Complexity of evolutionary equilibria in static fitness landscapes

ArXiv 1308.5094

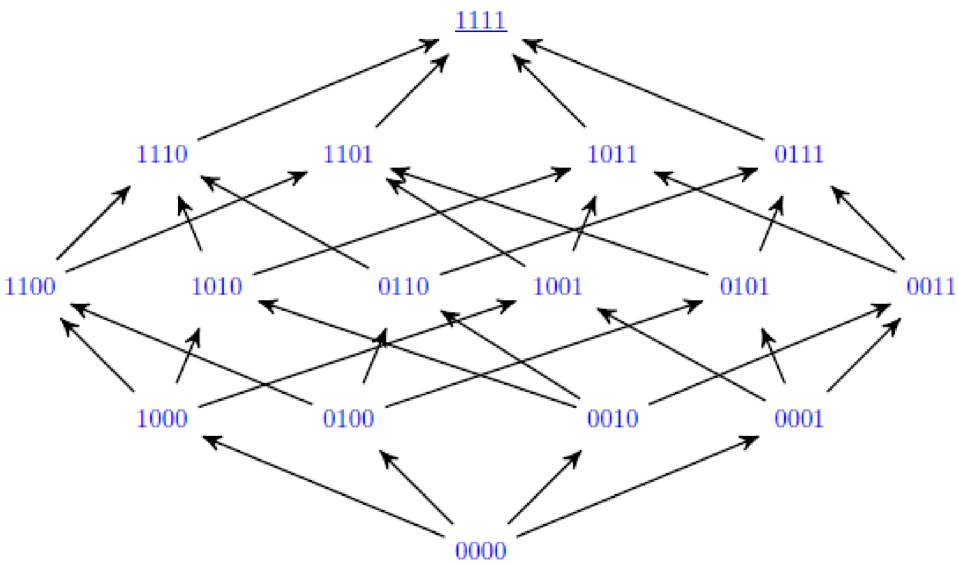
Artem Kaznatcheev
McGill University
egtheory.wordpress.com

"In a rugged field of this character selection will easily carry the species to the nearest peak, but there may be innumerable other peaks which are higher but which are separated by “valleys.” The problem of evolution as I see it is that of a mechanism by which the species may continually find its way from lower to higher peaks in such a field."

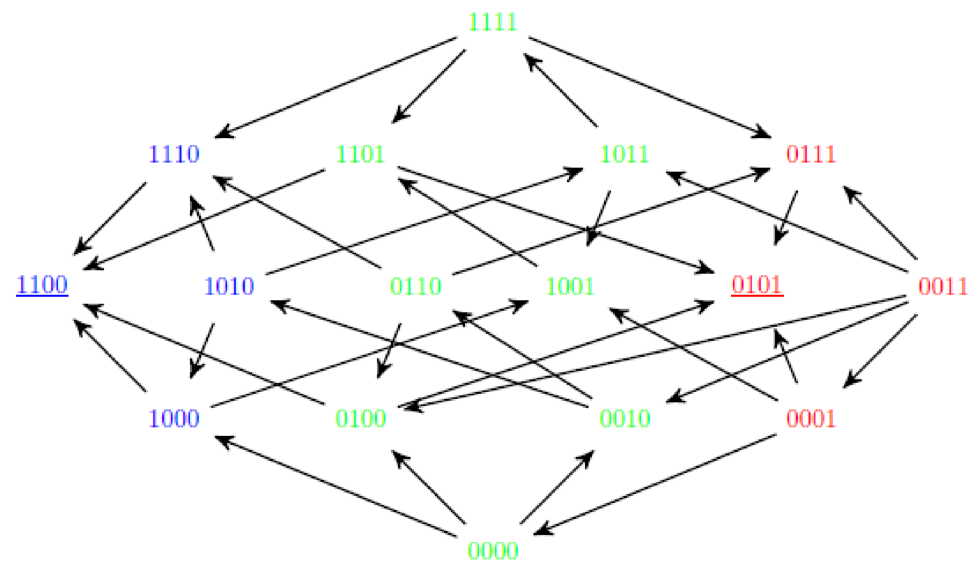
Wright, S. (1932).
The roles of mutation, inbreeding, crossbreeding, and selection in evolution.
Proceedings of the Sixth International Congress of Genetics, 356-366

http://egtheory.wordpress.com/2013/08/16/fitness-landscapes/

H. H. Chou, H. C. Chiu, N. F. Delaney, D. Segr, and C. J. Marx (2011).
Diminishing returns epistasis among benecial mutations decelerates adaptation.
Science, 6034:1190-1192.



(a) *Escherichia coli* β-lactamase



(b) *Plasmodium falciparum* dihydrofolate reductase

E. R. Lozovsky, T. Chookajorn ... and D. L. Hartl. (2009).
Stepwise acquisition of pyrimethamine resistance in the malaria parasite.
Proc. Natl. Acad. Sci. USA, 106:12025-12030.

http://egtheory.wordpress.com/2013/08/18/empirical-epistasis/

2.4 NK model of rugged fitness landscapes

Definition 5 ([KL87, KE89, Kau93]). The NK *model* is a fitness landscape on $\{0,1\}^n$. The n loci are arranged in a gene-interaction network where each locus x_i is linked to K other loci $x^i_1, ..., x^i_k$ and has an associated fitness contribution function $f_i : \{0,1\}^{K+1} \to \mathbb{R}_+$ Given a vertex $v \in \{0,1\}^n$, we define the fitness $f(x) = \sum_{i=1}^{n} f_i(x_i x^i_1 ... x^i_k)$.

S. Kauman and S. Levin. (1987)
Towards a general theory of adaptive walks on rugged landscapes.
Journal of Theoretical Biology, 128:11-45.

http://egtheory.wordpress.com/2013/08/25/nk-block-models/

Justifying asymptotic worst-case analysis to scientists

20

5

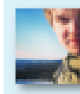
I've been working on on introducing some results from computational complexity into theoretical biology, especially evolution & ecology, with the goal of being interesting/useful to biologists. One of the biggest difficulties I've faced is in justifying the usefulness of asymptotic worst-case analysis for lower bounds. **Are there any article length references that justify lower bounds and asymptotic worst-case analysis to a scientific audience?**

I am really looking for a good reference that I can defer to in my writing instead of having to go through the justifications in the limited space I have available (since that is not the central point of the article). I am also aware of other kinds and paradigms of analysis, so **I am *not* seeking a**

cc.complexity-theory reference-request big-picture application-of-theory worst-case

mod share edit close delete flag

asked Jan 25 at 22:52



Artem Kaznatcheev ◆
5,399 ● 5 ● 37 ● 109

http://cstheory.stackexchange.com/q/20731/1037

Justifying asymptotic worst-case analysis to scientists

20

5

I've been working on on introducing some results from computational complexity into theoretical biology, especially evolution & ecology, with the goal of being interesting/useful to biologists. One of the biggest difficulties I've faced is in justifying the usefulness of asymptotic worst-case analysis for lower bounds. **Are there any article length references that justify lower bounds and asymptotic worst-case analysis to a scientific audience?**

I am really looking for a good reference that I can defer to in my writing instead of having to go through the justifications in the limited space I have available (since that is not the central point of the article). I am also aware of other kinds and paradigms of analysis, so **I am *not* seeking a**

cc.complexity-theory reference-request big-picture application-of-theory worst-case

mod share edit close delete flag

asked Jan 25 at 22:52
Artem Kaznatcheev
5,399 5 37 109

22 Worst-case behavior is impossible to justify ... Peter Shor Jan 25 at 23:15

6

My personal (and biased) take is that asymptotic worst-case analysis is a historical stepping stone to more practically useful kinds of analysis. It therefore seems hard to justify to practitioners.

Proving bounds for the worst case is often easier than proving bounds for even "nice" definitions of average case. Asymptotic analysis is also often much easier than proving reasonably tight bounds. Worst-case asymptotic analysis is therefore a great place to start.

answered Jan 30 at 20:59
András Salamon
9,846 2 30 98

http://cstheory.stackexchange.com/q/20731/1037

Theorem 7. *Finding a local optimum in the NK fitness landscape is PLS-complete.*

For $K \leq 1$, even a global optimum can be found in polynomial time [WTZ00], so the theorem is as strong as it can be.

Proof. Consider an instance of Weighted 2SAT with variables $x_1, ..., x_n$, clauses $C_1, ..., C_m$ and positive integer costs $c_1, ..., c_m$. We will build a landscape with $m + n$ loci, with the first m labeled $b_1, ..., b_m$ and the next n labeled $x_1, ..., x_m$. Each b_k will correspond uses the variables x_i and x_j (i.e., the first literal is either x_i or $\bar{x}_i$ and the set $i < j$ to avoid ambiguity). Define the corresponding fitness effect of the

$$f_k(0x_ix_j) = \begin{cases} c_k & \text{if } C_k \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases}$$

$$f_k(1x_ix_j) = f_k(0x_ix_j) + 1$$

D.S. Johnson, C.H. Papadimitriou, and M. Yannakakis. (1988)
How easy is local search?
Journal of Computer and System Sciences, 37:79-100.

A.A. Schaer and M. Yannakakis. (1991)
Simple local search problems that are hard to solve.
SIAM Journal on Computing, 20(1):56-87.

If P != PLS then

there exist NK fitness landscapes with K = 2 such that *no local fitness peak can be found in polynomial time*

If evolution only follows adaptive walks

there exist NK fitness landscapes with K = 2 (and initial wild types) such that *every adaptive path to any local fitness peak is of exponential length*

follow the adaptive paths then we can strengthen the result:

Corollary 8. *There is a constant $c > 0$ such that, for infinitely many n, there are instances of NK models (with $K \geq 2$) on $\{0,1\}^n$ and initial vertices v such that any adaptive path from v will have to take at least 2^{cn} steps before finding a fitness peak.*

Proof. If the initial vertex has $s = 11...1$ then there is a bijection between adaptive paths in the fitness landscape and any weight-increasing path for optimizing the weighted 2SAT problem. Thus, theorem 5.15 of [SY91] applies. □

http://egtheory.wordpress.com/2013/08/31/evolutionary-equilibria/

If we consider **fittest mutant** (very very large populations) wins then
there exist semi-smooth fitness landscapes such that it takes an *exponential number of steps* to get to the unique fitness peak.

Relevant CStheory results:

Klee, V., & Minty, G.J. (1972).
How good is the simplex algorithm?.
In Shisha, Oved. *Inequalities III*. 159–175.

Jeroslow, R.G. (1973).
The simplex algorithm with the pivot rule of maximizing criterion improvement.
Discrete Mathematics 4(4): 367-377.

If we consider **strong-selection weak-mutation dynamics** (random adaptive step) then
there exist semi-smooth fitness landscapes such that with high probability, the expected number of steps to get to the unique fitness peak is *exp(O(n^(1/3)))*.

Relevant CStheory results:

Matousek, J., & Szabo, T. (2006).
RANDOM EDGE can be exponential on abstract cubes.
Advances in Mathematics, 204, 262-277.

http://egtheory.wordpress.com/2013/09/12/semi-smooth-fitness-landscapes/

If we want to be within a few adaptive steps of a local fitness peak then

nope

Relevant CStheory results:
S.T. Fischer. (1995)
A note on the complexity of local search problems.
Information Processing Letters, 53:69-75.

If we want to be close in fitness to a local fitness peak then

nope, $f(x)/f(x^*) < 2^{(-kn)}$

Relevant CStheory results:
H. Klauck. (1996).
On the hardness of global and local approximation.
Algorithm Theory – SWAT: 88-99.

If we want to be within a few adaptive steps of a local fitness peak then

nope

Relevant CStheory results:
S.T. Fischer. (1995)
A note on the complexity of local search problems.
Information Processing Letters, 53:69-75.

If we want to be close in fitness to a local fitness peak then

nope, $f(x)/f(x^*) < 2^{(-kn)}$

Relevant CStheory results:
H. Klauck. (1996).
On the hardness of global and local approximation.
Algorithm Theory – SWAT: 88-99.

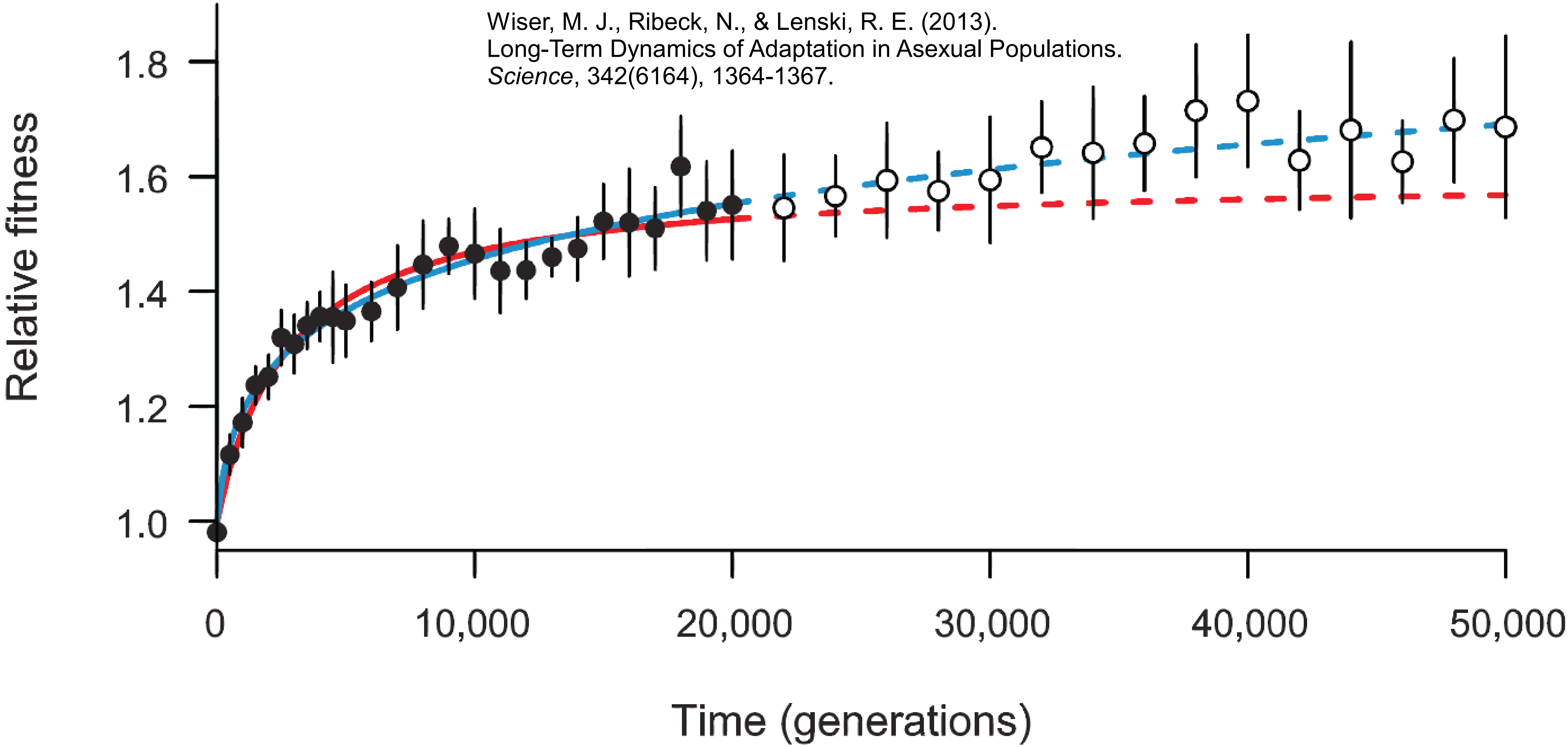
If we want want every neighbour *y* of our focal *x* to be within a small fitnss then

yup!

Relevant CStheory results:
J.B. Orlin, A.P. Punnen, and A.S. Schulz. (2004)
Approximate local search in combinatorial optimization.
SIAM J. Comput., 33:1201-1214.

Specifically, if $f(x) \geq (1 - s)f(y)$ then

we **can** find an approximate local fitness peak in *poly(n,* ***1/s****)* steps, and
we **cannot** find an approximate local fitness peak in *poly(n,* ***ln(1/s)****)* steps



YES: $s = \text{poly}(n, 1/T)$

NO: $s = \exp(-T)\text{poly}(n)$

LTEE: $s \sim a/T$

“The host is picking the fitness landscape on which the pathogen evolves“

- Carl T. Bergestrom