

EVOLUČNÁ ROBOTIKA

2. část'

A thick, horizontal yellow brushstroke with a textured, painterly appearance, spanning most of the width of the slide.

František Mráz, KSVI MFF UK

EVOLUČNÉ A NEURÓNOVÉ TECHNIKY



- **Základné okruhy**
 - ◆ **evolučné algoritmy – genetické algoritmy**
 - ◆ **neurónové siete**
 - ◆ **kombinácie evolučných algoritmov a neurónových sietí**
 - ◆ **genetické programovanie**

- **Evolúcia riadiacich systémov pre roboty už bola aplikovaná na**
 - ◆ riadenie neurónovými sieťami
 - ◆ parametre pevného riadiaceho programu
 - ◆ samotný počítačový program

GENETICKÉ ALGORITMY (GA)

- ⚙ Holland 60. roky 20. stor.
- **Populácia** umelých chromozómov sa cyklicky podrobuje
 - ◆ selektívnej reprodukcií preferujúcej výkonnejších jedincov a
 - ◆ náhodným zmenám
- **Umelý chromozóm** (genotyp) = reťazec symbolov kódujúci vlastnosti jedinca (fenotyp)
 - ◆ napr. binárna hodnota premennej alebo postupnosť hodnôt premenných
 - ◆ mnoho typov kódovaní
 - ◆ binárne, Greyov kód, reálne hodnoty
 - ◆ abecedy – napr. binárna, ternárna, ...

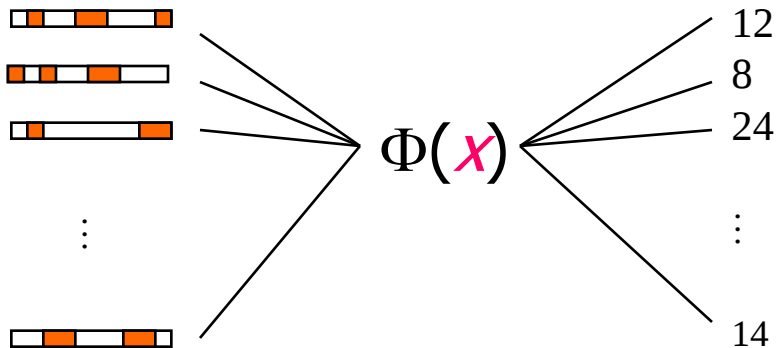
GA – fitness funkcia

- **Fitness funkcia (účelová, cieľová funkcia)** – kritérium výkonnosti, zobrazenie: genotyp $\rightarrow$ reálne (celé) číslo
 - ♦ čím vyššia hodnota, tým je jedinec lepší

$$\Phi: \text{genotyp} \rightarrow R$$

populácia

fitness



Genetické algoritmy

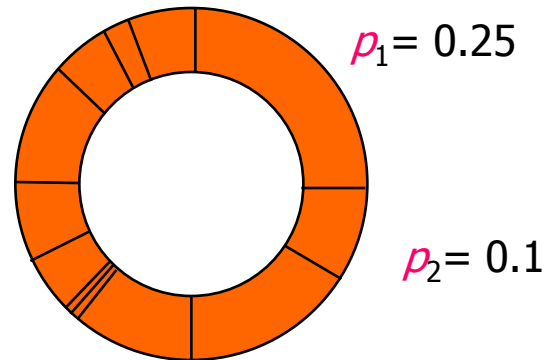
- **Jednoduchý genetický algoritmus** (Goldberg 1989)
 - ♦ vytvor populáciu N náhodne vygenerovaných chromozómov $x_1, x_2, \dots, x_N$
 - ♦ opakuj
 - ♦ dekóduj všetky chromozómy a spočítaj ich fitness $f_i = \Phi(x_i)$
 - ♦ vytvor novú populáciu selektívnou reprodukciou
 - ♦ rekombinuj chromozómy – kríženie
 - ♦ mutuj chromozómy
 - ♦ skonči, keď sa objaví hľadaný jedinec, alebo fitness najlepšieho nerastie

Jednoduchá selekcia

- Zo starej populácie vytvárame novú kopírovaním chromozómov tak, že čím lepší jedinec, tým viacej jeho kópií sa môže objaviť v novej populácii
- reprodukcia **ruletou**:
 - ◆ každý jedinec dostane na kole rulety priehradku veľkosti p_i úmernej hodnote jeho fitness funkcie – to bude pravdepodobnosť, že bude reprodukováný (predpokladáme nezápornú fitness funkciu)

$$p_i = \frac{f_i}{\sum_{j=1}^n f_j}$$

- ◆ ruletou sa otočí N krát



Varianty selekcie 1

■ Problémy jednoduchkej selekcie

a) keď všetci jedinci majú podobnú fitness

náhodné prehľadávanie s genetickým posunom

b) keď jeden až dvaja jedinci majú fitness o mnoho vyššiu než zvyšok populácie

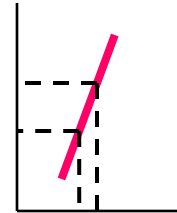
takmer všetci jedinci v novej generácii budú kópiou toho istého jedinca, predčasná konvergencia

■ Riešenie:

1. **škálovanie [scaling]** (typicky lineárna transformácia)

pre a) zväčšiť rozdiely, pre b) zmenšiť

2. **selekcia podľa poradia [rank based]** – jedinci sa zoradia podľa fitness a pravdepodobnosť reprodukcie je úmerná poradiu jedinca, nie jeho fitness



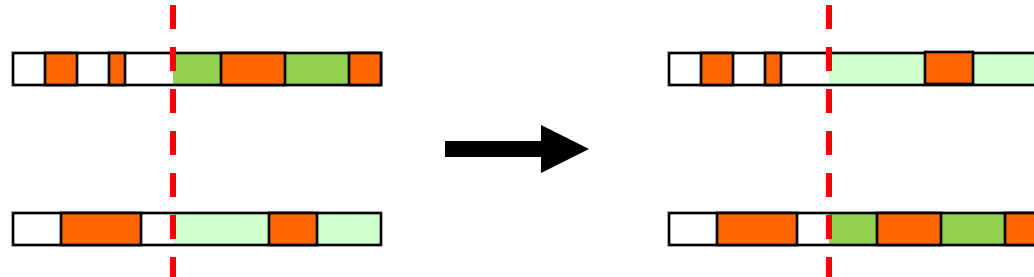
Varianty selekcie 2

3. **S orezávaním [truncation]** – jedincov usporiadame podľa veľkosti fitness, M najlepších jedincov okopírujeme O krát tak, že $N = M \times O$
 4. **Turnajom [tournament]** – náhodne sa vyberú 2 jedinci a vygeneruje sa náhodné číslo $r \in \langle 0,1 \rangle$, ak je $r < T$, kde $T \in \langle 0,1 \rangle$ je preddefinovaný parameter, tak bude okopírovaný jedinec s vyššou fitness inak ten druhý
- Pri ďalších genetických operáciách sa môže doteraz najlepší jedinec stratiť. Pomoc: **elitizmus [elitism]** – S najlepších jedincov je bez zmeny okopírovaných do novej generácie

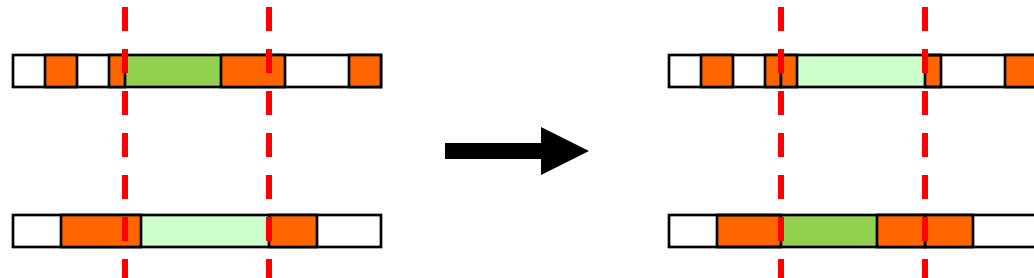
Kríženie

- Jedinci sú náhodne spárovaní a každý pár je s danou pravdepodobnosťou skrížený

Jednobodové
kríženie



Viacbodové
kríženie



Nedeterminizmus

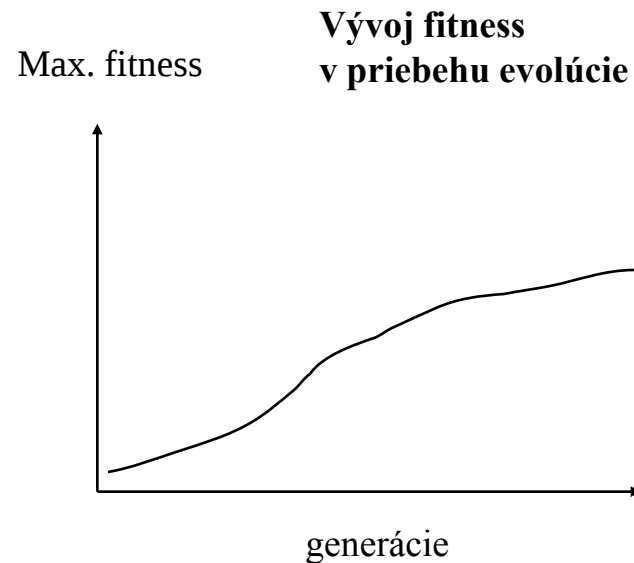
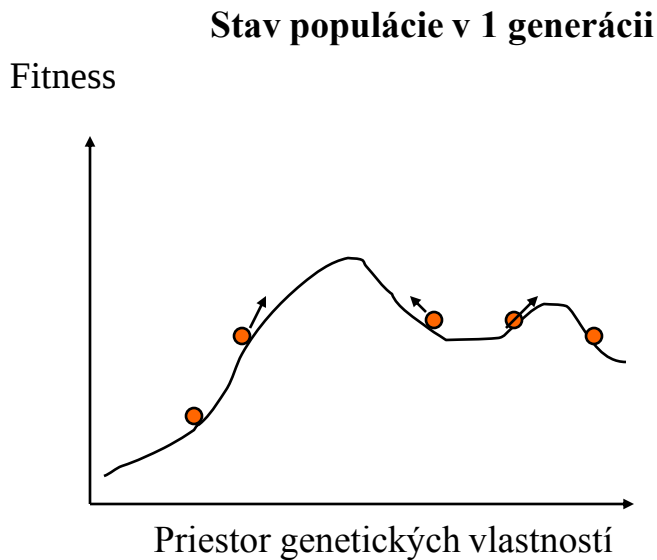


- celý GA je nedeterministický, používa náhodné veličiny
- typicky sa sleduje priemerná a maximálna fitness, GA sa zastaví, keď je dosiahnutá dopredu zadaná hodnota fitness, alebo zadaný počet generácií, alebo sa fitness v priebehu niekoľkých generácií nemení.
- celý GA sa spúšťa opakovane s rovnakými alebo s pozmenenými parametrami
- výsledkom je najlepší jedinec vybraný z finálnych generácií zo všetkých behov GA

Fitness krajina

– „fitness landscape“

- Fitness si môžeme predstaviť ako mnohorozmernú nadrovinu udávajúcu hodnotu fitness vo všetkých možných hodnotách genotypu – fitness krajina s kopcami a údoliami; pre chromozóm dĺžky 1 je to funkcia jednej premennej



Schémy

- **Chromozóm** = reťazec dĺžky m nad k -prvkovou abecedou Σ
- **Schéma** = reťazec dĺžky m nad $(k+1)$ -prvkovou abecedou $\Sigma \cup \{*\}$

- ♦ schéma s r hviezdikami reprezentuje k^r reťazcov
- ♦ každý reťazec dĺžky m je reprezentovaný $(k+1)^r$ schémami
- ♦ **rád** schémy S : $o(S)$ = počet pevných symbolov (nie $*$)
- ♦ (definičná) **dĺžka** schémy S : $\delta(S)$ = vzdialenosť medzi prvou a poslednou pevnou pozíciou v schéme
- ♦ Príklad:

	rád $o(S)$	definičná dĺžka $\delta(S)$
$S_1 = (** * 001 * 110)$	6	6
$S_2 = (**** 00 ** 0 *)$	3	4
$S_3 = (11101 ** 001)$	8	9

Schémy

- N = počet jedincov v populácii
- $\xi(S, t)$ = počet reťazcov v populácii v čase t , ktoré sú reprezentované schémou S
- $\Phi(S, t)$ = fitness schémy v čase t = priemerná fitness reťazcov reprezentovaných schémou S

$$\Phi(S, t) = \frac{1}{\xi(S, t)} \sum_{j=1}^{\xi(S, t)} \Phi(x_{i_j}),$$

kde $\{x_{i_1}, \dots, x_{i_{\xi(S, t)}}\}$ sú reťazce z $P(t)$ pokryté schémou S

- pravdepodobnosť, že $x \in P(t)$ bude vybrané selekciou do $P(t+1)$ je

$$p = \frac{\Phi(x)}{F(t)}, \quad \text{kde } F(t) = \sum_{j=1}^N \Phi(x_j) \text{ je fitness populácie}$$

Schémy a selekcia

- Po selekcii očakávame $\xi(S, t+1)$ reťazcov pokrytých schémou S

$$\xi(S, t+1) = \xi(S, t) \cdot pop_vel \cdot \frac{\Phi(S, t)}{F(t)} = \xi(S, t) \cdot \frac{\Phi(S, t)}{\bar{F}(t)}$$

kde $\bar{F}(t)$ je priemerná fitness populácie $P(t)$

- teda "nadpriemerná" schéma pokryje v nasledujúcej generácii viacej reťazcov – ak by platilo pre prvých t generácií

$$\Phi(S, t) \geq \bar{F}(t) \cdot (1 + \varepsilon) \text{ pre nejaké } \varepsilon > 0$$

- tak $\xi(S, t) \geq \xi(S, 0) \cdot (1 + \varepsilon)^t$ tj. počet reťazcov reprezentovaných "nadpriemernou" schémou rastie geometrickou postupnosťou

Schémy a kríženie

- Pravdepodobnosť, že schéma bude zmenená (zničená) krížением

$$\frac{\delta(S)}{m-1}$$



- teda pravdepodobnosť zachovania príslušnosti k schéme bude

$$1 - \frac{\delta(S)}{m-1}$$

- ◆ ALE:

- ◆ kríženie nastáva iba s pravdepodobnosťou p_k
- ◆ aj keď je bod kríženia "vo vnútri schémy", tak schéma môže prežiť

- teda pravdepodobnosť zachovania schémy je

$$p_z \geq 1 - p_k \cdot \frac{\delta(S)}{m-1}$$

- selekcia + kríženie:

$$\xi(S, t+1) \geq \xi(S, t) \cdot \frac{\Phi(S, t)}{F(t)} \cdot \left[1 - p_k \frac{\delta(S)}{m-1} \right]$$

Schémy a mutácia

- Ak mutácia zmení v reťazci jeden z pevných symbolov schémy, tak reťazec nebude schémou pokrytý; pravdepodobnosť, že mutácia nezmení žiaden z pevných symbolov schémy je

$$(1 - p_m)^{o(S)} \approx 1 - o(S) \cdot p_m \text{ pretože } p_m \ll 1$$

- selekcia + kríženie + mutácia:

$$\begin{aligned} \xi(S, t+1) &\geq \xi(S, t) \cdot \frac{\Phi(S, t)}{\bar{F}(t)} \cdot \left[1 - p_k \frac{\delta(S)}{m-1} \right] \cdot [1 - o(S) \cdot p_m] \\ &\geq \xi(S, t) \cdot \frac{\Phi(S, t)}{\bar{F}(t)} \cdot \left[1 - p_k \frac{\delta(S)}{m-1} - o(S) \cdot p_m \right] \end{aligned}$$

- schéma S "prežíva" dobre, ak je krátka ($\delta(S)$) a malého rádu ($o(S)$)

Veta o schémach

- **Veta o schémach (Schema Theorem):** Schémy, ktoré sú krátke, majú nízky rád a sú nadpriemerné, pokrývajú exponenciálne rastúci počet reťazcov v po sebe idúcich generáciách genetického algoritmu
- **Hypotéza o stavebných blokoch (Building block Hypothesis):** Genetický algoritmus hľadá riešenie blízke optimu reťazením krátkych schém s nízkym rádom nazývaných stavebné bloky.
 - ◆ nie je to veta!
 - ◆ niekedy stavebné bloky znemožnia nájdenie optima
 - ◆ Holland spočítal, že za určitých podmienok môže genetický algoritmus v jednom kroku spracovať až N^3 krátkych schém, ale funguje to iba ak je možné „dobré“ vlastnosti kódovať takýmito schémami – variácie GA

Sú GA dobrou simuláciou biologickej evolúcie?

- Evolúcia nekončí a nezačína vždy od nuly (bootstrapping)
 - ◆ **SAGA** – Species Adaptation Genetic Algorithm: premenná dĺžka genotypu, značne skonvergovaná populácia sa mení postupnými mutáciami, nie krížením
 - ◆ **Neutrálne siete** – drobné náhodné zmeny v genotype nemusia meniť fitness, ale môžu populáciu posunúť inam
 - ◆ fitness krajina môže byť sústava prepojených neutrálnych sietí s rôzne vysokými hodnotami fitness
 - ◆ „prázdne“ úseky DNA
- Ďalšie variácie GA – spojenie s „dospievaním“ a učením – fitness sa mení v priebehu života jedinca
- **ČO SA NEMENÍ:** a) genetická reprezentácia jedinca
b) autonómna interakcia s prostredím

Vylepšenia genetických algoritmov

■ Delenie na druhy

- ◆ na genotype sa definuje vzdialenosť = miera podobnosti jedincov
- ◆ 1 druh = množina podobných jedincov
- ◆ Špeciálny typ kríženia – iba v rámci druhu (predpoklad, že krížením podobných jedincov vznikne podobný, teda funkčný jedinec)
- ◆ Ochrana nových druhov
 - ◆ Po obmedzenú dobu od vzniku nového druhu jedinci pri selekcii sú vyberaní iba na základe fitness pre jedincov ich vlastného druhu

■ Premennivá veľkosť populácie

- ◆ Optimálna veľkosť populácie sa ťažko určuje
- ◆ Starnutie
 - ◆ pri vzniku nového jedinca je mu (podľa jeho fitness) nastavený max. vek = max. počet generácií, než bude jedinec z populácie vyradený – **veľmi ťažko sa nastavujú parametre!**

Vylepšenia genetických algoritmov



■ Kódovanie

◆ Konečná doména

- ◆ pevný počet bitov – čo ak počet prvkov domény nie je mocnina 2?

◆ Celé čísla

- ◆ binárne
- ◆ Grayov kód

◆ Reálne čísla

- ◆ s obmedzeným rozsahom – binárne kódovanie rovnomerne rozdelených deliacich bodov
- ◆ čísla s pevným počtom desatinných miest
- ◆ čísla s plávajúcou desatinnou čiarkou – ako robiť kríženie a mutácie?