

# Genetické algoritmy a jejich aplikace



PB016 Úvod do umělé inteligence

Martin Frodl

1. 12. 2009

# Obecné schéma genetického algoritmu

1. náhodně vygeneruj populaci řešení
2. pro každé řešení vypočítej **fitness**, tj. hodnotu nějaké ohodnocovací funkce
3. na základě fitness vyber dvojice ke zkřížení a vygeneruj jejich potomstvo
4. vytvoř novou populaci z generací rodičů a potomků
5. opět vypočítej fitness všech jedinců
6. pokud bylo nalezeno uspokojivé řešení (tj. s dostatečně vysokou fitness),
7. vrať výslednou populaci, jinak opakuj od bodu 3

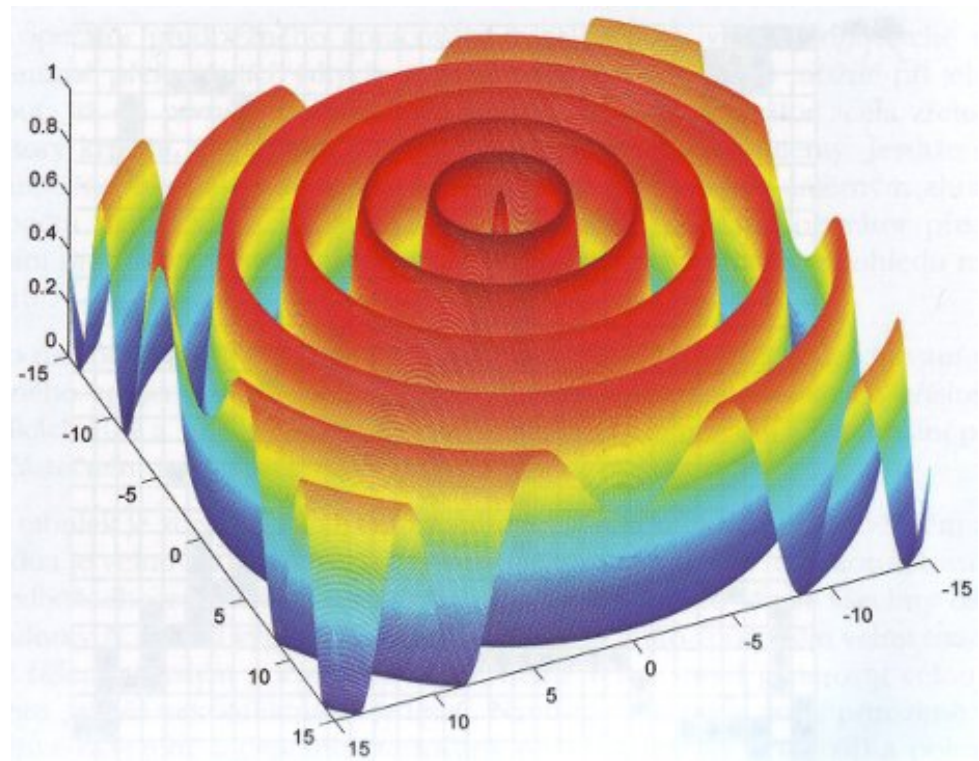
# Příklad: hledání extrémů funkce

**Úkol:** najděte globální maximum funkce  $f$

(omezme se na množinu  $[-100,100] \times [-100,100]$  a požadujeme přesnost na 4 desetinná místa)

$f$  je příkladem funkce, u které klasický „horolezecký“ přístup selhává. Naopak genetický algoritmus řešení s více než 90% pravděpodobností nalezne.

$$f(x, y) = \frac{1 - \sin^2 \sqrt{x^2 + y^2}}{1 + 0,001 \cdot (x^2 + y^2)}$$



# Příklad: hledání extrémů funkce

- **Vhodná reprezentace řešení.** Množinu si rozdělíme na  $(200 \cdot 10000)^2 = 4 \cdot 10^{12}$  dílků, kde každý představuje možné řešení. Řešení můžeme jednoznačně reprezentovat 42bitovou posloupností, protože

$$2^{41} \approx 2,20 \cdot 10^{12} < 4 \cdot 10^{12} < 4,40 \cdot 10^{12} \approx 2^{42}$$

Tato posloupnost se označuje jako **chromosom**.

(0,1,0,0,1,0,0,1,1,0,1,0,0,1,0,0,1,1,0,1,0,0,1,0,0,1,0,1,1,0,1,1,0,0,1,0,0,1,0,1,1,0)

- Pro každý chromosom existuje jednoznačná transformace do souřadnic  $(x,y)$ . Lze pro něj tedy vypočítat hodnotu  $f(x,y)$ , která je v tomto případě zároveň i ohodnocovací funkcí – fitness.
- Při vytváření potomstva může navíc dojít k *mutaci* (bodová změna na chromosomu) a *crossing-overu* (výměna odpovídajících úseků rodičovských chromosomů)

# Příklad: hledání extrémů funkce

## Crossing over

Rodičovské chromosomy:

(0,1,0,0,1,0,0,1,1,0,1,0,0,1,0,0,1,1,0,1,0,0,1,0,0,1,0,1,1,0,1,1,0,0,1,0,0,1,0,1,1,0)

(1,0,1,0,0,1,1,0,1,1,0,1,1,0,0,1,0,0,1,0,1,0,0,1,1,0,1,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1)

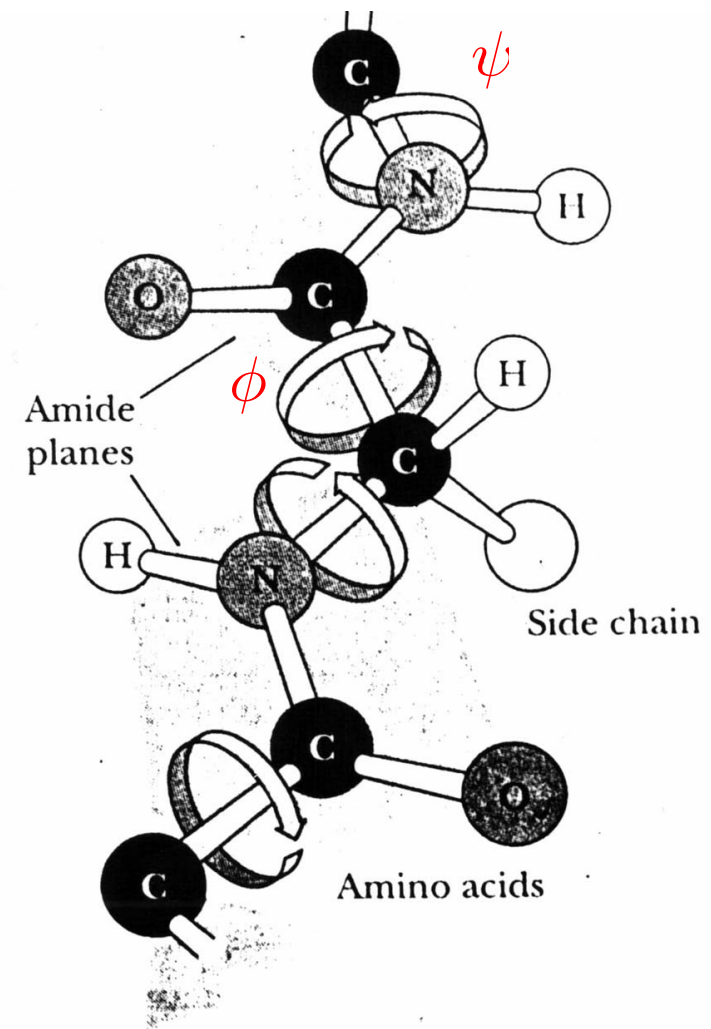
Chromosomy potomků:

(0,1,0,0,1,0,0,1,1,0,1,0,0,1,0,0,1,1,0,1,0,0,1,0,0,1,0,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1)

(1,0,1,0,0,1,1,0,1,1,0,1,1,0,0,1,0,0,1,0,1,0,0,1,1,0,1,1,1,0,1,1,0,0,1,0,0,1,0,1,1,0)

# Aplikace: předpověď 3D struktury proteinů

- **Úkol:** na základě sekvence aminokyselin vypočítej tvar proteinu s nejnižší energií
- **Naivní reprezentace řešení:** vektor kartézských souřadnic každého atomu
- **Lepší reprezentace:** vektor tzv. **torzních úhlů**  $\phi$ ,  $\psi$  mezi jednotlivými aminokyselinovými zbytky
- **Ohodnocovací funkce:** celková energie molekuly (čím nižší, tím lepší řešení)



# Aplikace: předpověď chování cen akcií

- **Úkol:** odhadni hodnotu akcií nějaké firmy na 12 týdnů dopředu
- **Vstupní data:** *technické a fundamentální* informace o daných akciích
- **Reprezentace řešení:** pravidlo (pravidla) rozhodující o akci na základě vstupních informací
  - **Pravidlo 1:** Pokud  $P/E > 30$ , pak prodej.
  - **Pravidlo 2:** Pokud  $P/E < 40$  a tempo růstu  $> 40\%$ , pak nakup
- Chromosom bude místo binární posloupnosti reprezentován syntaktickým stromem
- **Ohodnocovací funkce:** čistý výnos v případě aplikace daného pravidla

# Další oblasti použití genetických algoritmů

- **Design** materiálů a tvarů **závodních automobilů** i běžných dopravních prostředků
- Strukturní a funkcionální design staveb, strojů, výrobních jednotek atd.
- Směrování paketů v telekomunikačních sítích (**routing**)
- **Šifrování a dešifrování**
- Návrh **syntetických molekul** (léčiva, materiály atd.)