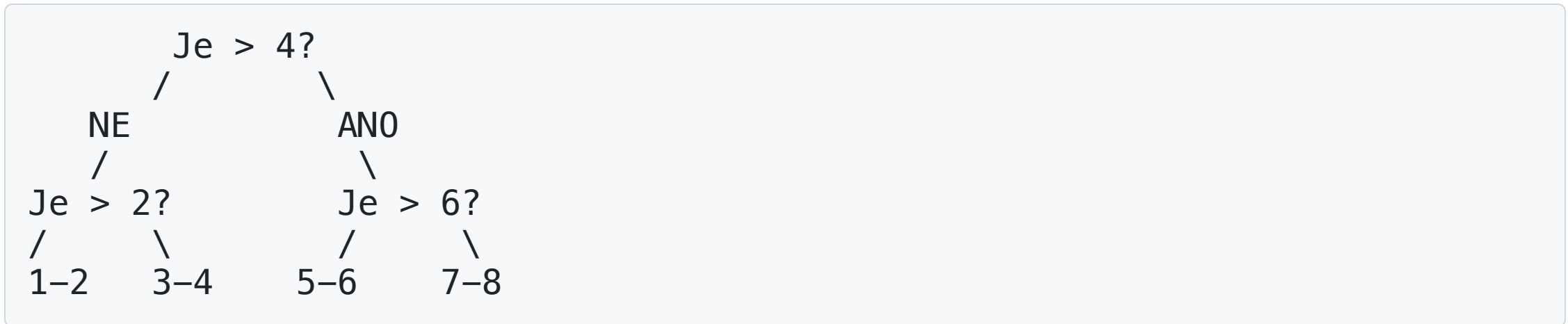
## Výběr oběda



*Každá cesta od kořene k listu představuje jednu možnost*

# Strom pro hádání čísla

## Rozsah 1-8



### Vlastnosti:

- Každá otázka pŕlí interval
- Maximálně 3 otázky
- Všechny cesty podobně dlouhé

# Výška stromu

## Důležitý koncept

### Definice:

- Nejdelší cesta od kořene k listu
- = Maximální počet otázek

### Příklady:

- 8 možností → výška 3
- 16 možností → výška 4
- 32 možností → výška 5

*Výška =  $\log_2(\text{počet možností})$*

# Optimální strom

## Charakteristiky

### 1. Je vyvážený

- Podobná délka všech cest

### 2. Efektivně dělí možnosti

- Každá otázka pólí množinu

### 3. Minimalizuje výšku

- Nejkratší možná cesta ke všem listům



# Tvorba stromu

## Praktický postup

1. Spočítej počet možností
2. Urči potřebnou výšku stromu
3. Identifikuj otázky půlící množinu
4. Postupně buduj strom odshora

### Tip:

*Začni otázkou, která rozdělí množinu na co nejpodobnější části*

# Použití stromů

## Praktické aplikace

1. Diagnostika problémů
2. Klasifikace dat
3. Rozhodovací algoritmy
4. Expertní systémy
5. Strojové učení

*Stromy pomáhají systematizovat rozhodování*

# Shrnutí

## Klíčové koncepty

1. Strom vizualizuje rozhodovací proces
2. Výška stromu určuje počet otázek
3. Optimální strom je vyvážený
4. Každá otázka by měla půlit množinu
5. Praktické využití v mnoha oblastech

*V další části si ukážeme, jak měřit množství informace v rozhodovacím procesu...*

# Měření informace

**Jak kvantifikovat množství informace?**

# Od stromů k měření

## Co nám stromy prozradily?

### Pozorování:

- 1 otázka → rozliší 2 možnosti
- 2 otázky → rozliší 4 možnosti
- 3 otázky → rozliší 8 možností
- 4 otázky → rozliší 16 možností

*Vidíte vzorec?*

# Bit jako jednotka informace

## Základní definice

**1 bit = informace, která:**

- Rozliší mezi 2 možnostmi
- Zodpoví 1 otázku typu ANO/NE
- Vyloučí polovinu možností

**Příklady 1 bitu:**

- Hod mincí (panna/orel)
- Sudé/liché číslo
- Zapnuto/vypnuto

# Vztah otázek a bitů

## Obecný vzorec

n otázek = n bitů =  $2^n$  možností

## Příklady:

- 2 bity → 4 možnosti
- 3 bity → 8 možností
- 4 bity → 16 možností

*Každý další bit zdvojnásobí počet rozlišitelných možností*

# Výpočet potřebných bitů

Pro daný počet možností

Vzorec:

Počet bitů =  $\log_2(\text{počet možností})$

Příklady:

- 8 možností → 3 bity
- 16 možností → 4 bity
- 1000 možností → 10 bitů

*Zaokrouhlujeme nahoru!*



# Příklad: Hrací karty

## Kolik bitů potřebujeme?

### Analýza:

1. Barva (4 možnosti):

- 2 bity

2. Hodnota (13 možností):

- 4 bity

3. Celá karta:

- 6 bitů celkem

*Dokážeme zakódovat všech 52 karet*

# Přechod od bitů k neurčitosti

## Jak bity souvisí s nejistotou?

### Bity jako základní jednotka informace

- Bit je základní jednotka informace, kterou potřebujeme k odlišení mezi dvěma možnostmi.
- Každý bit **vylučuje určitou část možností, čímž snižuje naši nejistotu.**

### Příklady:

- Kódování hracích karet: potřebujeme 6 bitů k rozlišení všech karet.
- Čím více možností musíme rozlišit, tím více bitů potřebujeme.

*Ale co kdyby možnosti nebyly stejně pravděpodobné?*

# Neurčitost a pravděpodobnost

## Jaký je vztah mezi pravděpodobností a neurčitostí?

### Když jsou možnosti **stejně pravděpodobné**:

- Hod kostkou (šest možností) → Každé číslo má stejnou šanci:  $1/6$ .
- Potřebujeme určitý počet bitů, abychom určili konkrétní výsledek.

### Když jsou možnosti **různě pravděpodobné**:

- Například při předpovědi počasí: Když hrajem hru myslí si zvíře se studenty Přírodovědecké fakulty:-)
- **Méně pravděpodobné události** přinášejí více informace, pokud nastanou, protože **odstraňují více nejistoty**.

*Jak můžeme kvantifikovat tuto neurčitost? Odpovědí je entropie.*

# Entropie jako míra neurčitosti

## Jak přesně měříme neurčitost?

### Co je entropie?

- Entropie je **měřítka nejistoty** nebo neurčitosti o systému před tím, než obdržíme konkrétní informaci.
- Půjdeš večer? Ano, Ne, Možná
- Vyjadřuje, kolik bitů potřebujeme k vyjádření všech možných stavů systému.

### Příklad:

- Hod mincí: **Entropie = 1 bit** (panna nebo orel, obě stejně pravděpodobné).
- Hod kostkou: **Entropie = 2,58 bitů** (šest možností, všechny stejně pravděpodobné).
- Čím více možností a čím jsou tyto možnosti méně jisté, tím **vyšší entropie**.

*Pojďme se podívat na konkrétní příklad výpočtu entropie pro reálný systém*

# Informační entropie

## Míra neurčitosti systému

Vzorec:

$$H = -\sum P(s) \times \log_2 P(s)$$

### Příklad: Hod mincí

- $P(\text{panna}) = 1/2$
- $P(\text{orel}) = 1/2$
- $H = 1 \text{ bit}$

*Dokonalá neurčitost = maximální entropie*

# Rovnoměrné rozdělení

**Když jsou možnosti stejně pravděpodobné**

**Příklad: Hod kostkou**

- Šestka: 16.7% šance
- Ostatní čísla: každé 16.7%

**Důsledek:**

- Informace "Padla šestka"  $\neq$  "Nepadla šestka"
- Méně pravděpodobná událost = více informace

# Nerovnoměrné rozdělení

## Když jsou možnosti stejně pravděpodobné

- Na vánoční besídku roku 2013 přijede poslanec (nebo poslankyně) z Karlovarského kraje.
- Chceme zjistit, **kolik informace získáme** o stranické příslušnosti tohoto poslance.

[zdroj příkladu](#)

## Otázka

Kolik bitů informace se dozvíme z této zprávy?

# Krok 1: Výpočet entropie před zprávou

## 1. Definice entropie:

- Entropie (H) měří nejistotu. Čím vyšší je entropie, tím větší je nejistota.
- Vzorec pro entropii:

$$H = - \sum p(x) \cdot \log_2(p(x))$$

## 2. Rozdělení poslanců:

- Celkový počet poslanců je 200, jejich rozdělení mezi strany je nerovnoměrné: 33, 50, 14, 14, 47, 26, 16.

## 3. Výpočet:

- Pravděpodobnost každé strany se vypočítá jako podíl počtu poslanců na celkovém počtu.



## Krok 2: Výsledek entropie před zprávou

- Dosazením pravděpodobností do vzorce:

$$H_{\text{pred}} = - \left( \frac{33}{200} \cdot \log_2 \frac{33}{200} + \frac{50}{200} \cdot \log_2 \frac{50}{200} + \dots \right)$$

- Počet výpočtů je omezen na sedm různých pravděpodobností (pro každou stranu).
- Výsledek:

$$H_{\text{pred}} \approx 2,63 \text{ b}$$

Entropie před přijetím zprávy je 2,63 b, což znamená vyšší nejistotu ohledně stranické příslušnosti.

# Krok 3: Výpočet entropie po zprávě

## 1. Redukce možností:

- Víme, že poslanec zastupuje Karlovarský kraj.
- V této skupině je pouze 5 poslanců s četnostmi podle stran: 2, 2, 1.

## 2. Výpočet pravděpodobností:

- Pravděpodobnosti jsou nyní:  $(\frac{2}{5}, \frac{2}{5}, \frac{1}{5})$ .

## 3. Entropie po zprávě:

$$H_{po} = - \left( \frac{2}{5} \cdot \log_2 \frac{2}{5} + \frac{2}{5} \cdot \log_2 \frac{2}{5} + \frac{1}{5} \cdot \log_2 \frac{1}{5} \right)$$

- Výsledek:

$$H_{po} \approx 1,52 \text{ b}$$

# Konečný výsledek: Množství informace

## 1. Rozdíl entropií:

- Množství informace získané z nové zprávy je rozdíl mezi entropií před a po:

$$I = H_{\text{pred}} - H_{\text{po}} = 2,63 - 1,52 = 1,11 \text{ b}$$

## 2. Interpretace:

- Zjištění stranické příslušnosti poslance nám poskytlo **1,11 bitu informace**.
- Tento výsledek ukazuje množství odstraněné nejistoty díky nové zprávě.

**Závěr:** Čím je informace specifická nebo nepravděpodobná, tím více bitů informací nám poskytuje.

# Praktické využití

## Kde se s bity setkáme?

### 1. Digitální úložiště:

- 1 bajt = 8 bitů
- 1 KB = 1024 bajtů

### 2. Datové přenosy:

- Rychlost v bitech za sekundu
- Např. 100 Mb/s internet

### 3. Kompresce dat:

- Odstranění nadbytečných bitů
- Efektivnější kódování

# Shrnutí

## Klíčové koncepty

1. Bit je základní jednotka informace
2.  $n$  bitů rozliší  $2^n$  možností
3. Počet bitů =  $\log_2(\text{možností})$
4. Pravděpodobnost ovlivňuje množství informace
5. Praktické využití v IT

*V další části si ukážeme rozdíl mezi daty a informací...*

# Data vs. Informace

**Kdy se data stávají informací?**

# Co jsou data?

## Základní charakteristika

### Data jsou:

- Posloupnost znaků nebo signálů
- Bez kontextu nemají význam
- Surový materiál pro informaci

### Příklad:

01000001

*Co to znamená?*

# Různé interpretace dat

## Stejná data, různý význam

Data: "01000001"

- Jako číslo: 65
- Jako ASCII: písmeno "A"
- Jako teplota: 65°C
- Jako věk: 65 let

*Význam závisí na interpretaci*



# Co je informace?

## Data + Kontext

### Informace vzniká když:

1. Známe formát dat
2. Víme, co reprezentují
3. Rozumíme kontextu
4. Můžeme data interpretovat

### Příklad:

65 + °C = informace o teplotě varu vody

# Kódování informace

## Různá data, stejná informace

### Číslo 42 jako:

- Binárně: 00101010
- Desítkově: 42
- Římsky: XLII
- Text: "čtyřicet dva"
- Hex: 2A

*Různé reprezentace, stejný význam*

# Kontext mění význam

## Důležitost interpretace

### Příklad: "110"

- V binární soustavě: 6
- V desítkové soustavě: 110
- Jako číslo pokoje: místnost 110
- Jako teplota: 110°C
- Jako čas: 1:10

*Stejná data → různé informace*

# Ztráta kontextu

## Proč je důležitá dokumentace

### Co znamenají tato data?

23, 45, 67, 89

### Možné interpretace:

- Seznam čísel?
- Souřadnice?
- Měření?
- Časové údaje?

*Bez kontextu jsou data nepoužitelná*

# Shrnutí

## Klíčové poznatky

1. Data jsou nosičem informace
2. Informace vzniká interpretací dat
3. Stejná data → různé informace
4. Různá data → stejná informace
5. Kontext je klíčový

**Děkuji za pozornost**