Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

CS4618 Artificial Intelligence I

Today: Asymmetric Mutations
Crossover

Thomas Jansen

November 23rd

Announcements

Fact some asked for a tutorial/review of CS4618
before the in-class test

Consequence open QA session
date, time and place to be decided
Doodle on course web site
if you have not done so, enter your data 'immediately'
since I will decide today at 3pm

Reminder in-class test next Friday
look for sample paper on course web site
late next Monday

Plans for Today

Remember: Asymmetric Mutations

Standard Bit Mutations

1. For $i \in \{1, 2, \ldots, n\}$
2. With probability $1/n$
3. set $y[i] := 1 - x[i]$
4. else set $y[i] := x[i]$

(1+1) EA

Asymmetric Mutations

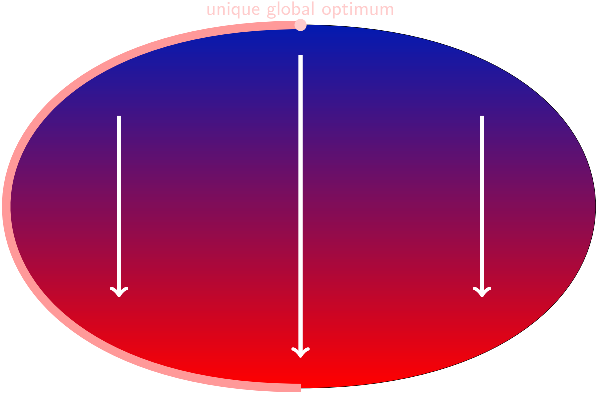
1. For $i \in \{1, 2, \ldots, n\}$
2. If $x[i] = 1$
3. then $p_m = 1/(2\,|x|)$
4. else $p_m = 1/(2(n - |x|))$.
5. With probability p_m
6. set $y[i] := 1 - x[i]$
7. else set $y[i] := x[i]$

asymmetric (1+1) EA

Plateau

Def. $\textsc{Plateau}\colon \{0,1\}^n \to \{0,1,\ldots,n+1\}$

$$\textsc{Plateau}(x) = \begin{cases} n+1 & \text{if } x = 1^n \\ n & \text{if } x = 1^i0^{n-i},\ i \in \{0,1,\ldots,n-1\} \\ n - |x| & \text{otherwise} \end{cases}$$



Remember: Results on PLATEAU

Def. $\text{PLATEAU}\colon \{0,1\}^n \to \{0,1,\dots,n+1\}$

$$\text{PLATEAU}(x) = \begin{cases} n+1 & \text{if } x = 1^n \\ n & \text{if } x = 1^i 0^{n-i},\ i \in \{0,1,\dots,n-1\} \\ n - |x| & \text{otherwise} \end{cases}$$

- $\mathrm{E}\left(T_{(1+1)\ \text{EA},\text{PLATEAU}}\right) = \Theta(n^3)$
  $\widehat{=}$efficient
- $\mathrm{Prob}\left(T_{\text{asym. } (1+1)\ \text{EA},\text{PLATEAU}} = n^{o\left(n^{1/6}\right)}\right) = 2^{-\Omega\left(n^{1/6}\right)}$
  $\widehat{=}$completely inefficient

Is this only due to the slight variation bias
or more due to the plateau-structure?

Consider asymmetric (1+1) EA on $\text{PLATEAU}_{a_{10}}$
with $a_{10} = 101010\cdots10$

Clear $\mathrm{E}\left(T_{(1+1)\ \text{EA},\text{PLATEAU}_{a_{10}}}\right) = \Theta(n^3)$

Asymmetric Mutations on $\textsc{Plateau}_{a_{10}}$

Consider asymmetric (1+1) EA on $\textsc{Plateau}_{a_{10}}$
with $a_{10} = 101010 \cdots 10$

Consider $1^i 0^{n-i} \oplus a_{10} = 1^i 0^{n-i} \oplus 10101010 \cdots 1010$
$= 010101 \cdots 1010$

Thus $\forall x = 1^i 0^{n-i}\colon \ \left|1^i 0^{n-i} \oplus a_{10}\right| = n - \left|1^i 0^{n-i} \oplus a_{10}\right| (\pm 1)$

Thus Prob (increase number of 1-bits | change number of 1-bits)
$\approx \mathrm{Prob}$ (decrease number of 1-bits | change number of 1-bits)
everywhere on the plateau

Thus fair random walk on the plateau
for the asymmetric (1+1) EA on $\textsc{Plateau}_{a_{10}}$

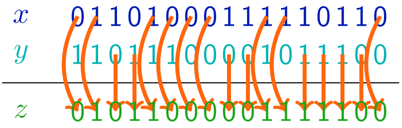
Thus $\mathrm{E}\left(T_{\text{asym. (1+1) EA},\textsc{Plateau}_{a_{01}}}\right) = \Theta(n^3)$

Reminder: Crossover Operators

In general $z = \text{crossover}(x, y)$ with $x, y, z \in \{0,1\}^n$

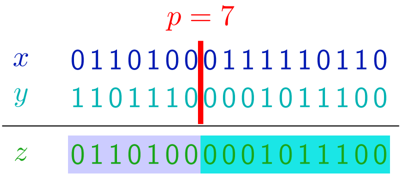
Uniform Crossover

For $i = 1$ to n do
  With probability $1/2$
  set $z[i] = x[i]$
  else set $z[i] = y[i]$



1-Point Crossover

Select $p \in \{0, 1, \ldots, n\}$ u. a. r.
For $i = 1$ to p do
  Set $z[i] = x[i]$.
For $i = p + 1$ to n do
  Set $z[i] = y[i]$.



On Uniform and 1-Point Crossover

Similarities

- both produce one offspring from two parents
- both respect parents: $\forall i\colon x[i] = y[i] \Rightarrow z[i] = x[i] = y[i]$
- both allow for variability:
  $\forall i\colon x[i] \neq y[i] \Rightarrow 0 < \text{Prob}\left(z[i] = 0\right) < 1$

Differences

- variability in offspring
  - uniform crossover: $\forall x, y\colon \ |\{z \mid z = \text{crossover}(x, y)\}| = 2^{\mathsf{H}(x,y)}$
  - 1-point crossover: $\forall x, y\colon \ |\{z \mid z = \text{crossover}(x, y)\}| \leq n + 1$
- positional dependencies
  - uniform crossover: $\forall x, y,$ permutations σ:
    $\{z \mid z = \text{crossover}(x, y)\} = \left\{\sigma^{-1}(z) \mid z = \text{crossover}(\sigma(x), \sigma(y))\right\}$
  - 1-point crossover: $\forall x, y,$ permutations σ:
    $\{z \mid z = \text{crossover}(x, y)\} \neq \left\{\sigma^{-1}(z) \mid z = \text{crossover}(\sigma(x), \sigma(y))\right\}$

Goal demonstrate understanding of uniform and 1-point crossover by presenting appropriate examples where crossover is essential for efficiency

Steady State Genetic Algorithm

Steady State GA

1. Choose $x_1, x_2, \ldots, x_\mu \in \{0,1\}^n$ uniformly at random.
2. Repeat
3. With probability p_c select $z_1, z_2 \in \{x_1, x_2, \ldots, x_\mu\}$ u. a. r.
   $z :=$ crossover(z_1, z_2)
   Else select $z \in \{x_1, x_2, \ldots, x_\mu\}$ uniformly at random.
4. $y :=$ standard bit mutation(z)
5. Select new $x_1, x_2, \ldots, x_\mu$ out of best $x_1, x_2, \ldots, x_\mu, y$.

7 Until 'decide to stop'

8. Output x_i with $f(x_i) = \max\{f(x_j) \mid 1 \leq j \leq \mu\}$.

Parameters: crossover probability $p_c \in [0,1]$
population size $\mu \in \mathbb{N}$

Ideas for an Example Function for 1-Point Crossover

Observation $\forall x, y\colon\ |\{z \mid z = \text{crossover}(x,y)\}| \leq n+1$
$\Rightarrow$ each possible offspring occurs with good probability

Observation $\text{H}(x,y)$ large $\Rightarrow \min\{\text{H}(x,z), \text{H}(y,z)\}$ may be large

Remember mutations very sensitive with respect to Hamming distance
'jumps' of size k exponentially unlikely in k
$\text{Prob}(\text{H}(\text{mutation}(x), x) \geq k) = e^{-\Omega(k)}$

Idea construct example function with easy to find local optima
with large Hamming distance to global optimum
where the Hamming distance is easy to bridge for crossover

Example Function for 1-Point Crossover

Definition $b(x) = \max\{l \mid \exists i\colon x[i] = x[i+1] = \cdots = x[i+l-1] = 1\}$
length of longest block of 1-bits in x

Example $n = 6$, $\textsc{OneMax}(x) = 4$

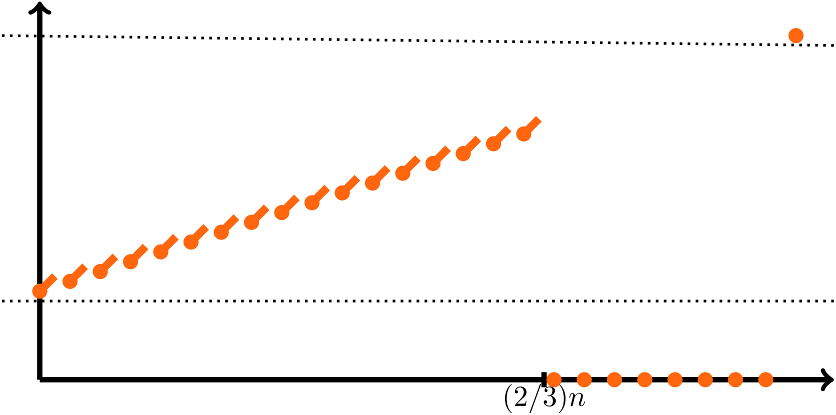
$b(x) = 4$ 001111, 011110, 111100
$b(x) = 3$ 010111, 011101, 100111, 101110, 111001, 111010
$b(x) = 2$ 011011, 101011, 101101, 110011, 110101, 110110

Definition $f_1\colon \{0,1\}^n \to \mathbb{N}_0$ with

$$f_1(x) = \begin{cases} n^2 + 1 & \text{if } x = 1^n \\ n \cdot |x| + b(x) & \text{if } |x| \leq (2/3)n \\ 0 & \text{otherwise} \end{cases}$$

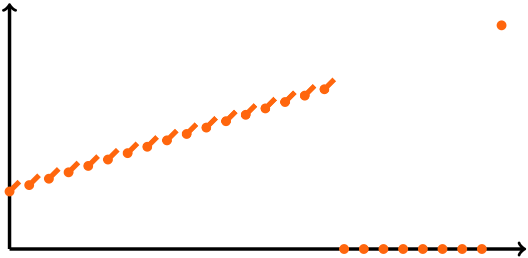
Example Function for 1-Point Crossover

$$f_1(x) = \begin{cases} n^2 + 1 & \text{if } x = 1^n \\ n \cdot |x| + b(x) & \text{if } |x| \leq (2/3)n \\ 0 & \text{otherwise} \end{cases}$$



Without Crossover on f_1

$$f_1(x) = \begin{cases} n^2+1 & \text{if } x = 1^n \\ n\cdot|x| + b(x) & \text{if } |x| \leq (2/3)n \\ 0 & \text{otherwise} \end{cases}$$



Remember without crossover only mutation remains

Observations

- Prob (initial population left of 'gap') $= 1 - 2^{-\Omega(n)}$
- simultaneous mutation of $n/3$ bits needed
- probability for such mutations $2^{-\Omega(n)}$
- $\Rightarrow$ exponential time with overwhelming probability

With 1-Point Crossover on f_1

$$f_1(x) = \begin{cases} n^2+1 & \text{if } x = 1^n \\ n \cdot |x| + b(x) & \text{if } |x| \le (2/3)n \\ 0 & \text{otherwise} \end{cases}$$

Theorem

Consider the steady state GA with crossover probability $p_c \le 1-\varepsilon$ (constant $\varepsilon > 0$) and population size $\mu > n/3$.
$\mathrm{E}\left(T_{\mathsf{GA},f_1}\right) = O\left(\mu n^3 + \mu^2/p_c\right)$

Proof

Phase 1 until all have $(2/3)n$ 1-bits
$\mathrm{E}\left(T\right) = O(\mu \cdot n \log n)$
not worse than μ times ONEMAX
mutation suffices, no crossover with prob. $\ge \varepsilon = \Omega(1)$

At $(2/3)n$ 1-Bits

Phase 2 until all 1-bits in one block
$E(T) = O(\mu \cdot n^3)$
prob. move a 1-bit $\geq (1/n^2)(1-1/n)^{n-2}$
mutation suffices, no crossover with prob. $\geq \varepsilon = \Omega(1)$

Phase 3 until all local optima
$E(T) = O(\mu \cdot n^3)$
prob. move a block $\geq (1/n^2)(1-1/n)^{n-2}$
mutation suffices, no crossover with prob. $\geq \varepsilon = \Omega(1)$

Phase 4 until global optimum
$E(T) = O(\mu^2/p_c)$
prob. pick two good parents and crossover $\geq (1/\mu^2)p_c(1/3)$
crossover needed

Thus
$$E(T_{\mathsf{GA},f_1}) = O(\mu n \log n + \mu n^3 + \mu^2/p_c) = O(\mu n^3 + \mu^2/p_c)$$

□

Ideas for an Example Function for Uniform Crossover

Observation $\forall x, y\colon\ |\{z \mid z = \text{crossover}(x, y)\}| = 2^{\mathsf{H}(x,y)}$
$\Rightarrow$ probability for a specific offspring may be tiny

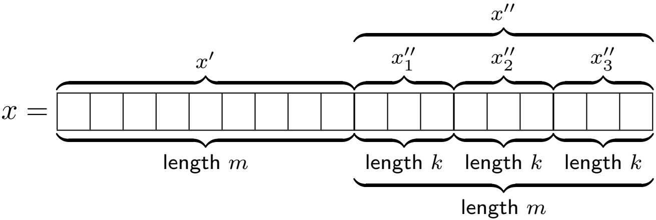
Observation $\mathsf{H}(x, y)$ large $\Rightarrow \min\{\mathsf{H}(x, z), \mathsf{H}(y, z)\}$ may be large

Remember mutations very sensitive with respect to Hamming distance
'jumps' of size k exponentially unlikely in k
$\text{Prob}(\mathsf{H}(\text{mutation}(x), x) \geq k) = e^{-\Omega(k)}$

Idea construct example function with easy to find local optima
with large Hamming distance to global optimum
where the Hamming distance is easy to bridge for crossover
$\widehat{=}$ with large global optimum
$\widehat{=}$ many neighbouring global optima

Preparing an Example Function for Uniform Crossover

Notation partition bit string $x = x'x'' = x'x''_1x''_2x''_3$



$$n = 2m = m + m = m + 3k = m + k + k + k$$

Definition circle $C = \{1^i0^{m-i}, 0^i1^{m-i} \mid i \in \{1, 2, \ldots, m\}\}$
target $T = \{x''_1x''_2x''_3 \mid |x''_1| = |x''_2| = |x''_3| = \lfloor k/2 \rfloor\}$
$\mathsf{H}(x, A) = \min\{\mathsf{H}(x, y) \mid y \in A\}$

Example Function for Uniform Crossover

Definition circle $C = \{1^i 0^{m-i}, 0^i 1^{m-i} \mid i \in \{1, 2, \dots, m\}\}$
target $T = \{x''_1 x''_2 x''_3 \mid |x''_1| = |x''_2| = |x''_3| = \lfloor k/2 \rfloor\}$
$\mathsf{H}(x, A) = \min\{\mathsf{H}(x, y) \mid y \in A\}$

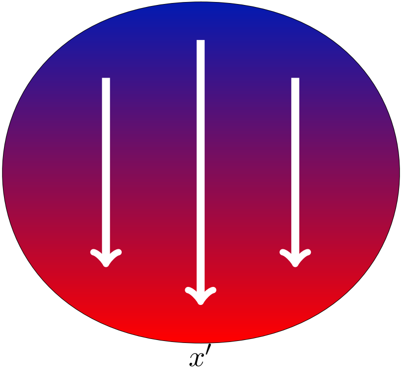
Definition $f_2 \colon \{0,1\}^n \to \mathbb{N}_0$ with

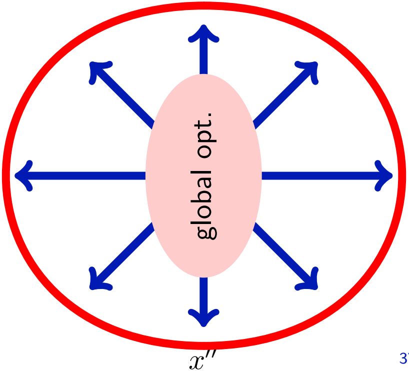
$$f_2(x) = \begin{cases} n - \mathsf{H}(x'', C) & \text{if } x' \neq 0^m \text{ and } x'' \notin C \\ 2n - \mathsf{H}(x', 0^m) & \text{if } x'' \in C \\ 0 & \text{if } x' = 0^m \text{ and } x'' \notin (C \cup T) \\ 3n & \text{if } x' = 0^m \text{ and } x'' \in T \end{cases}$$

Example Function for Uniform Crossover

Definition $f_2\colon \{0,1\}^n \to \mathbb{N}_0$ with

$$f_2(x) = \begin{cases} n - \mathsf{H}\left(x'', C\right) & \text{if } x' \neq 0^m \text{ and } x'' \notin C \\ 2n - \mathsf{H}\left(x', 0^m\right) & \text{if } x'' \in C \\ 0 & \text{if } x' = 0^m \text{ and } x'' \notin (C \cup T) \\ 3n & \text{if } x' = 0^m \text{ and } x'' \in T \end{cases}$$





Summary & Take Home Message

Things to remember

- asymmetric mutations
- carefully biasing variation
- 1-point crossover
- insights and design of appropriate example functions
- immense potential speed-up by crossover

Take Home Message

- Incorporating domain knowledge is usually beneficial.
- Biasing variation can speed search considerably.
- Think and check what you've done.
- Crossover is unique to evolutionary algorithms.
- Crossover can speed-up search dramatically.
- Problem structure needs to be appropriate for crossover to work.
- Encoding issues become more important when using crossover.