

MI-SPOL-7

Experimentální vyhodnocení algoritmů, zejména randomizovaných.

Algoritmy

Aproximativní algoritmus:

- Každou instanci řeší v poly čase s relativní chybou ϵ

Randomizovaný algoritmus:

- Algorismus, který je založen na náhodné volbě
- Vstupy: vstup instance, náhodná čísla
- V průměrném případě statistická chyba
- **Monte Carlo:**
 - dosažený výsledek je náhodná proměnná, čas běhu pevný
 - *přesný čas na projetí okruhu formulí, ale nevím, jak se umístím*
 - náhodné procházky s omezeným počtem kroků
- **Las Vegas:**
 - vždy přesné řešení, čas běhu náhodná proměnná
 - *kasíno mě zaručeně obere, ale nevím kdy*
 - Quicksort
- **Výhody:**
 - strukturní jednoduchost
 - očekávaná kvalita výsledků může být lepší než zaručená kvalita aproximativních algoritmů
 - zlepšení kvality nezávislým opakováním
- **Nevýhody:**
 - Musím se vzdát buď deterministického času, nebo přesnosti řešení

Experiment:

- otázka
 - *co chci zjistit?*

- **plán experimentu**

- spuštění algoritmu:

$instance_1 \rightarrow výsledek_1$

$instance_2 \rightarrow výsledek_2$

...

- *otázka*: závislost výstupních veličin na vstupních veličinách

odpověď: kvantitativní výrok (vzorec) o té závislosti

- **provedení experimentu**

$veličina_a \ instance_1 \rightarrow veličina_b \ výsledku_1$

$veličina_a \ instance_2 \rightarrow veličina_b \ výsledku_2$

...

- **interpretace výsledků**

- odpověď různé kvality

- $veličina_b = F(veličina_a)$

Prakticky použitelná odpověď:

- vyžaduje zobecnění (extrapolaci)
- nezávislá na veličinách, které v ní nejsou zahrnuty (vyloučení ostatních vlivů, předpoklad jejich statistického rozložení)
- musí vypovídat o významné závislosti (testy hypotéz)

Co se hodnotí:

- **kvalita řešení**: absolutní (pokud znám správná řešení), relativní (pokud porovnávám dvě heuristiky)
- **výpočetní náročnost**: náročné na měření -- čas výpočtu zahrnuje všechny vlivy (i implementačně/platformě závislé)

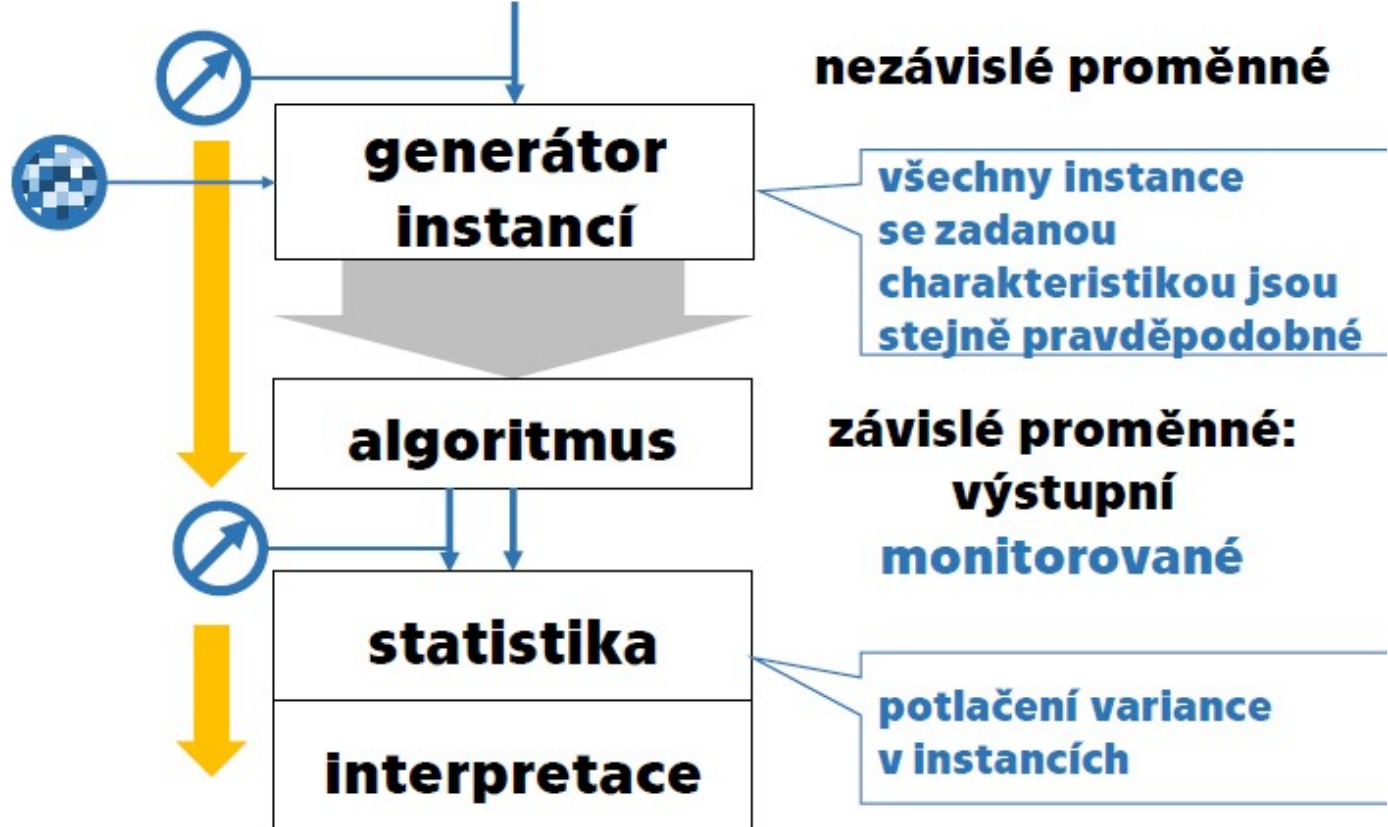
Měřítka výpočetní složitosti: počet testovaných stavů

Výběr instancí:

- **Náhodně generované**: škálovatelné, rovnoměrně rozdělené
- **Náhodně generované, s ohledem na experiment**: přesné pozorování chování algoritmu
- **Standardní benchmarky**: příklady z praxe, nerovnoměrné rozdělení

Obecné vyhodnocení algoritmu:

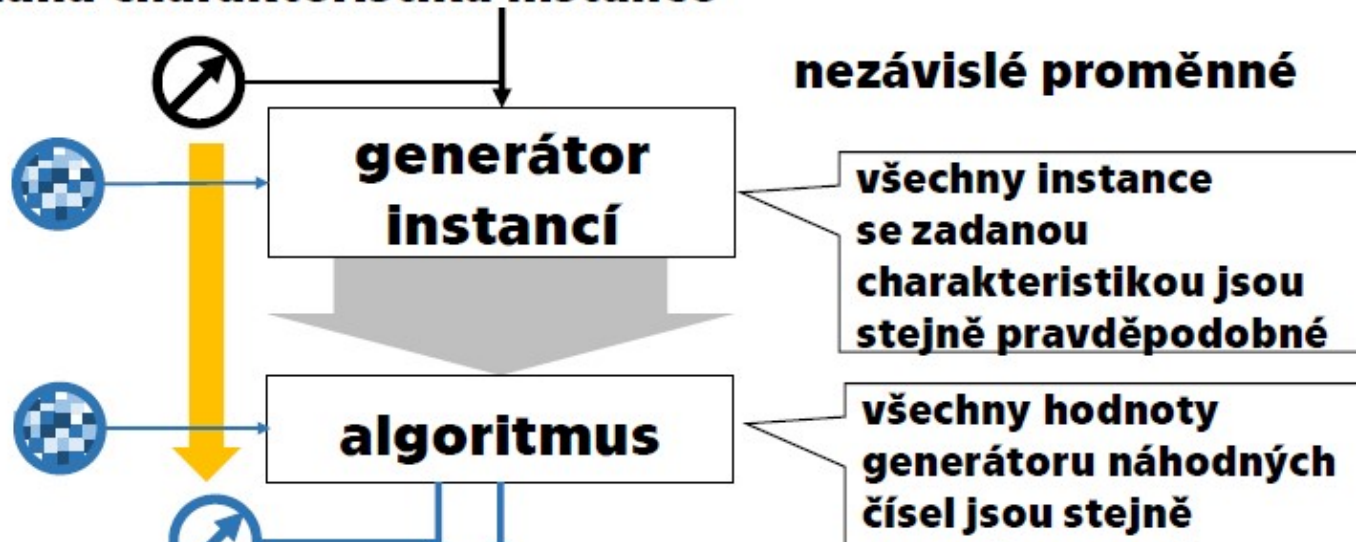


zadaná charakteristika instance

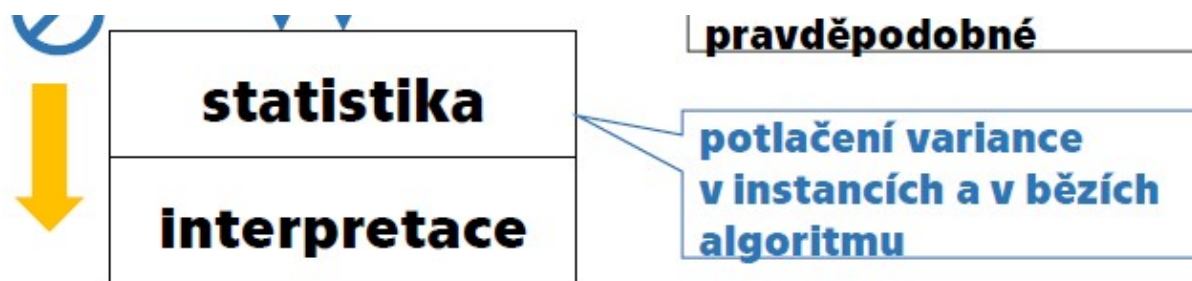
Statistika: jedna hodnota zadané charakteristiky, např. průměr, medián

Srovnání: A je systematicky lepší než B : dominance

Hlubší analýza: charakteristika instancí, kde je A lepší

Randomizovaný algoritmus**zadaná charakteristika instance**

7



Dva zdroje variance (generátor instancí a sám algoritmus) -- dva stupně zpracování

Statistické rozložení nemusí být stejné

Srovnání na jedné instanci:

- přes inicializaci RNG
- A systematicky lepší než B : dominance (na zkoušených instancích)

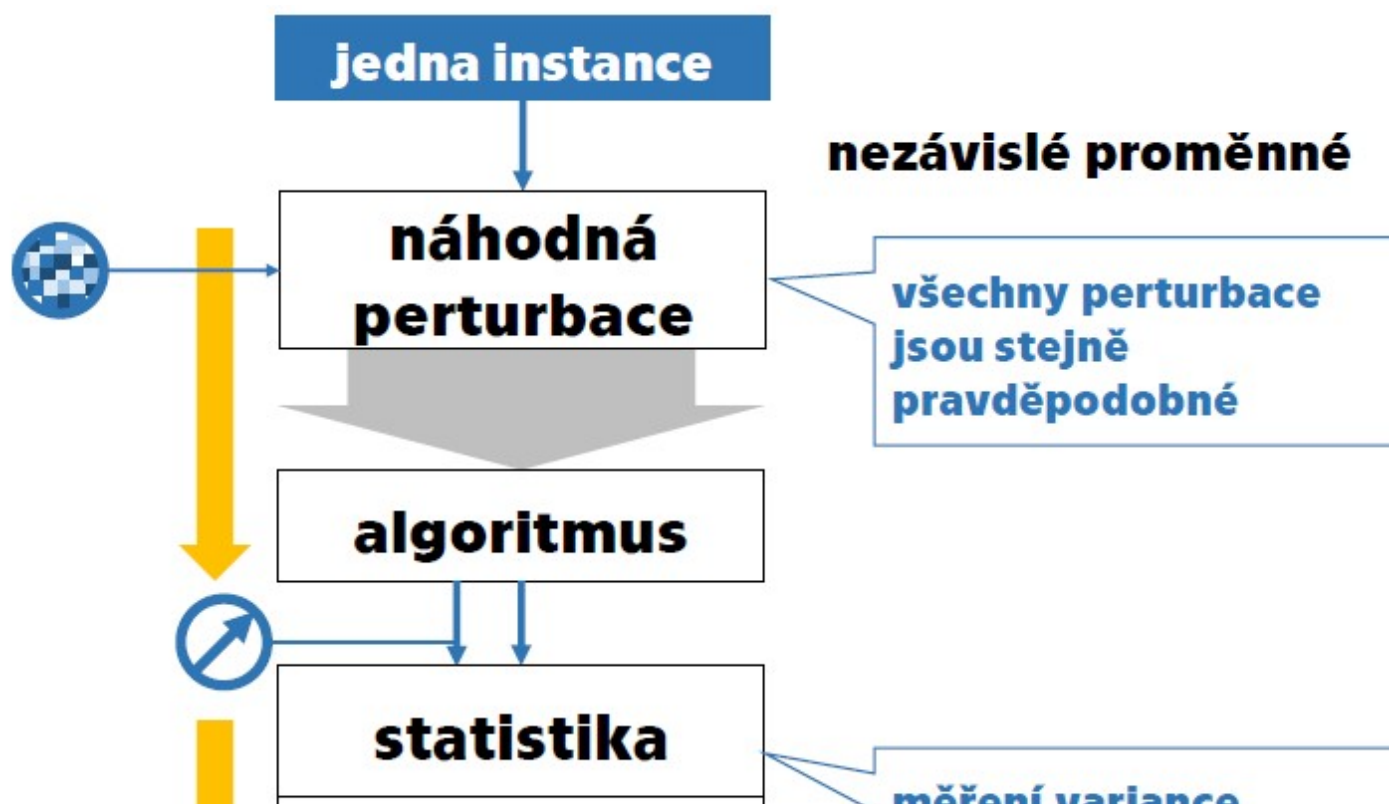
Pro které hodnoty RNG dělat hlubší analýzu?

Robustnost heuristiky: závislost práce heuristiky na lexikálním uspořádání dat, které nemění význam

(zpřeházení klauzulí SAT -> jiné řešení v jiném čase,

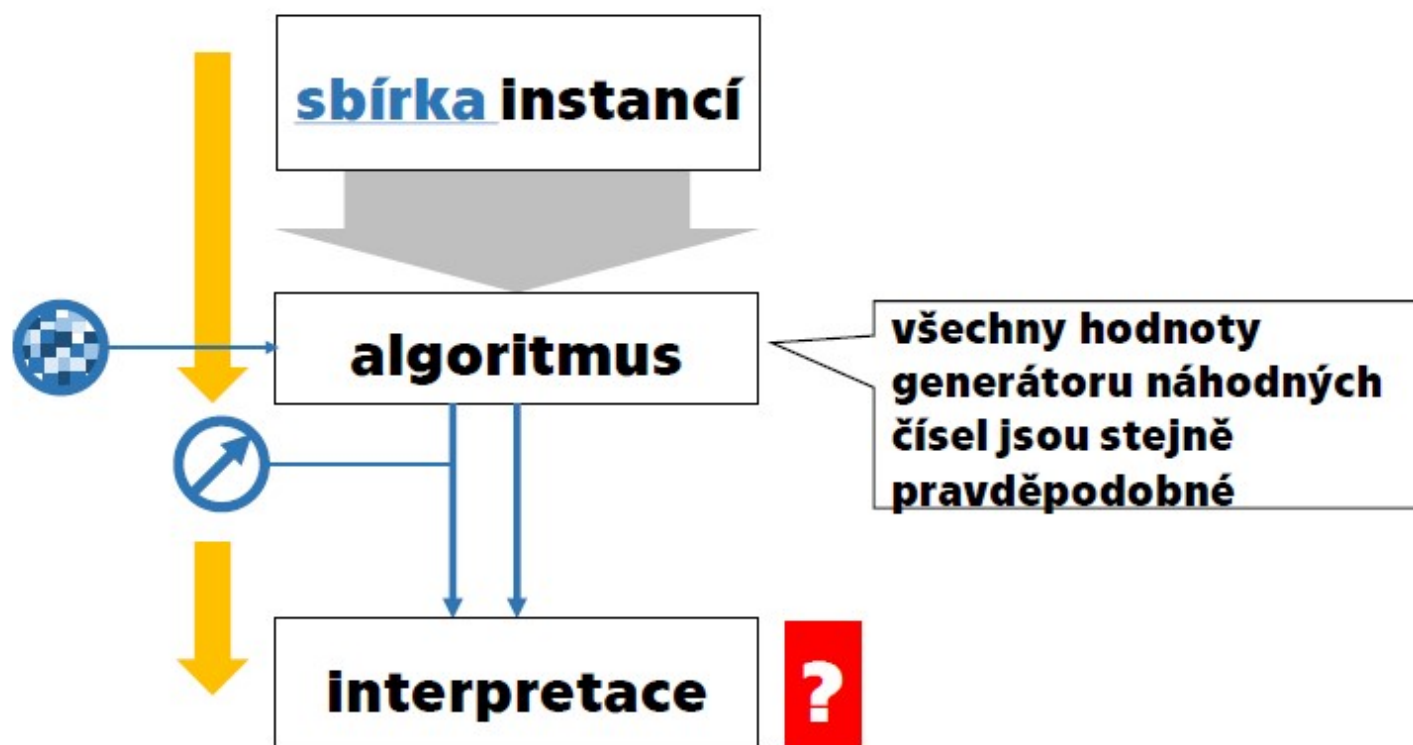
zpřeházení pořadí deklarace proměnných -> rychlejší program)

Měření robustnosti:





Inženýrská algoritmika



- žádná známá standardní sada instancí **nemá zaručenou statistickou reprezentativnost**:
 - neplatí statistické předpoklady
 - nefunguje průměrování
 - nelze eliminovat neznámé zdroje šumu
 - ze statistiky zbyl pouze existenční kvantifikátor (existuje instance, na které...)
- srovnání: **vzhledem k dané sadě** instancí
- **relevantnost pro praxi**: otázka na autora sady

Vizualizace

Pro heuristiky pracující iterativně s **jednou aktuální konfigurací**

Nejčastěji **vývoj optimalizačního kritéria** s pořadím iterace

Lze odpozorovat některé vlastnosti:

- heuristika skončí u **nejbližšího trochu dobrého** řešení a jiná ani nezkoumá
- heuristika **bezcílně bloudí** prostorem konfigurací
- heuristika nachází **čím dál horší** řešení
- heuristika hledá správně, ale **skončí moc brzy**

Práce s heuristikou

- **white box evaluation:**
 - omezená sada instancí
 - detailní měření
 - vhled, porozumění
 - modifikace heuristiky
- **black box evaluation:**
 - plná sada instancí
 - měření výsledků
 - ověření kvality a výkonu
 - žádné modifikace heuristiky

Více parametrů heuristiky:

- obecně nejsou nezávislé -- nutno ověřit
- nastavování více parametrů = cesta prostorem konfigurací heuristiky

Praktické nasazení heuristik

- praktické požadavky: rychlost, kvalita, žádná intervence koncového uživatele
- heuristiky:
 - výměnné části, parametry, vzájemná závislost
- práce s heuristikou a experimentální vyhodnocení:
 - vhled do činnosti heuristiky
 - white box, black box
 - standardní zkušební úlohy

- otázka, plán, experiment
- vizualizace