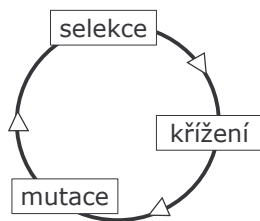


Simulovaná evoluce II.

- princip a analogie
- základní typy simulované evoluce
- kódování instancí
- evoluční operátory
- selekce a selekční tlak
- řízení generací
- podmínky ukončení
- práce s omezujícími podmínkami
- schémata
- teorie stavebních bloků
- zavádějící (klamně) funkce
- fast messy GA
- Bayesovská optimalizace
- paralelizace

Schémata a teorie stavebních bloků

GA jako zpětnovazební dynamický systém



Jaká informace obíhá ve smyčce?

Jaké je dynamické chování smyčky?

Schémata

Myšlenka: řešení je sestaveno ze stavebních bloků

1	0	1	0	0	0	1
1	1	1	0	1	0	0

co mají tyto chromozómy společného?

*	*	1	*	*	0	*
1	*	1	0	*	0	*

← například } schémata
← nebo

délka=5 řád=4

Význam schémat

- každý genotyp o n genech obsahuje 2^n schémat
- každý výpočet zdatnosti získává údaj o průměrné zdatnosti všech schémat
- zpracování populace o velikosti M znamená zpracování $\Omega(M^3)$ schémat (implicitní paralelismus GA)

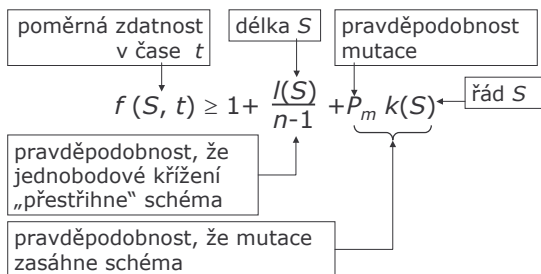
Dynamika schémat

- krátká schémata malého řádu přežívají ve smyčce snáze
- větší schémata potřebují lepší zdatnost

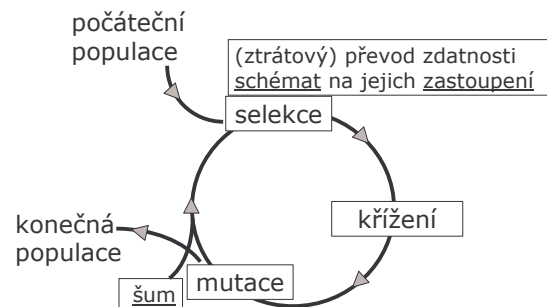
Věta o schématech

Selekce s pravděpodobností úměrnou zdatnosti, jednobodové křížení.

Zastoupení schématu S se stane početnější, když



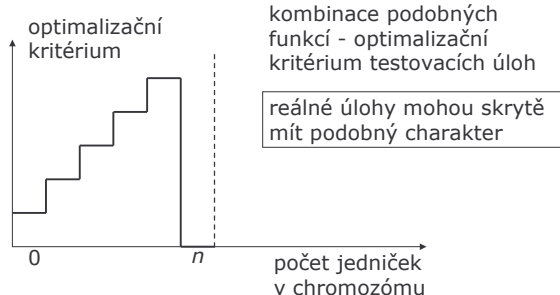
Schématata v GA cyklu



Linkage problem

- Životnost a vývoj schémat závisí na jejich délce, nikoliv pouze řádu
- Řešení:
 - dynamické přeuspořádání (linkage learning)
 - explicitní práce se stavebními bloky
- Platnost hypotézy stavebních bloků?

Zavádějící, klamné funkce (deceptive functions)



Kompetentní GA (Goldberg a další)

- klasický GA nezachází optimálně se schémata vyšších řádů, protože o nich neví; možné směry:
- pracovat se stavebními bloky explicitně jako s oddělenými fragmenty genetické informace ($\rightarrow$ messy GA)
- pracovat s pravděpodobnostním modelem vazeb mezi hodnotami proměnných, které vytvářejí stavební bloky ($\rightarrow$ bayesovská optimalizace)

Fast Messy GA

Kódování a hodnocení

místo hodnota místo hodnota

Fragment genetické informace

- někde nedospecifikovaný
- někde přespecifikovaný

- Referenční jedinec poskytne informaci pro nedospecifikovaná místa
- Bereme v úvahu první výskyt hodnoty pro dané místo genotypu

Struktura

tak, aby každá kombinace měla stejnou pravděpodobnost

opakovaná selekce a zkracování na cílovou délku (řád) k

operátory rozdělení a spojení (cut and splice)

začíná nová éra

$k=1$

generování stavebních bloků

filtrace stavebních bloků

rekombinace

nový referenční jedinec

$k=k+1$

Generování a filtrace stavebních bloků

- Cíl: získat kvalitní stavební bloky
- Metoda:
 - náhodně vygenerovat řetězce délky blízké délce chromozomu
 - selekcí zvýšit podíl kvalitních fragmentů
 - náhodně řetězce zkrátit
 - opakovat selekci a zkracování, až se dosáhne zadané délky (řádu)
 - pro hodnocení fragmentů řádu k je použit referenční jedinec zkonstruovaný pomocí fragmentů řádu $k+1$

Rekombinace

- Rozdělení (cut) obou rodičů v náhodných bodech
- Spojení (splice) rozdělených fragmentů
- Obě operace mají nezávisle nastavenou pravděpodobnost

Rozdělení a spojení

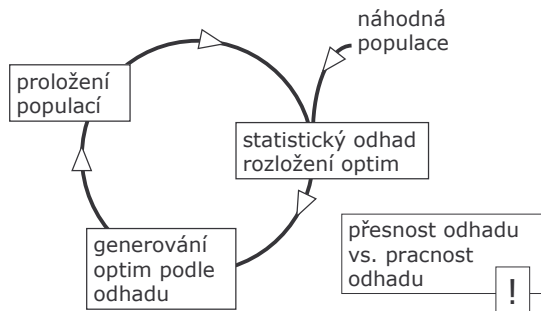
A B

C D

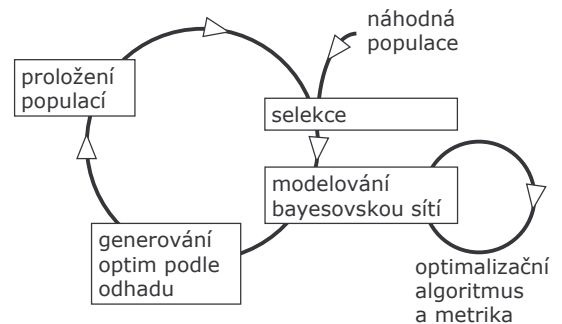
AB	
AC	
AD	
BA	
BC	
BD	
CA	
CB	
CD	
DA	
DB	
DC	

Algoritmy založené na statistických modelech závislosti

Princip

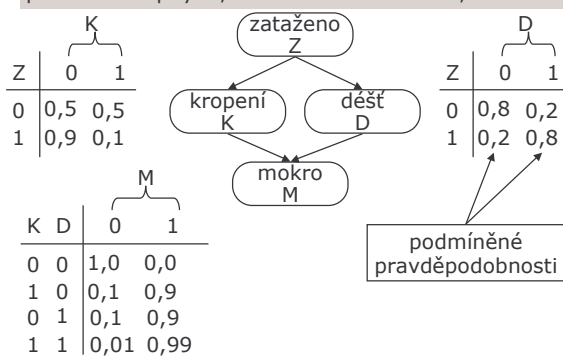


Bayesovská optimalizace (The Bayesian Optimization Algorithm, BOA)



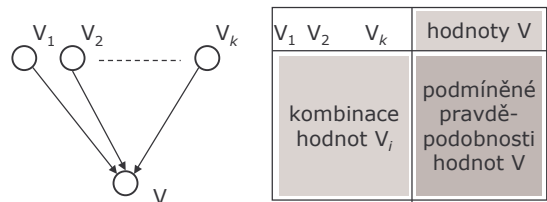
Bayesovská síť

podle K. Murphyho, U. of British Columbia, 1998



Bayesovská síť

Každá proměnná reprezentace modelována uzlem orientovaného acyklického grafu



Užití bayesovských sítí

- Najít nejpravděpodobnější vysvětlení pozorovaných pravděpodobností (UI)
 - je mokro; je to spíš deštěm nebo kropením?
- Zjistit pravděpodobnosti hodnot proměnných v závislosti na vstupních pravděpodobnostech
 - jaká je pravděpodobnost, že je mokro, není-li zataženo?

Optimalizace bayesovských sítí

- Metrika: jak dobře vystihuje měřená síť současnou populaci
- Algoritmus:
 - pouze nejlepší nebo jiná lokální heuristika
 - operace: přidání hrany, obrácení hrany, odebrání hrany

Paralelizace

Granularita

- Jedna aktuální generace v paralelním systému
 - sdílená paměť
 - výpočet zdatnosti, křížení, mutace se snadno škálují s počtem procesorů
 - minimální a maximální zdatnost, ranking: nutná komunikace, logaritmická složitost
- Jedna aktuální generace na procesor
 - nutná občasná výměna genetického materiálu (např. se sousedními procesory)
 - napomáhá zachovat diverzitu genetického materiálu
 - tzv. Island Parallel GA