Theory and Practice of Optimal Mutation Rate Control in Hamming Spaces of DNA Sequences

Roman V. Belavkin[1] Alastair Channon[2] Elizabeth Aston[2]
John Aston[3] Christopher G. Knight[4]

[1]School of Engineering and Information Sciences
Middlesex University, London NW4 4BT, UK

[2]Research Institute for the Environment, Physical Sciences and Applied Mathematics
Keele University, ST5 5BG, UK

[3]Department of Statistics
University of Warwick, CV4 7AL, UK

[4]Faculty of Life Sciences
University of Manchester, M13 9PT, UK

August 11, 2011, ECAL

Evolution as an Information Dynamic System

- EPSRC Sandpit '*Math of Life*' (July, 2009):

- Three year project (2010–12)

  Middlesex University : Roman Belavkin
  University of Warwick : John Aston
  University of Keele : Alastair Channon & Elizabeth Aston
  University of Manchester : Chris Knight & Rok Krasovec

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.
- Important for adaptation of organisms (Fisher, 1930; Orr, 2005).

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.
- Important for adaptation of organisms (Fisher, 1930; Orr, 2005).

Operations research

- Setting $\mu = 1/l$ (Mühlenbein, 1992; Ochoa et al., 1999; Eigen et al., 1988).

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.
- Important for adaptation of organisms (Fisher, 1930; Orr, 2005).

Operations research

- Setting $\mu = 1/l$ (Mühlenbein, 1992; Ochoa et al., 1999; Eigen et al., 1988).
- Control (Ackley, 1987; Fogarty, 1989; Yanagiya, 1993; Bäck, 1993; Vafaee et al., 2010).

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.
- Important for adaptation of organisms (Fisher, 1930; Orr, 2005).

Operations research

- Setting $\mu = 1/l$ (Mühlenbein, 1992; Ochoa et al., 1999; Eigen et al., 1988).
- Control (Ackley, 1987; Fogarty, 1989; Yanagiya, 1993; Bäck, 1993; Vafaee et al., 2010).

Biology

- Controlled (to a degree) by the organism (e.g. DNA repair, Hakem, 2008).

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.
- Important for adaptation of organisms (Fisher, 1930; Orr, 2005).

Operations research

- Setting $\mu = 1/l$ (Mühlenbein, 1992; Ochoa et al., 1999; Eigen et al., 1988).
- Control (Ackley, 1987; Fogarty, 1989; Yanagiya, 1993; Bäck, 1993; Vafaee et al., 2010).

Biology

- Controlled (to a degree) by the organism (e.g. DNA repair, Hakem, 2008).
- Closely related species can have different μ (e.g. bacterium *Deinococcus radiodurans*, Cox et al., 2010).

Introduction: Optimal Mutation Rates

- Mutation is an innovation process in GA search.
- Important for adaptation of organisms (Fisher, 1930; Orr, 2005).

Operations research

- Setting $\mu = 1/l$ (Mühlenbein, 1992; Ochoa et al., 1999; Eigen et al., 1988).
- Control (Ackley, 1987; Fogarty, 1989; Yanagiya, 1993; Bäck, 1993; Vafaee et al., 2010).

Biology

- Controlled (to a degree) by the organism (e.g. DNA repair, Hakem, 2008).
- Closely related species can have different μ (e.g. bacterium *Deinococcus radiodurans*, Cox et al., 2010).
- May depend on changes in the environment (Bjedov et al., 2003).

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Reproduction $\omega_s \mapsto \omega_{s+1}$

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Reproduction $\omega_s \mapsto \omega_{s+1}$

- $P(x_{s+1} \mid x_s)$ conditional probability of $x_s \mapsto x_{s+1}$.

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Reproduction $\omega_s \mapsto \omega_{s+1}$

- $P(x_{s+1} \mid x_s)$ conditional probability of $x_s \mapsto x_{s+1}$.
- $T := \Big(P(x_{s+1} \mid x_s)\Big)$ — Markov operator, $p_s := P(x_s)$

$$p_{s+1} = Tp_s = \sum_{x_s} P(x_{s+1} \mid x_s)\, P(x_s) \quad \Rightarrow \quad p_{s+t} = T^t p_s$$

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Reproduction $\omega_s \mapsto \omega_{s+1}$

- $P(x_{s+1} \mid x_s)$ conditional probability of $x_s \mapsto x_{s+1}$.
- $T := \Big(P(x_{s+1} \mid x_s)\Big)$ — Markov operator, $p_s := P(x_s)$

$$p_{s+1} = T p_s = \sum_{x_s} P(x_{s+1} \mid x_s)\, P(x_s) \quad \Rightarrow \quad p_{s+t} = T^t p_s$$

- Adaptation $\mathbb{E}\{x_{s+t}\} \geq \mathbb{E}\{x_s\}$, where $\mathbb{E}\{x_s\} := \sum x_s\, P(x_s)$.

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Reproduction $\omega_s \mapsto \omega_{s+1}$

- $P(x_{s+1} \mid x_s)$ conditional probability of $x_s \mapsto x_{s+1}$.
- $T := \Big(P(x_{s+1} \mid x_s)\Big)$ — Markov operator, $p_s := P(x_s)$

$$p_{s+1} = Tp_s = \sum_{x_s} P(x_{s+1} \mid x_s)\, P(x_s) \quad \Rightarrow \quad p_{s+t} = T^t p_s$$

- Adaptation $\mathbb{E}\{x_{s+t}\} \geq \mathbb{E}\{x_s\}$, where $\mathbb{E}\{x_s\} := \sum x_s\, P(x_s)$.

Control

- μ — parameter controlling $P_\mu(x_{s+1} \mid x_s)$.

Preliminaries

Individuals and Fitness

Let Ω — all individual organisms, $f : \Omega \to \mathbb{R}$ fitness function, $x = f(\omega)$.

Reproduction $\omega_s \mapsto \omega_{s+1}$

- $P(x_{s+1} \mid x_s)$ conditional probability of $x_s \mapsto x_{s+1}$.
- $T := \Big(P(x_{s+1} \mid x_s)\Big)$ — Markov operator, $p_s := P(x_s)$

$$p_{s+1} = Tp_s = \sum_{x_s} P(x_{s+1} \mid x_s)\, P(x_s) \quad \Rightarrow \quad p_{s+t} = T^t p_s$$

- Adaptation $\mathbb{E}\{x_{s+t}\} \geq \mathbb{E}\{x_s\}$, where $\mathbb{E}\{x_s\} := \sum x_s\, P(x_s)$.

Control

- μ — parameter controlling $P_\mu(x_{s+1} \mid x_s)$.
- $\mu(x)$ — control function, $T_{\mu(x)}$, $\mathbb{E}_{\mu(x)}\{x_{s+t}\}$.

Optimal Fitness Value Functions

Instantaneous

- Maximum adaptation in no more than λ generations

$$\overline{x}(\lambda) := \sup_{\mu(x)} \{\mathbb{E}_{\mu(x)}\{x_{s+t}\} : t \leq \lambda\}$$

Optimal Fitness Value Functions

Instantaneous

- Maximum adaptation in no more than λ generations

$$\overline{x}(\lambda) := \sup_{\mu(x)} \{\mathbb{E}_{\mu(x)}\{x_{s+t}\} : t \leq \lambda\}$$

- Minimum number of generations to achieve adaptation υ

$$\overline{x}^{-1}(\upsilon) := \inf_{\mu(x)} \{t \geq 0 : \mathbb{E}_{\mu(x)}\{x_{s+t}\} \geq \upsilon\}$$

Optimal Fitness Value Functions

Instantaneous

- Maximum adaptation in no more than λ generations

$$\overline{x}(\lambda) := \sup_{\mu(x)} \{\mathbb{E}_{\mu(x)}\{x_{s+t}\} : t \leq \lambda\}$$

- Minimum number of generations to achieve adaptation υ

$$\overline{x}^{-1}(\upsilon) := \inf_{\mu(x)} \{t \geq 0 : \mathbb{E}_{\mu(x)}\{x_{s+t}\} \geq \upsilon\}$$

Cumulative

$$\sup_{\mu(x)} \sum_{\lambda=0}^{t} \mathbb{E}_{\mu(x)}\{x_{s+\lambda}\} \leq \sum_{\lambda=s}^{t} \overline{x}(\lambda)$$

Optimal Fitness Value Functions

Information Dynamics (Belavkin, 2010, 2011)

- Maximum adaptation in no more than λ bits between p_s and p_{s+t}:

$$\overline{x}(\lambda) := \sup_{\mu(x)} \{\mathbb{E}_{\mu(x)}\{x_{s+t}\} : \mathbb{E}\{\log(p_{s+t}/p_s)\} \leq \lambda\}$$

- Minimum number of bits to achieve adaptation υ

$$\overline{x}^{-1}(\upsilon) := \inf_{\mu(x)} \{\mathbb{E}\{\log(p_{s+t}/p_s)\} : \mathbb{E}_{\mu(x)}\{x_{s+t}\} \geq \upsilon\}$$

Cumulative

$$\sup_{\mu(x)} \sum_{\lambda=0}^{t} \mathbb{E}_{\mu(x)}\{x_{s+\lambda}\} \leq \sum_{\lambda=s}^{t} \overline{x}(\lambda)$$

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \ldots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \ldots, \alpha_l)$.

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \dots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \dots, \alpha_l)$.
Hamming space : $\mathcal{H}_\alpha^l := \{1, \dots, \alpha\}^l$ with metric $d(a, b) := |\{i : a_i \neq b_i\}|$.

Relatively Monotonic Landscapes

Representation : alphabet $\{1,\dots,\alpha\}$, genotypes $\omega \iff (\alpha_1,\dots,\alpha_l)$.
Hamming space : $\mathcal{H}_\alpha^l := \{1,\dots,\alpha\}^l$ with metric $d(a,b) := |\{i : a_i \neq b_i\}|$.

Question

Rugged landscape?

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \dots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \dots, \alpha_l)$.
Hamming space : $\mathcal{H}^l_\alpha := \{1, \dots, \alpha\}^l$ with metric $d(a, b) := |\{i : a_i \neq b_i\}|$.

$$a \lesssim b \xrightarrow{\quad f \quad} f(a) \leq f(b)$$

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \dots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \dots, \alpha_l)$.
Hamming space : $\mathcal{H}_\alpha^l := \{1, \dots, \alpha\}^l$ with metric $d(a, b) := |\{i : a_i \neq b_i\}|$.

$$\begin{tikzcd}
a \lesssim b \arrow[r, "f"] \arrow[d] & f(a) \leq f(b) \\
a, b \in \mathcal{H}_\alpha^l &
\end{tikzcd}$$

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \dots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \dots, \alpha_l)$.
Hamming space : $\mathcal{H}_\alpha^l := \{1, \dots, \alpha\}^l$ with metric $d(a, b) := |\{i : a_i \neq b_i\}|$.

$$
\begin{tikzcd}
a \lesssim b \arrow[r, "f"] \arrow[d] & f(a) \leq f(b) \\
a, b \in \mathcal{H}_\alpha^l \arrow[r, "d"] & -d(a, \top) \leq -d(b, \top) \arrow[u]
\end{tikzcd}
$$

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \dots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \dots, \alpha_l)$.
Hamming space : $\mathcal{H}_\alpha^l := \{1, \dots, \alpha\}^l$ with metric $d(a, b) := |\{i : a_i \neq b_i\}|$.

$$
\begin{tikzcd}
a \lesssim b \arrow[r, "f"] \arrow[d] & f(a) \leq f(b) \\
a, b \in \mathcal{H}_\alpha^l \arrow[r, "d"] & -d(a, \top) \leq -d(b, \top) \arrow[u]
\end{tikzcd}
$$

Definition (Relatively Monotonic Landscape)

f is locally monotonic (isomorphic) relative to a metric d, if there exist $B(\top, l) := \{\omega : d(\top, \omega) \leq l\}$, $\top = \sup \Omega$, such that $\forall\, a, b \in B(\top, l)$:

$$-d(\top, a) \leq -d(\top, b) \quad \Longrightarrow (\iff) \quad f(a) \leq f(b)$$

Relatively Monotonic Landscapes

Representation : alphabet $\{1, \dots, \alpha\}$, genotypes $\omega \iff (\alpha_1, \dots, \alpha_l)$.
Hamming space : $\mathcal{H}_\alpha^l := \{1, \dots, \alpha\}^l$ with metric $d(a, b) := |\{i : a_i \neq b_i\}|$.

$$
\begin{tikzcd}
a \lesssim b \arrow[r, "f"] \arrow[d] & f(a) \leq f(b) \\
a, b \in \mathcal{H}_\alpha^l \arrow[r, "d"] & -d(a, \top) \leq -d(b, \top) \arrow[u]
\end{tikzcd}
$$

Definition (Relatively Monotonic Landscape)

f is locally monotonic (isomorphic) relative to a metric d, if there exist $B(\top, l) := \{\omega : d(\top, \omega) \leq l\}$, $\top = \sup \Omega$, such that $\forall\, a, b \in B(\top, l)$:

$$-d(\top, a) \leq -d(\top, b) \quad \Longrightarrow (\iff) \quad f(a) \leq f(b)$$

Example (Needle in a haystack)

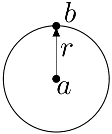
$f(\omega) = 1$ if $d(\top, \omega) = 0$; $f(\omega) = 0$ otherwise.

Mutation and Adaptation in a Hamming Space

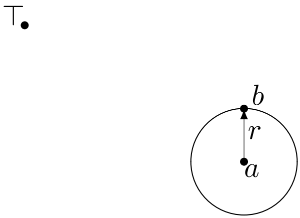
$\bullet a$

Mutation and Adaptation in a Hamming Space

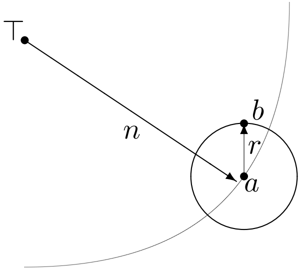
Mutation and Adaptation in a Hamming Space



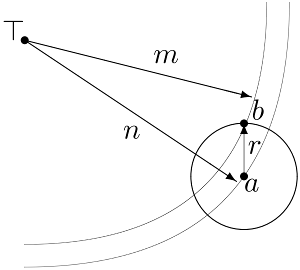
Mutation and Adaptation in a Hamming Space



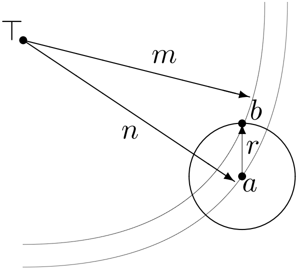
Mutation and Adaptation in a Hamming Space



Mutation and Adaptation in a Hamming Space

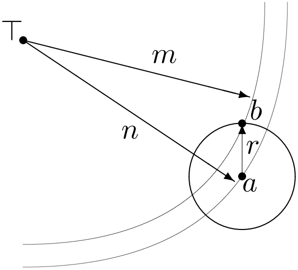


Mutation and Adaptation in a Hamming Space



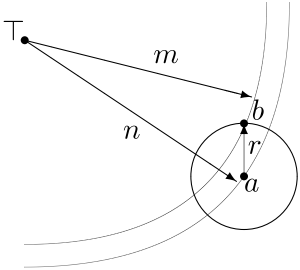
- $a \mapsto b \in S(a, r)$.

Mutation and Adaptation in a Hamming Space



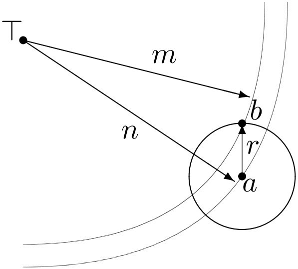
- $a \mapsto b \in S(a, r)$.
- r is mutation radius

Mutation and Adaptation in a Hamming Space



- $a \mapsto b \in S(a, r)$.
- r is mutation radius
- $P_{\mu}(m \mid n) =$?

Mutation and Adaptation in a Hamming Space

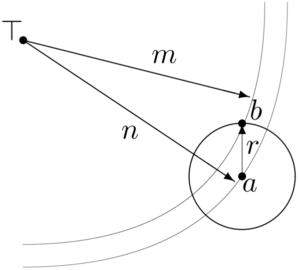


- Expand for all $r \in [0, l]$:

$$P_{\mu}(m \mid n) = \sum_{r=0}^{l} P(m \mid n, r)\, P_{\mu}(r \mid n)$$

- $a \mapsto b \in S(a, r)$.
- r is mutation radius
- $P_{\mu}(m \mid n) =$?

Mutation and Adaptation in a Hamming Space



- Expand for all $r \in [0, l]$:

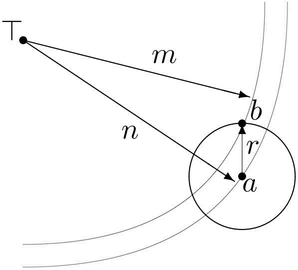
$$P_\mu(m \mid n) = \sum_{r=0}^{l} P(m \mid n, r)\, P_\mu(r \mid n)$$

- In a Hamming space $\mathcal{H}_\alpha^l$:

$$P_\mu(r \mid n) = \binom{l}{r} \mu(n)^r (1 - \mu(n))^{l-r}$$

- $a \mapsto b \in S(a, r)$.
- r is mutation radius
- $P_\mu(m \mid n) = ?$

Mutation and Adaptation in a Hamming Space



- $a \mapsto b \in S(a, r)$.
- r is mutation radius
- $P_\mu(m \mid n) = ?$

- Expand for all $r \in [0, l]$:

$$P_\mu(m \mid n) = \sum_{r=0}^{l} P(m \mid n, r)\, P_\mu(r \mid n)$$

- In a Hamming space $\mathcal{H}^l_\alpha$:

$$P_\mu(r \mid n) = \binom{l}{r} \mu(n)^r (1 - \mu(n))^{l-r}$$

- and

$$P(m \mid n, r) = \frac{|S(\top, m) \cap S(a, r)|_{d(\top,a)=n}}{|S(a, r)|}$$

Adaptation in One Generation

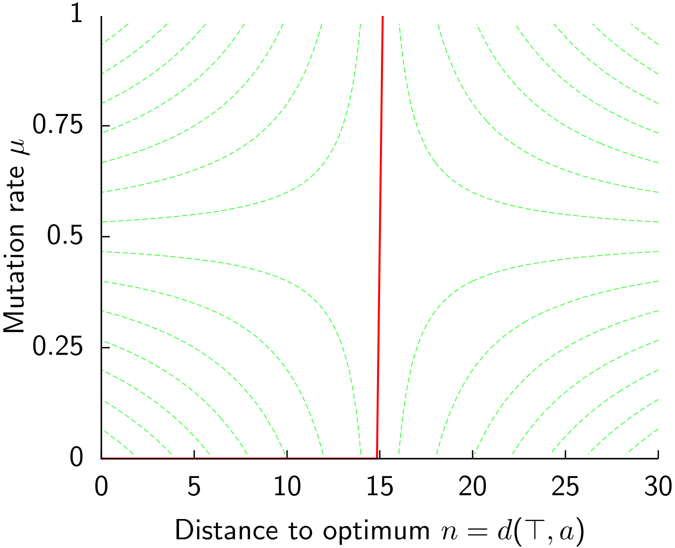
- Minimize $\mathbb{E}\{n_{s+t}\}$ subject to $t \leq 1$.

Adaptation in One Generation

- Minimize $\mathbb{E}\{n_{s+t}\}$ subject to $t \leq 1$.
- In this case the optimal function is

$$\mu(n) := \begin{cases} 0 & \text{if } n < l(1 - 1/\alpha) \\ \frac{1}{2} & \text{if } n = l(1 - 1/\alpha) \\ 1 & \text{otherwise} \end{cases}$$

Step function



Maximizing Probability of Success

- Probability of ‘success’ $P_{\mu}(m < n \mid n)$ (Bäck, 1993, for $\mathcal{H}_2^l$).

Maximizing Probability of Success

- Probability of 'success' $P_\mu(m < n \mid n)$ (Bäck, 1993, for $\mathcal{H}_2^l$).
- Define $\hat{\mu}(n)$ such that

$$P_{\hat{\mu}}(m < n \mid n) = \max_{\mu} P_\mu(m < n \mid n)$$

Maximizing Probability of Success

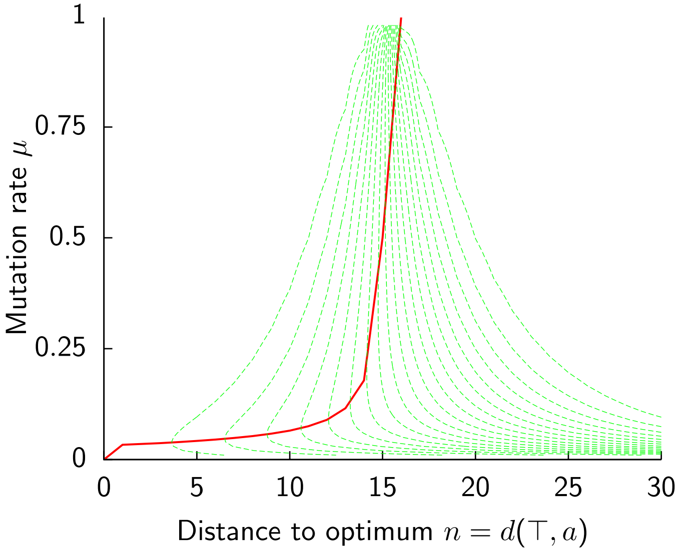
- Probability of 'success' $P_\mu(m < n \mid n)$ (Bäck, 1993, for $\mathcal{H}_2^l$).
- Define $\hat{\mu}(n)$ such that

$$P_{\hat{\mu}}(m < n \mid n) = \max_\mu P_\mu(m < n \mid n)$$

- This corresponds to maximization of $\mathbb{E}\{u(m,n)\}$, where

$$u(m,n) := \begin{cases} 1 & \text{if } m < n \\ 0 & \text{otherwise} \end{cases}$$

$\max P_{\mu}(m < n \mid n)\}$



Mutation to Optimum

- Probability of mutating directly to optimum

$$P_{\mu}(m = 0 \mid n) = (\alpha - 1)^{-n} \mu^{n} (1 - \mu)^{l-n}$$

Mutation to Optimum

- Probability of mutating directly to optimum

$$P_{\mu}(m = 0 \mid n) = (\alpha - 1)^{-n} \mu^{n} (1 - \mu)^{l-n}$$

- Maximization conditions $P'_{\mu} = 0$ and $P''_{\mu} \leq 0$ give $n - l\mu = 0$ or

$$\mu(n) = \frac{n}{l}$$

Mutation to Optimum

- Probability of mutating directly to optimum

$$P_{\mu}(m = 0 \mid n) = (\alpha - 1)^{-n} \mu^{n} (1 - \mu)^{l-n}$$

- Maximization conditions $P'_{\mu} = 0$ and $P''_{\mu} \leq 0$ give $n - l\mu = 0$ or

$$\mu(n) = \frac{n}{l}$$

Remark

- *For $n = 1$ we have $\mu = 1/l$ (error threshold).*

Mutation to Optimum

- Probability of mutating directly to optimum

$$P_{\mu}(m = 0 \mid n) = (\alpha - 1)^{-n}\mu^{n}(1 - \mu)^{l-n}$$

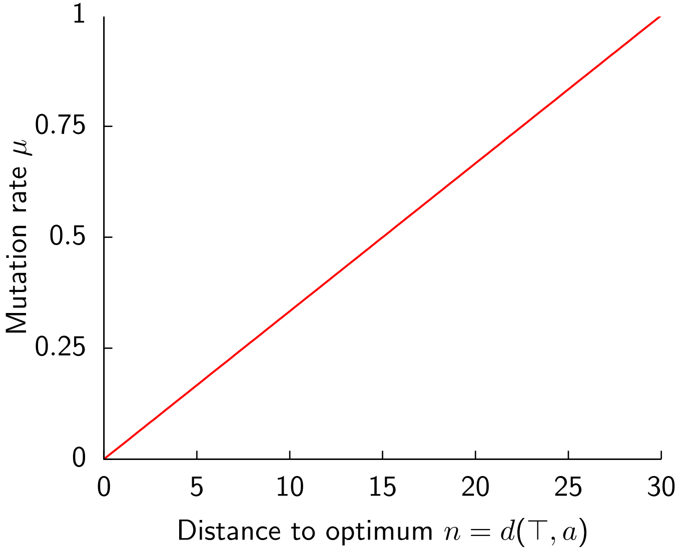
- Maximization conditions $P'_{\mu} = 0$ and $P''_{\mu} \leq 0$ give $n - l\mu = 0$ or

$$\mu(n) = \frac{n}{l}$$

Remark

- *For $n = 1$ we have $\mu = 1/l$ (error threshold).*
- *Optimal for Boolean landscapes (Needle in a haystack).*

Linear function



Information Heuristics $t \leq \lambda \iff I_{KL}(p_{s+t}, p_s) \leq \lambda$

- The optimal μ corresponds to CDF of $P_0(m)$:

$$\mu(n) = P_0(m < n) = \sum_{m=0}^{n-1} P_0(m)$$

Information Heuristics $t \leq \lambda \iff I_{KL}(p_{s+t}, p_s) \leq \lambda$

- The optimal μ corresponds to CDF of $P_0(m)$:

$$\mu(n) = P_0(m < n) = \sum_{m=0}^{n-1} P_0(m)$$

- $P_0(m)$ is computed from uniform distribution $P_0(\omega) = \alpha^{-l}$:

$$P_0(m) = \binom{l}{m} \left(1 - \frac{1}{\alpha}\right)^m \left(\frac{1}{\alpha}\right)^{l-m} = \binom{l}{m} \frac{(\alpha - 1)^m}{\alpha^l}$$

Information Heuristics $t \leq \lambda \iff I_{KL}(p_{s+t}, p_s) \leq \lambda$

- The optimal μ corresponds to CDF of $P_0(m)$:

$$\mu(n) = P_0(m < n) = \sum_{m=0}^{n-1} P_0(m)$$

- $P_0(m)$ is computed from uniform distribution $P_0(\omega) = \alpha^{-l}$:

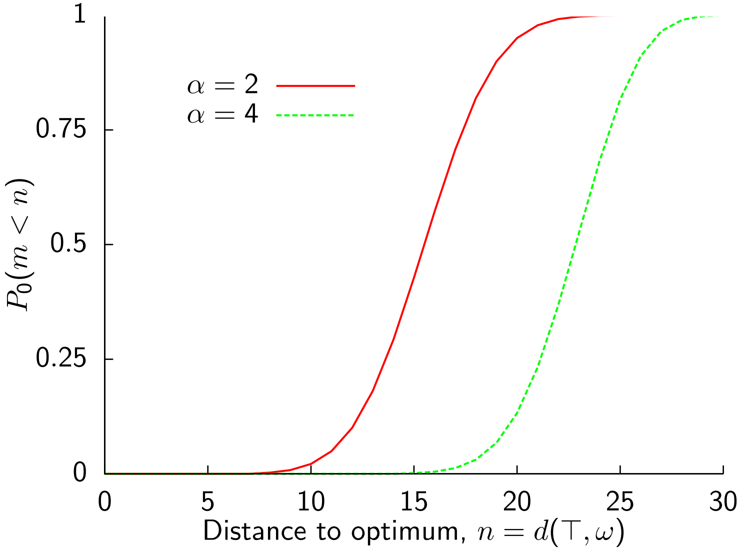
$$P_0(m) = \binom{l}{m} \left(1 - \frac{1}{\alpha}\right)^m \left(\frac{1}{\alpha}\right)^{l-m} = \binom{l}{m} \frac{(\alpha - 1)^m}{\alpha^l}$$

Informed Mutation Rate

In a (weakly) monotonic landscape we can use CDF of empirical frequency P_e of observed fitness values:

$$P_0(m) \iff P_e(x) \quad \text{and} \quad P_0(m < n) \iff P_e(x_r > x)$$

'Informed' Mutation function



Practice: Evolving Optimal Mutation Rates

Inner GA

Genotypes : sequences in $\mathcal{H}_\alpha^l$.

Populations : 100 individuals.

Generations : $t = 500$.

Evolution : mutation only.

Objective : maximize $x = f(\omega)$.

Practice: Evolving Optimal Mutation Rates

Inner GA

Genotypes : sequences in $\mathcal{H}_\alpha^l$.

Populations : 100 individuals.

Generations : $t = 500$.

Evolution : mutation only.

Objective : maximize $x = f(\omega)$.

Meta GA

Genotypes : functions $\mu(x)$, $\mu \in [0, 1]$.

Populations : 100 individuals.

Generations : $t = 5 \cdot 10^5$.

Evolution : tournament selection, recombination, mutation.

Objective : maximize $\mathbb{E}\{x\}$ in Inner GA at the last generation.

Experimental Results

1. $\mathcal{H}_2^{30}$ (i.e. $\alpha = 2$, $l = 30$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.

Experimental Results

1. $\mathcal{H}_2^{30}$ (i.e. $\alpha = 2$, $l = 30$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.
2. $\mathcal{H}_4^{10}$ (i.e. $\alpha = 4$, $l = 10$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.

Experimental Results

1. $\mathcal{H}_2^{30}$ (i.e. $\alpha = 2$, $l = 30$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.
2. $\mathcal{H}_4^{10}$ (i.e. $\alpha = 4$, $l = 10$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.
3. $\mathcal{H}_4^{10}$ (i.e. $\alpha = 4$, $l = 10$) and fitness $f(\omega)$ defined by a complete DNA-protein affinity landscape for 10-base-pair sequences (Rowe et al., 2010), which we refer to as the aptamer landscape.

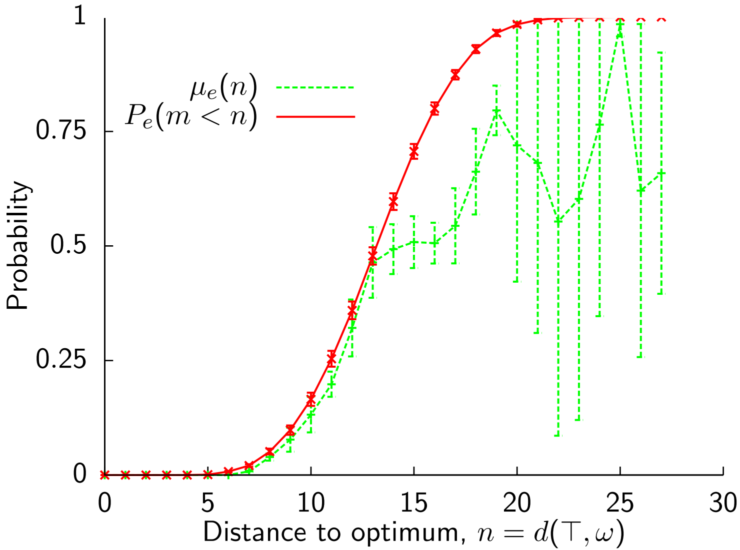
Experimental Results

1. $\mathcal{H}_2^{30}$ (i.e. $\alpha = 2$, $l = 30$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.
2. $\mathcal{H}_4^{10}$ (i.e. $\alpha = 4$, $l = 10$) and fitness $f(\omega) = -d(\top, \omega)$, where d is Hamming metric.
3. $\mathcal{H}_4^{10}$ (i.e. $\alpha = 4$, $l = 10$) and fitness $f(\omega)$ defined by a complete DNA-protein affinity landscape for 10-base-pair sequences (Rowe et al., 2010), which we refer to as the aptamer landscape.

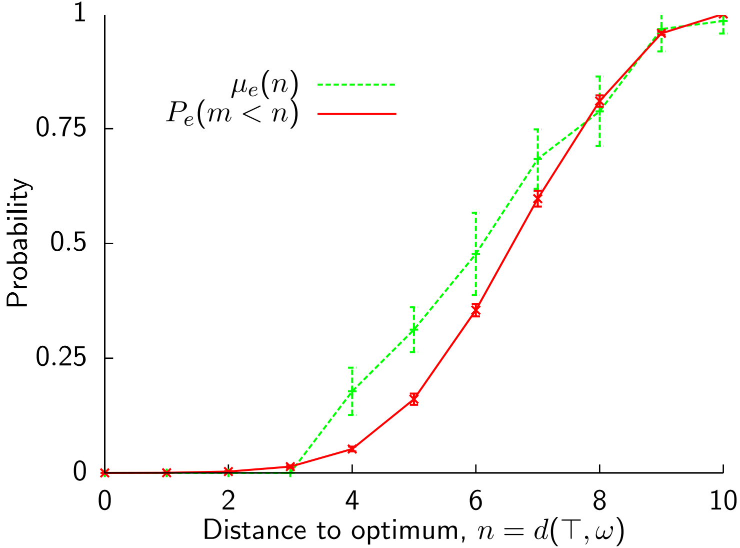
Output

- $\mu_e(x)$ — evolved mutation rate functions.
- $P_e(x_r > x)$ — CDFs of empirical distributions $P_e(x)$ of fitness.

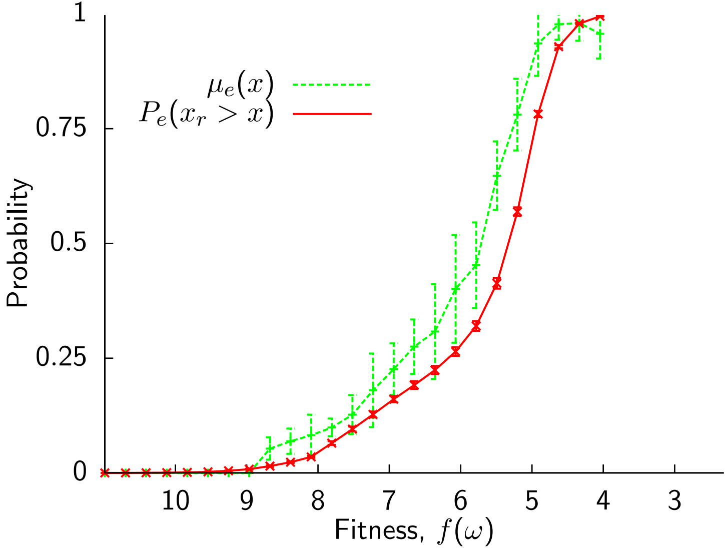
$\mathcal{H}_2^{30}$, fitness $f(\omega) = -d(\top, \omega)$



$\mathcal{H}_4^{10}$, fitness $f(\omega) = -d(\top, \omega)$



$\mathcal{H}_4^{10}$, fitness the aptamer landscape (Rowe et al., 2010)



Conclusions and Questions

- Analytical formulae for $P_{\mu}(m \mid n)$ in $\mathcal{H}^l_\alpha$.

Conclusions and Questions

- Analytical formulae for $P_{\mu}(m \mid n)$ in $\mathcal{H}^l_\alpha$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).

Conclusions and Questions

- Analytical formulae for $P_{\mu}(m \mid n)$ in $\mathcal{H}^l_\alpha$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).
- Exact optimization is hard, but possible in some cases and approximate for others.

Conclusions and Questions

- Analytical formulae for $P_{\mu}(m \mid n)$ in $\mathcal{H}^l_\alpha$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).
- Exact optimization is hard, but possible in some cases and approximate for others.
- New heuristic $\mu(x) = P_e(x_r > x)$ based on empirical CDF and information dynamics.

Conclusions and Questions

- Analytical formulae for $P_{\mu}(m \mid n)$ in $\mathcal{H}^l_\alpha$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).
- Exact optimization is hard, but possible in some cases and approximate for others.
- New heuristic $\mu(x) = P_e(x_r > x)$ based on empirical CDF and information dynamics.
- Evolved control functions in Protein-DNA affinity landscapes confirm our conjecture that natural fitness is (weakly) monotonic relative to $\mathcal{H}^l_4$ of DNA sequences ($\Rightarrow$ control of μ benefits natural evolution).

Conclusions and Questions

- Analytical formulae for $P_{\mu}(m \mid n)$ in $\mathcal{H}_{\alpha}^{l}$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).
- Exact optimization is hard, but possible in some cases and approximate for others.
- New heuristic $\mu(x) = P_e(x_r > x)$ based on empirical CDF and information dynamics.
- Evolved control functions in Protein-DNA affinity landscapes confirm our conjecture that natural fitness is (weakly) monotonic relative to $\mathcal{H}_4^l$ of DNA sequences ($\Rightarrow$ control of μ benefits natural evolution).

Question

- *Including control of mutation rate adds cost in complexity.*

Conclusions and Questions

- Analytical formulae for $P_\mu(m \mid n)$ in $\mathcal{H}^l_\alpha$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).
- Exact optimization is hard, but possible in some cases and approximate for others.
- New heuristic $\mu(x) = P_e(x_r > x)$ based on empirical CDF and information dynamics.
- Evolved control functions in Protein-DNA affinity landscapes confirm our conjecture that natural fitness is (weakly) monotonic relative to $\mathcal{H}^l_4$ of DNA sequences ($\Rightarrow$ control of μ benefits natural evolution).

Question

- *Including control of mutation rate adds cost in complexity.*
- *Does the gain in performance outweigh this cost?*

Conclusions and Questions

- Analytical formulae for $P_\mu(m \mid n)$ in $\mathcal{H}^l_\alpha$.
- Defined relatively monotonic landscapes to clarify the role of a representation space (i.e. 'rugged' is relative).
- Exact optimization is hard, but possible in some cases and approximate for others.
- New heuristic $\mu(x) = P_e(x_r > x)$ based on empirical CDF and information dynamics.
- Evolved control functions in Protein-DNA affinity landscapes confirm our conjecture that natural fitness is (weakly) monotonic relative to $\mathcal{H}^l_4$ of DNA sequences ($\Rightarrow$ control of μ benefits natural evolution).

Question

- *Including control of mutation rate adds cost in complexity.*
- *Does the gain in performance outweigh this cost?*
- *Have biological organisms evolved such controls?*

Ackley, D. H. (1987). An empirical study of bit vector function optimization. In L. Davis (Ed.), *Genetic algorithms and simulated annealing* (pp. 170–204). Pitman.

Bäck, T. (1993). Optimal mutation rates in genetic search. In S. Forrest (Ed.), *Proceedings of the 5th International Conference on Genetic Algorithms* (pp. 2–8). Morgan Kaufmann.

Belavkin, R. V. (2010). Information trajectory of optimal learning. In M. J. Hirsch, P. M. Pardalos, & R. Murphey (Eds.), *Dynamics of information systems: Theory and applications* (Vol. 40). Springer.

Belavkin, R. V. (2011). On evolution of an information dynamic system and its generating operator. *Optimization Letters*, 1–14. (10.1007/s11590-011-0325-z)

Bjedov, I., Tenaillon, O., Gerard, B., Souza, V., Denamur, E., Radman, M., et al. (2003). Stress-induced mutagenesis in bacteria. *Science*, *300*(5624), 1404-9.

Cox, M. M., Keck, J. L., & Battista, J. R. (2010). Rising from the ashes: DNA repair in Deinococcus radiodurans. *PLoS Genet*, *6*(1), e1000815.

Eigen, M., McCaskill, J., & Schuster, P. (1988). Molecular quasispecies. *Journal of Physical Chemistry*, *92*, 6881-6891.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Oxford University Press.

Fogarty, T. C. (1989). Varying the probability of mutation in the genetic algorithm. In J. D. Schaffer (Ed.), *Proceedings of the 3rd International Conference on Genetic Algorithms* (pp. 104–109). Morgan Kaufmann.

Hakem, R. (2008). DNA-damage repair; the good, the bad, and the ugly. *Embo J*, *27*(4), 589-605.

Mühlenbein, H. (1992, September). How genetic algorithms really work: Mutation and hillclimbing. In einhard Männer & B. Manderick (Eds.), *Parallel problem solving from nature 2* (pp. 15–26). Brussels, Belgium: Elsevier.

Ochoa, G., Harvey, I., & Buxton, H. (1999). Error thresholds and their relation to optimal mutation rates. In *Proceedings of the Fifth European Conference on Artificial Life (ECAL'99)* (Vol. 1674, pp. 54–63). Berlin: Springer-Verlag.

Orr, H. A. (2005). The genetic theory of adaptation: a brief history. *Nat Rev Genet*, *6*(2), 119-27.

Rowe, W., Platt, M., Wedge, D. C., Day, P. J., & Kell, D. B. (2010). Analysis of a complete DNA-protein affinity landscape. *Journal of Royal Society Interface*, *7*(44), 397–408.

Vafaee, F., Turán, G., & Nelson, P. C. (2010). Optimizing genetic operator rates using a Markov chain model of genetic algorithms. In M. Pelikan & J. Branke (Eds.), (pp. 721–728). ACM.

Yanagiya, M. (1993). A simple mutation-dependent genetic algorithm. In S. Forrest (Ed.), *Proceedings of the 5th International Conference on Genetic Algorithms* (p. 659). Morgan Kaufmann.