Evolution with culture†

María-Elena Solano

Abstract I study a simple variant of Holland’s Genetic Algorithm (GA) equipped with a fixed-rate ‘Boltzmann’ selection rule \( p_{\text{select}} \propto e^{-\eta(1-\text{Fitness})} \), meant to capture the ‘viral’ spread of adaptive innovations characteristic of cultural (aka memetic) evolution. Using the recent groundbreaking reduction by Chastain et al. (2014) and Meir & Parkes (2015) of evolution to coordination game dynamics with multiplicative weights updates, I show that, for a fixed (optimal) value of \( \eta \), the time-averaged expected fitness of populations evolved under this variant follows \( O(1/\sqrt{T}) \) no-regret dynamics —whereas GA ‘merely’ achieves \( O(1/T) \) diminishing regret dynamics, ultimately bounded by the selection strength of the fitness landscape. Experiments on a challenging Ms. Pac-Man grammatical evolution (GE) task, however, provide mixed evidence for these theoretical results.

1 Introduction

What makes cultural (aka memetic) evolution different? I claim the answer is ‘virality’: unlike genes, highly adaptive memes quickly spread person-to-person throughout society in a contagion-like manner. As a result of this ‘accelerated’ spread, I argue, highly adaptive memeplexes ‘survive’ the exponential halving induced by recombination. As Dawkins put it (1976, §11):

[If you contribute to the world’s culture, if you have a good idea, compose a tune, invent a sparking plug, write a poem, it may live on, intact, long after your genes have dissolved in the common pool.]

To model this claim quantitatively, consider the following simple variant of Holland’s Genetic Algorithm (GA) (1975): instead of selecting chromosomes \( c \in \mathcal{C} \) using a fitness-proportional rule

\[
p_{\text{select}} \propto \text{Fitness}(c),
\]

select according to the exponentially-weighted (aka ‘Boltzmann’) rule

\[
p_{\text{select}} \propto e^{-\eta(1-\text{Fitness}(c))},
\]

for some fixed \( \eta \). Henceforth, I use the term ‘MA’ to refer to this variant\(^1\).

---

†I am grateful to Lisa Meeden for her invaluable guidance, suggestions and mentorship. Any errors remain my own. I thank Andrew Ruether for his generous assistance with setting up the experiments, which was made possible through the XSEDE Campus Champions program. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

1Short for ‘memetic algorithm’. This is not to suggest, however, that this is a novel variant in any way —cf. §1.1.
fitness regret after \( T \) generations:

**Theorem 1.** Let \( s = \max \{ \text{Fitness}(\cdot) \} - \min \{ \text{Fitness}(\cdot) \} \) be the selection strength of the fitness landscape. Then with high probability, after \( T > 0 \) generations, GA’s time-averaged expected fitness regret is at most

\[
s^2 + O(1/T).
\]

**Theorem 2.** Let \( \eta = 4|\mathcal{C}|/(1 - \tilde{f}_{\min})^{1/2} \) and \( \tilde{f}_{\min} \leq \text{Fitness}(\cdot) \in [0, 1] \). Then with high probability, after \( T > 0 \) generations, MA’s time-averaged expected fitness regret is at most

\[
O \left( \sqrt{(1 - \tilde{f}_{\min})/T} \right) + O(1/T).
\]

In other words, whereas GA ‘merely’ follows \( O(1/T) \) diminishing regret dynamics ultimately bounded by the selection strength of the fitness landscape \( s \), MA follows \( O(1/\sqrt{T}) \) no-regret dynamics —or up to \( O(1/T) \) for \( \tilde{f}_{\min} \to 1 \). Hence, for ‘jagged’ landscapes with high \( s \), MA will likely outperform GA. On the other hand, on ‘flat’ landscapes with low \( s \), GA will converge faster than MA.

Note the resemblance with the canonical regret bounds of the celebrated Multiplicative Weights (MW) class of algorithms (cf. Arora, Hazan & Kale 2012). Indeed, Theorems 1 and 2 derive directly —with minor differences—from the recent groundbreaking reduction by Chastain et al. (2014) and Meir & Parkes (2015) of evolution to a coordination game between genes with MW updates, plus Freund & Schapire (1997)’s regret bounds for their exponentially-weighted variant of MW (aka ‘Hedge’). The (short) proof is presented in §3.

Also note the bound in Theorem 2 is ‘opportunistic’, in the sense it relies on setting \( \eta \) to its (approximate) optimal value \( \eta^* \approx \Theta(\ln |\mathcal{C}|) \). Nevertheless, I argue this value admits a parsimonious interpretation: namely, the rate needed to (roughly) maximize the entropy of the resulting marginal allele densities subject to the information encoded in the fitness landscape. This (heuristic) claim is made more precise in §3.5.

Another implication of the connection to MW is that these bounds are tight up to additive constants (cf. Arora, Hazan & Kale 2012, §4)—so they should hold well in practice. Yet as it turns out, they do not. This claim is verified empirically in §4.

Lastly, I note these regret bounds are largely impractical for evolutionary computing (EC) applications, where the main concern is to maximize the best fitness of the population, not the mean. Furthermore, they only hold with high probability due to the distorting effect of mutation. These and other open problems are discussed in §5.

### 1.1 Related work

Exponentially-weighted (aka ‘Boltzmann’) selection rules have been studied extensively in the EC literature (cf. Mitchell 1995, §5.4). Goldman (1990) described a novel (approximate) reduction of GA to simulated annealing by implementing a tournament scheme with exponentially-weighted selection probabilities and a time-varying rate \( \eta = \eta_t \), roughly corresponding to the (inverse of the) ‘temperature’
of the search. De la Maza & Tidor (1991) showed that GA with time-annealed Boltzmann selection outperforms ‘canonical’ GA on a set of benchmark tasks, and suggested superior entropy control as the likely source of the advantage. In the context of statistical mechanics, Prügel-Bennett & Shapiro (1994) reduced GA with fixed-rate Boltzmann selection —as in (2)— and Gaussian initial fitness distribution to Derrida’s Random Energy Model (1981), a tractable class of spin-glass dynamics with stable thermal equilibrium —and hence a unique optimal choice of \( \eta = \eta^* \).

The contagion-like, accelerated dynamics of cultural transmission vis-à-vis genetic transmission has been studied extensively in the broader sociobiology literature (cf. Boyd, Richerson & Henrich 2011, pp.431-432). Cavalli-Sforza & Feldman (1981, §1.9) study a simple model of frequency-proportionate cultural transmission that results in S-shaped, logistic-like trait frequencies. Boyd & Richerson (1983, §5) describe a stylized dual-inheritance model of population dynamics in which fitness-proportionate cultural transmission leads to an exponential increase in trait frequency, in direct proportion to the variance of the fitness landscape.

In a different line of research, Kanade (2010, §5.4) reduced a parallel variant of Feldman’s correlational statistical query (CSQ) learning (Feldman 2008, §3) with polynomially many parallel steps to an extension of Valiant’s evolvability (2009) with recombination and polylog number of generations. Interestingly, under this reduction, evolvability with recombination and weak selection can only simulate a somewhat weaker variant of CSQ that requires polylog overhead.

2 Preliminaries

Throughout this paper, I use \(|\cdot|\) to denote set cardinality, and \(\diamond\) to denote vector concatenation. Otherwise, I mostly follow the notation in Meir & Parkes (2015) (where applicable).

2.1 Setup

Consider an initial population \(O(t=0)\) or finite multiset of chromosomes \(c = (a_1, \ldots, a_m) \in C\) with population densities \(p_c(t=0) \sim D(t=0)\), each comprising \(m\) genes (or memes) \(a_i \in A\) drawn from a finite set of \(n\) integer-valued alleles \(A = \{1, \ldots, n\}\). Given a bounded fitness function \(f : C \to [0, 1]\), we want to evolve the population: that is, given an evolutionary algorithm EA, to iteratively produce a chain of populations

\[
O^{(0)} \xrightarrow{\text{EA}} O^{(1)} \xrightarrow{\text{EA}} \ldots \xrightarrow{\text{EA}} O^{(T)}
\]

such that, after \(T\) generations, the expected fitness of chromosomes in the last population \(O^{(T)}\)

\[
E_{c \sim P^{(T)}}[f(c)] = \sum_{c \in C} f(c) p_c^{(T)}
\]

(3)

\[
= \frac{1}{|O^{(T)}|} \sum_{c \in O^{(T)}} f(c)
\]

(4)

is as high as ‘possible’. In an adversarial environment, a desirable goal would be to bound the ‘instan-
taneous’ fitness regret (cf. Cesa-Bianchi and Lugosi 2006, §I)
\[ E_{c \sim P(T)}[f(c)] - E[\hat{f}(T)], \] (5)
where \( \hat{f}(T) \) is the highest achievable fitness by the EA after \( T \) steps.

In this paper, however, I focus on the more opaque —and, in a sense, weaker— notion of time-averaged fitness regret
\[ \frac{1}{T} \sum_{t=1}^{T} E_{c \sim P(t)}[f(c)] - \frac{1}{T} \sum_{t=1}^{T} E[\hat{f}(T)], \] (6)
which, although much less interpretable than (5), is much more common in the online learning literature —and indeed, is the usual form of the canonical bounds I am drawing from. The extension of Theorems 1 and 2 to the much more natural notion of ‘instantaneous’ regret is discussed further in §6.

2.2 Evolutionary Algorithms

Let me formally state Holland’s ‘canonical’ GA:

**Algorithm 1 (GA) (Holland 1975)**
*Input:* Fitness function \( f : \mathcal{C} \to \mathbb{R}_+ \), initial population \( \mathcal{O}^{(0)} \) of chromosomes drawn from \( \mathcal{C} \), number of generations \( T \), mutation and crossover rates \( p_m, p_r \in [0, 1] \).
1. For \( t = 1, \ldots, T \):
   i. Let \( \mathcal{O}^{(t)} = \) new population of \( |\mathcal{O}^{(t-1)}| \) chromosomes obtained by repeating steps (a)-(c):
      a. Select two chromosomes from \( \mathcal{O}^{(t-1)} \), with probability \( f(\cdot)/\sum_{c \in \mathcal{O}^{(t-1)}} f(c) \).
      b. Cross the two resulting chromosomes, with probability \( p_r \).
      c. Mutate the two resulting chromosomes, with probability \( p_m \).
2. Output \( \mathcal{O}^{(T)} \).

—and its ‘memetic’ variant, MA:

**Algorithm 2 (MA)**
*Input:* Fitness function \( f : \mathcal{C} \to [0, 1] \), initial population \( \mathcal{O}^{(0)} \) of chromosomes drawn from \( \mathcal{C} \), number of generations \( T \), mutation and crossover rates \( p_m, p_r \in [0, 1] \), free parameter \( \eta \).
1. For \( t = 1, \ldots, T \):
   i. Let \( \mathcal{O}^{(t)} = \) new population of \( |\mathcal{O}^{(t-1)}| \) chromosomes obtained by repeating steps (a)-(c):
      a. Select two chromosomes from \( \mathcal{O}^{(t-1)} \), with probability \( e^{-\eta(1-f(\cdot))}/\sum_{c \in \mathcal{O}^{(t-1)}} e^{-\eta(1-f(c))} \).
      b. Cross the two resulting chromosomes, with probability \( p_r \).
      c. Mutate the two resulting chromosomes, with probability \( p_m \).
2. Output \( \mathcal{O}^{(T)} \).

Note the only difference between GA and MA —other than normalization— is in the Select step, parameterized by \( \eta \). Also note the Cross step denotes single-point crossover, as in Holland’s original GA.

Let me also state the following ‘generalized’ variant of GA (henceforth GA\(_{\text{na}}\)) with no mutation and ar-
bitary homologous regions—which corresponds to the multi-gene version of the evolutionary model studied in Meir & Parkes (2015, §A):

Algorithm 3 (GA$_{na}$)

\textbf{Input}: Fitness function }$f : C \to \mathbb{R}_+\text{, initial population } \mathcal{O}^{(0)}\text{ of chromosomes drawn from } C\text{, number of generations } T\text{, crossover rate } p_r \in [0, 1]$.  

1. For }$t = 1, \ldots, T$:
   i. Let }$\mathcal{O}^{(t)} = \text{ new population of } |\mathcal{O}^{(t-1)}| \text{ chromosomes obtained by repeating steps (a)-(c)}:$
   a. Select two chromosomes from }$\mathcal{O}^{(t-1)}$, with probability }$f(\cdot) / \sum_{c \in \mathcal{O}^{(t-1)}} f(c)$.  
   b. Cross the two chromosomes (on a random partition of }$\{1, \ldots, m\}$, with probability }$p_r$.  
2. Output }$\mathcal{O}^{(T)}$.

Note that single-point crossover corresponds to the case where the recombinable homologous regions are limited to the pairs }$\{\emptyset, \{1, \ldots, m\}\}$, }$\{\{1\}, \{2, \ldots, m\}\}$, }$\{\{1, 2\}, \{3, \ldots, m\}\}$, and so forth.

2.3 Regret Bounds for Coordination Games with Multiplicative Weights Updates

Consider a repeated coordination game between }$m$ players }$l \in \{1, \ldots, m\}$, }$n$ actions }$a \in \{1, \ldots, n\}$, and shared payoff }$\Delta(e) = \Delta((a_1, \ldots, a_m)) \in [0, 1]$; where at time }$t > 0$ each player }$l$ updates their (initially uniform) mixed strategies }$X_l^{(t)}$ according to the exponentially-weighted multiplicative weights (MW) update rule (aka ‘Hedge’) (Freund & Schapire 1996):

\[
X_l^{(t)}(a) \propto X_l^{(t-1)}(a) E[e^{-\eta(1-\Delta(e))} | a_l = a] \tag{7}
\]

—so that, intuitively, the weight of each pure strategy }$a$ is updated multiplicatively in proportion to its payoff }$\Delta(e)$, by an exponential amount controlled by }$\eta$.

Surprisingly, using this update rule with }$\eta \approx \Theta(\ln n^m)$ ensures that after }$T$ rounds, the average payoff regret vis-à-vis that of the best mixed strategies }$\bar{X}_l$

\[
\frac{1}{T} \sum_{t=1}^{T} \Delta(e | a_l \sim X_l^{(t)}) - \frac{1}{T} \sum_{t=1}^{T} \Delta(e | a_l \sim \bar{X}_l) \tag{8}
\]

converges to zero at a worst-case rate of roughly }$O(1/\sqrt{T})$, or as high as }$O(1/T)$ for }$\bar{\Delta}_{\text{min}} \to 1$, where }$\bar{\Delta}_{\text{min}} \leq \Delta(\cdot)$ is a prior lower-bound on the instantaneous shared payoff:

\textbf{Lemma 1}. Let }$\eta = 4(\ln n^m / (1 - \bar{\Delta}_{\text{min}}))^{1/2}$ and }$\bar{\Delta}_{\text{min}} \leq \Delta(\cdot) \in [0, 1]$. Then after }$T$ rounds, the average payoff regret under the Hedge update rule as described above is at most

\[
O \left( \sqrt{(1 - \bar{\Delta}_{\text{min}})/T} \right) + O(1/T). \tag{9}
\]

\textbf{Proof}: Follows immediately from Lemma 4 in Freund & Schapire (1996) —with gains instead of losses— applied to Corollary 4 in Freund & Schapire (1999) (extended to the }$m$ players case), and refined by the constant factor in (Stoltz & Lugosi 2005) to ensure convergence to a correlated equilibrium by also bounding ‘internal regret’ (cf. Cesa-Bianchi et al. 2006, §4.4).
The resemblance with the bound in Theorem 2 is not accidental: MA’s population dynamics reduce (with high probability) to a coordination game with Hedge update dynamics, so the payoff bound in Lemma 1 also implies a bound on the time-averaged expected fitness regret of populations evolved under MA. This claim will be made precise in the next section.

3 Analysis

In this section I restate and sketch the main results of this paper, Theorems 1 and 2. At a high level, they follow directly from Meir & Parkes (2015) and Chastain et al. (2014)’s groundbreaking reduction of evolution to coordination games using MW-like update rules. Omitted steps are given in §A.

3.1 Meir & Parkes (2015) and Chastain et al. (2014)’s Reduction

Recently, Meir & Parkes (2015) extended Chastain et al. (2014)’s observation that under GA\textsubscript{na}, the marginal allele density of locus $l$ after $T$ generations

$$\sum_{c \in C | a_l = a} \rho_c^{(T)},$$

is precisely the mixed strategy of player $l$ after $T > 0$ rounds of a coordination game similar to the one described in §2.3 above, except with two-sided payoffs/losses, and with the MW-like update rule —aka ‘polynomial weights’ or ‘PW’ (Meir & Parkes 2015, §2.3)—

$$X_l^{(T)}(a) \propto X_l^{(T-1)}(a)\mathbb{E}[(1 + \eta)\Delta(c) | a_l = a],$$

with the ‘learning rate’ $\eta = s$ set to the selection strength of the fitness landscape $s = \max f(\cdot) - \min f(\cdot)$, and the payoffs/losses normalized to $\Delta = (f - 1)/s \in [-1, 1]$:

**Lemma 2.** After $T$ generations, the marginal allele densities under GA\textsubscript{na} are identical to the mixed strategies after $T$ rounds of a two-sided coordination game with payoffs/losses $\Delta = (f - 1)/s$, and PW updates with ‘learning rate’ $\eta = s$.

**Proof:** Follows immediately from Corollary 1 in Meir & Parkes (2015), extended to the multiple genes case via Proposition A.3 in (ibid).

Furthermore, since the PW rule in (10) also ensures diminishing regret dynamics (cf. Meir & Parkes 2015, §5.1), this reduction immediately implies a regret bound on $f$ —albeit a less advantageous one than the one in Lemma 1, due to the inability to ‘tune’ $s$ in advance (cf. Corollary 3 in Meir & Parkes 2015):

**Lemma 3.** After $T$ generations, GA\textsubscript{na}’s time-averaged fitness regret is at most

$$s^2 + O(1/T).$$
Proof: Follows directly from Lemma 2 above and Corollary 3 in Meir & Parkes (2015).

Put in another way, GA na’s marginal population dynamics can be seen as a PW coordination game between genes where the strategies are the alleles, the shared payoffs/losses are the (centered, normalized) fitnesses of the chromosomes, and the ‘learning rate’ is the selection strength of the fitness landscape. As a result, PW’s diminishing average payoff regret also implies diminishing time-averaged fitness regret at rate \(O(1/T)\), albeit ultimately bounded by \(s^2\) (cf. Meir & Parkes 2015, §5).

In the next two subsections, I extend this reduction to GA and MA, respectively.

3.2 Extension to GA

Crucially, the reduction above can be extended to the case of GA with little modification.

To see why, note that under single-point crossover, the population dynamics (before mutation) are almost identical to the “SR dynamics” in Equation A.8 in Meir & Parkes (2015) —except for the inner summation term, which should now range over all pairs of homologous regions recombinable under single-point crossover (namely \(\emptyset, \{1,\ldots,m\}\), \(\{1\},\ldots,\{m\}\), \(\{1,2\},\ldots,\{3,\ldots,m\}\), and so forth) instead:

**Lemma 4.** For \(t > 0\), the population densities (before mutation) under GA are given by

\[
\tilde{p}_c^{(t+1)} = p_c \sum_{d \in \mathcal{C}} \frac{f(c_{J_1} \diamond d_{J_2})}{|\mathcal{R}|} \sum_{J \in \mathcal{R}} f(d_{J_1} \diamond c_{J_2}) p_{d_{J_1} \diamond c_{J_2}}^{(t)} \left( \frac{Z_f^{(t)}}{f(c)} \right)^2 + (1 - p_r) \frac{f(c) p_c^{(t)}}{Z_f^{(t)}},
\]

where \(\mathcal{R} = \{\emptyset, \{1,\ldots,m\}\}, \{1\},\ldots,\{m\}\), \(\{1,2\},\ldots,\{3,\ldots,m\}\), \(\ldots\) is the set of all pairs of recombinable homologous regions, and \(Z_f^{(t)} = \sum_{c \in \mathcal{C}} p_c^{(t)} f(c)\) is the expected fitness of chromosome \(c\) at generation \(t\).

**Proof:** Summing over all possible recombinations and normalizing gives the result.

From Lemma 4, it can be shown (cf. §A.1) that the marginal allele densities (after mutation) factor into the same form of Lemma A.2 in Meir & Parkes (2015) —albeit only with high probability, due to the distorting effect of mutation:

**Lemma 5.** For \(T > 0\), with high probability, GA’s marginal allele densities equal GA na’s.

**Proof:** See §A.1.

Lemma 5, in turn, readily implies Lemmas 2 and 3 above also apply to the case of GA —albeit, again, only with high probability. The result is Theorem 1 (restated here as Theorem 3):

**Theorem 3.** With high probability, GA’s time-averaged fitness regret is at most

\[
s^2 + O(1/T).
\]
Proof: Lemma 5 above implies Lemma 2 (and hence Lemma 3) applies to the case of GA (with high probability). The result follows.

In other words, GA’s population dynamics also reduce to a PW coordination game between genes. As a result, GA also follows $O(1/T)$ diminishing regret dynamics, bounded by $s^2$.

3.3 Extension to MA

In the case of MA, by a similar reasoning, note that its marginal population dynamics are precisely the update dynamics of the coordination game with exponentially-weighted (aka Hedge) update rules described above in §2.3. To see why, note that its population dynamics obey the following lemma (cf. Meir & Parkes 2015, Equation A.8):

**Lemma 6.** For $t > 0$, the population densities (before mutation) under MA are given by

$$p_{c}^{(t+1)} = p_r \sum_{d \in \mathcal{C}} \frac{1}{|\mathcal{R}|} \sum_{J \in \mathcal{R}} e^{-\eta(1-f(c_{J_1} \diamond d_{J_2}))} \frac{p_{c_{J_1} \diamond d_{J_2}}^{(t)} e^{-\eta(1-f(d_{J_1} \diamond c_{J_2}))}}{Z_{\text{exp}}^{(t)} e^{-\eta(1-f(c))}}$$

where $\mathcal{R} = \{\emptyset, \{1, \ldots, m\}, \{1\}, \{2, \ldots, m\}, \{1, 2\}, \{3, \ldots, m\}, \ldots\}$ is the set of all pairs of recombinable homologous regions, and $Z_{\text{exp}}^{(t)} = \sum_{c \in \mathcal{C}} p_{c}^{(t)} e^{-\eta(1-f(c))}$.

Proof: Summing over all possible recombinations and normalizing gives the result.

As a result, it can be shown (cf. §A.2) that the marginal allele densities (after mutation) obey the dynamics below (cf. Meir & Parkes 2015, Lemma A.2)—albeit, again, only with high probability, due to the distorting effect of mutation:

**Lemma 7.** With high probability, MA’s marginal allele densities equal

$$\frac{1}{Z_{\text{exp}}^{(T)}} \sum_{c \in \mathcal{C}} p_{c}^{(T)} e^{-\eta(1-f(c))}$$

Proof: See §A.2.

In turn, Lemma 7 readily implies that, with high probability, the marginal allele densities under MA are precisely the strategies obtained under the coordination game with Hedge updates as described in §2.3 (cf. Meir & Parkes 2015, Proposition A.3):

**Lemma 8.** With high probability, after $T > 0$ generations, the marginal allele densities under MA are identical to the mixed strategies after $T$ rounds of the one-sided coordination game described in §2.3, with payoffs $\Delta = f$, and Hedge updates with ‘learning rate’ $\eta$. 

8
Proof: By Lemma 7 (w.h.p.),
\[
\sum_{c \in \mathcal{C} | a_l = a} p_c^{(T+1)} = \frac{1}{Z^{(T)}} \sum_{c \in \mathcal{C} | a_l = a} p_c^{(T)} e^{-\eta(1-f(c))}
\]
\[
= \frac{1}{Z^{(T)}} \sum_{c \in \mathcal{C} | a_l = a} \left( \sum_{c \in \mathcal{C} | a_l = a} p_c^{(T)} p_c^{(T)} | a_l = a \right) e^{-\eta(1-f(c))}
\]
\[
= \frac{1}{Z^{(T)}} \left( \sum_{c \in \mathcal{C} | a_l = a} p_c^{(T)} \right) \sum_{c \in \mathcal{C} | a_l = a} e^{-\eta(1-f(c))} p_c^{(T)} | a_l = a
\]
\[
= \frac{1}{Z^{(T)}} \left( \sum_{c \in \mathcal{C} | a_l = a} p_c^{(T)} \right) \mathbb{E}[e^{-\eta(1-f(c))} | a_l = a],
\]
which equals the Hedge dynamics in (7) with \( \sum_{c \in \mathcal{C} | a_l = a} p_c^{(T+1)} = X_1^{(T)}(a) \).

We can then apply Hedge’s regret bound in Lemma 1 above to MA’s time-averaged fitness regret, which gives Theorem 2 (restated here as Theorem 4):

**Theorem 4.** Let \( \eta = 4(\ln |\mathcal{C}|/(1-\tilde{f}_{\min}))^{1/2} \) and \( \tilde{f}_{\min} \geq f(\cdot) \in [0, 1] \). Then with high probability, after \( T > 0 \) generations, MA’s time-averaged expected fitness regret is at most
\[
O \left( \sqrt{\frac{(1 - \tilde{\Delta}_{\min})/T}{T}} \right) + O(1/T).
\]

Proof: Lemma 8 implies Lemma 1 also holds for MA (w.h.p.) with \( \Delta = f \) and \( n^m = |\mathcal{C}|. \) The result follows.

In other words, MA’s population dynamics can be seen as a Hedge coordination game between genes where the payoffs are the fitnesses, and the ‘learning rate’ is the tunable parameter \( \eta \) —which, if set to its optimal rate \( \eta \approx \Theta(\ln |\mathcal{C}|) \), yields no regret dynamics —albeit at a slower worst-case convergence rate of \( O(1/\sqrt{T}) \). Note, however, that convergence may be as fast as \( O(1/T) \) for \( \tilde{f}_{\min} \to 1 \).

### 3.4 Taking Stock

The bounds in Theorems 1 and 2 show that, with high probability, MA’s time-averaged expected fitness reaches the best achievable local minimum at a worst-case rate \( O(1/\sqrt{T}) \) —or as high as \( O(1/T) \) for \( \tilde{f}_{\min} \to 1 \). GA, on the other hand, reaches a ‘glass ceiling’ bounded by the selection strength \( s \) —albeit at the much faster rate \( O(1/T) \). The performance gap between the two will be particularly salient for ‘difficult’ landscapes with high selection strength \( s \) (Figure 1, left). However, in the weak selection limit \( s \to 0 \), the ‘glass ceiling’ becomes immaterial, and GA reaches the best achievable local minimum at a much faster rate than MA (Figure 1, right). I note that evolutionary models often assume weak selection —if only for analytical convenience (cf. Bürger 2011, §3.4).

Also note that, as pointed out earlier, the MW regret bounds in Lemmas 1 and 3 are tight up to constant
3.5 A heuristic interpretation of $\eta^*$

Interpretively, the bound in Theorem 2 is rather unintuitive, since it only holds when $\eta$ is set to the somewhat opaque optimal value of $\eta^* \approx \Theta(\ln |C|)$. Nevertheless, I claim this order of magnitude admits a ‘natural’ interpretation: namely, the rate that (roughly) maximizes the entropy of the marginal allele densities subject to the information encoded in the fitness landscape — thus resulting in a maximally non-committal proxy for fitness that threads well the ‘thermal equilibrium’ between exploration and exploitation.

My argument is heuristic. Let $\tilde{X}_t = e^{-\lambda q(l)} / \sum_{l' \in \{1, \ldots, m\}} e^{-\lambda q(l')}$ be the MaxEnt density for some quantity $q(l)$, and let $X_t^{(T)} = \sum_{c \in C} p_c^{(T)}$ be the marginal allele densities after $T$ generations. Then by Lemma 2 in Schapire & Freund (1999), under MA with $\eta = \eta^*$, after $T > 0$ generations the relative entropy between the two is bounded by (w.h.p.)

$$\text{RE} \left( \tilde{X}_t \bigg| X_t^{(T)} \right) \leq \eta T \mathbb{E}[q(l)] - \frac{1 - \eta^2}{2\eta} \sum_{t=1}^T \mathbb{E}_{c \sim \mathcal{P}(0)} [f(c) | c_t = l] + \ln |C|.$$  

Therefore, on every generation $T > 0$, with high probability, the marginal allele densities can be seen as a MaxEnt density that best reflects the quantity

$$\mathbb{E}[q(l)] \approx (1 - O(1/\sqrt{T})) \left( \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{c \sim \mathcal{P}(0)} [f(c) | c_t = l] \right) + O(1/\sqrt{T}) \ln |C|.$$  

In other words, with high probability, the marginal allele densities approximately converge — at rate $O(1/\sqrt{T})$ — to the MaxEnt density that best reflects the expected time-averaged fitnesses for the allele.
This is not to say, however, that actual cultures do follow these dynamics. But it does show that they should from an information-theoretical point of view, insofar as they ensure the marginal allele densities approximately converge to an optimal encoding of the fitness landscape.

4 Experiments

Do these bounds hold in practice? As it turns out, the evidence is mixed.

4.1 Ms. Pac-Man Task

For my main experiment, I compared the performance of MA vs. GA on the task of evolving high-scoring controllers for Ms. Pac-Man: a challenging sequel to the original Pac-Man game with nondeterministic ghosts and complex scoring rules. For comparability purposes, I used the same parameters as in a similar experiment by Galván-López et al. (2010) except for: the selection method (which naturally changes), and the chromosome size (which was not reported therein).

The fitness landscape is defined by:

$$f(c) = \frac{1}{\max\{\text{Score}(\cdot)\}} \text{Score}(\text{GE}_G(c)),$$

where \(\text{Score} (\cdot)\) denotes the (normalized) score of the given controller on a single Ms. Pac-Man game against ‘legacy’ ghosts on a benchmark Ms. Pac-Man simulator (cf. Lucas 2010; Rohlfshagen et al. 2013), and the chromosomes \(c \in \mathcal{C} = \mathbb{Z}_{10}^{10} = \{0, \ldots, 2^{16} - 1\}\) are 16-bit unsigned integer vectors of length \(m = 10\). Controllers, in turn, are encoded using the ‘grammatical evolution’ (GE) scheme proposed in O’Neill & Ryan (2003): that is, as the result of a parameterized leftmost derivation (of length at most \(m \times d = 10 \times 3\)) on a given context-free grammar \(G\) such that, on the \(i\)-th rule application, the \((c(i \mod 10) \mod k)\)-th alternative out of \(k\) available is chosen — or a ‘null’ controller if the derivation length exceeds \(m\) times \(d = 3\) ‘wrap-arounds’.

\[
\begin{align*}
V_0 & \to B|BC \\
B & \to \text{if}(F\text{EF})\{V_0\} | \text{if}(F\text{EF})\{D\} \\
C & \to \text{else}\{D\} | \text{else}\{V_0\} \\
D & \to \text{go}_\text{to}(\text{pill}) | \text{go}_\text{to}(\text{ghost}) \\
E & \to < | > | == \\
F & \to d\text{.ghost} | d\text{.pill} | \text{dist} \\
\end{align*}
\]

Grammar \(\mathcal{G} = \{V, \Sigma, R, V_0\}\): 

\[
\begin{align*}
\text{Chromosome } c &= (15, 92, 33, 81, 17, 46): \\
15: & B \\
92: & \text{if}(F\text{EF})\{D\} \\
33: & \text{if}(\text{dist}\text{EF})\{D\} \\
81: & \text{if}(\text{dist}>F)\{D\} \\
17: & \text{if}(\text{dist}>\text{d\_pill})\{D\} \\
46: & \text{if}(\text{dist}>\text{d\_pill})\{\text{go}_\text{to}(\text{ghost})\} \\
\end{align*}
\]

Figure 2: Toy example of GE encoding. Adapted from Galván-López et al. (2010).

To fix ideas, consider the toy example in Figure 2, adapted from Galván-López et al. (2010). Given a grammar \(\mathcal{G} = \{V, \Sigma, R, V_0\}\) (left) and chromosome \(c = (15, 92, 33, 81, 17, 46)\) (right), we compute
GE\(_G(c)\) as follows:

- On the first rule application \(V_0 \rightarrow B|BC\), we choose the \((15 \mod 2)\)-st or first alternative —which yields
  \[GE\(_G(c)\) = B.\]

- On the second rule application \(B \rightarrow \text{if}(FEF)\{V_0\} | \text{if}(FEF)\{D\}\), we choose the \((92 \mod 2)\)-st or first alternative —which yields
  \[GE\(_G(c)\) = \text{if}(FEF)\{D\}.\]

- On the third rule application \(F \rightarrow d\_\text{ghost} \mid d\_\text{pill} \mid \text{dist}\), we choose the \((34 \mod 3)\)-st or third alternative —which yields
  \[GE\(_G(c)\) = \text{if}(\text{dist}EF)\{D\}.\]

- And so on until no non-terminals are left, ‘cycling’ through the chromosome up to \(d = 3\) times if necessary. However, if after 3 ‘wrap-arounds’ there are still non-terminals left, return NULL.

The actual grammar used is given in §B.1. Note the basis functions already encapsulate considerable high-level functionality, so the generalizability of these results to more general problem domains should not be overstated.

In addition to the fitness function \(f(\cdot)\) and chromosome set \(c \in C\) described above, I used the following parameters for both MA and GA —as per Galván-López et al. (2010, §4)—: initial random population of size \(|P_0| = 100\), number of generations \(T = 50\), mutation probability \(p_m = 0.1\), crossover probability \(p_r = 0.7\), and \(\bar{f}_{\text{min}} = 0\).

**Figure 3:** Empirical time-averaged (left) and ‘instantaneous’ (right) expected fitness of MA (blue) vs. GA (red) on the Ms. Pac-Man controller task (\(N = 100\) replications).

The results for \(N = 100\) Monte Carlo replications are given in Figure 3 (left). The evidence is mixed. Even though technically MA’s time-averaged expected fitness is ‘significantly’ higher than GA’s on every generation (one-sided KS test statistic \(D_{N,N}^+ \geq 0.66\), significant at the \(\alpha < 10^{-19}\) level), there
are important discrepancies with theory. Given the difficulty of the task, one would have expected a ‘jagged’ landscape (i.e. high selection strength \( s \)) and a low minimum achievable fitness \( \hat{f}_{\text{min}} \), which should result in both: (a) a large performance gap between MA and GA, and (b) a slower convergence rate of \( O(1/\sqrt{T}) \) for MA (as opposed to \( O(1/T) \) for GA). Instead, the (implied) gap or ‘glass ceiling’ in performance between MA and GA after \( T \) generations concentrates barely above 0.05; and the time-averaged expected fitnesses under GA do not seem to be following a \( O(1/T) \) law, but a \( O(1/\sqrt{T}) \) law instead that seems, at any rate, slower than MA’s.² These results strongly suggest that the leading constants in the bounds in Theorems 1 and 2 might be more significant than expected.

For comparison, the ‘instantaneous’ expected fitness curves are given in Figure 3 (right). Note, however, the bounds in Theorems 1 and 2 do not explicitly analyze this quantity —so I refrain from interpreting these results.

### 4.2 Sum-Of-Bits Task

For my second experiment, I compared the performance of MA vs. GA on the much simpler sum-of-bits task from CS63 last semester, with joint parameters \( c \in \mathcal{C} = \bigotimes_{i=1}^{15} \{0, 1\}, f(c) = f((a_1, \ldots, a_{15})) = \frac{1}{15} \sum_{i=1}^{15} a_i, |\mathcal{P}_0| = 50, T = 50, p_m = 0.01, p_r = 0.6 \) and \( \hat{f}_{\text{min}} = 0 \).

![Figure 4](image-url)

**Figure 4**: Empirical time-averaged (left) and ‘instantaneous’ (right) expected fitness of MA (blue) vs. GA (red) on the sum-of-bits task \((N = 100\) replications).

The results for 100 Monte Carlo replications are given in Figure 4 (left). Again, the results are mixed: even though MA’s time-averaged expected fitness does ‘significantly’ outperform GA’s on every generation (one-sided KS test statistic \( D_{N,N}^+ \geq 0.86 \), significant at the \( \alpha < 10^{-33} \) level), and there is some

²More concretely, fitting a \( \beta_1 - \beta_2 x^{-1/2} \) curve to each of the \( N = 100 \) time-averaged expected fitness curves under GA yields \( R^2 \) values in the (interquartile) range of \([0.94, 0.98]\), a mean of 0.96, and a max of 0.99; while fitting a \( \beta_1 - \beta_2 x^{-1} \) curve instead yields \( R^2 \) values in the (interquartile) range of \([0.47, 0.58]\), a mean of 0.52, and a maximum of 0.73.

Interestingly, in the case of MA, fitting a \( \beta_1 - \beta_2 x^{-1/2} \) curve to each of its \( N = 100 \) time-averaged expected fitness curves yields \( R^2 \) values in the (interquartile) range of \([0.96, 0.99]\), a mean of 0.98, and a maximum of 0.99; while fitting a \( \beta_1 - \beta_2 x^{-1} \) curve \( R^2 \) values in the (interquartile) range of \([0.78, 0.94]\), a mean of 0.89, and a maximum of 0.99—indicative perhaps of a non-trivial \( O(1/T) \) term. Note, however, these observations are largely heuristic and qualitative.
evidence of a ‘glass ceiling’ or gap between MA and GA; the apparent convergence rates do not match the theoretical predictions of MA converging slower than GA. For comparison, the (uninterpreted) ‘instantaneous’ expected fitness curves are given in Figure 4 (right).

5 Discussion and Open Problems

Theorems 1 and 2 above suggest that in ‘jagged’ landscapes far from the weak selection regime, the accelerated spread of memes in cultural evolution may improve its convergence properties vis-à-vis genetic evolution. Analytically, these results follow readily from the recent groundbreaking reduction of Meir & Parkes (2015) and Chastain et al. (2014) of evolution’s marginal population dynamics to the update dynamics of a coordination game between genes/memes with MW-like update rules. Interpretively, however, the reduction is hardly intuitive, the empirical evidence is mixed, and the semantics of the improvement mechanism are far from clear.

Can these bounds be improved? Certainly they could—and should—be restated in finite-sample terms—especially since the core analysis relies so heavily on the ‘true’ population densities $p_c$. Lemma 5 in Freund & Schapire (1999) provides a useful starting point here, as it bounds Hedge’s cumulative payoff deviations from their expected value by $O(1/\sqrt{T})$ (with high probability).

Another extension worth investigating further is translating the time-averaged regret bounds in Theorems 1 and 2 to the more natural (albeit much less studied) notion of instantaneous regret in (5). Note, however, this is generally unfeasible unless we are willing to either: (a) accept some kind of polynomially-discounted, time-averaged regret notion instead (cf. Cesa-Bianchi & Lugosi 2006, §2.11), or (b) make further assumptions about the fitness landscape (cf. Hazan & Seshadhri 2007). Having said that, the heuristic observation that MA also seems to (slightly) outperform GA in terms of ‘instantaneous’ regret on both the Ms. PacMan task and the sum-of-bits task (see also §B.2) should, at the very least, motivate further research attention on the connection between the two notions.

Also note that, for EC purposes, these bounds are largely impractical since they focus on the expected, time-averaged fitness of evolved populations, not its ‘instantaneous’ maximum. To get such a bound, techniques for studying the dynamics of second-order quantities under MW might be necessary—so the work in Cesa-Bianchi et al. (2007) might provide a useful starting point. Note, however, that these bounds rely on adaptive, time-varying learning rates $\eta_t$—which may be rather difficult to ‘motivate’ interpretively.

Lastly, from a sociobiological perspective, it is an open question whether or not meme transmission in actual cultures follows an exponential law with rate $\eta^* \approx \Theta(\ln |C|)$. Indeed, merely quantifying the value of $\eta$ in actual cultures seems like a formidable empirical task. Having said that, the normative interpretation proposed in §3.5 does suggest intriguing avenues of inquiry. For instance, is $\eta$ constant across the anthropological record? Does $\eta$ change across societies? If not, can we distinguish ‘developed’ from ‘developing’ societies by their differences in $\eta$? Or, in extremis, is societal collapse associated with sub-optimal values of $\eta$? To be able to address these issues quantitatively would surely be worth the empirical effort.
Appendix

A Proofs of Lemmas 5 and 7

The proofs below rely on the following crucial lemma:

**Lemma 9.** With high probability, after $T$ generations, the $\ell_2$ distortion induced by mutation on the marginal allele densities of both GA and MA vanishes with size.

**Proof:** To lighten the notation, I substitute $X_T$ for $\sum_{c \in C|a_1=a} \hat{p}^{(T)}_c$—so that $\hat{X}_T$ and $X_T$ denote the marginal allele densities of GA (or MA) before and after mutation, respectively.

Assume without loss of generality that $X_0 = \hat{X}_0 = 0$ (almost surely). Note for $T > 0$, the martingale differences $\frac{1}{2} \|(X_T - \hat{X}_T) - (X_{T-1} - \hat{X}_{T-1})\|_2$ are bounded a.s. by

$$\frac{1}{2} \|(X_T - \hat{X}_T) - (X_{T-1} - \hat{X}_{T-1})\|_2 = \frac{1}{2} \|(X_T - X_{T-1}) - (\hat{X}_T - \hat{X}_{T-1})\|_2$$

$$= \frac{1}{2} \|(X_T - X_{T-1}) - (\tilde{Q}X_T - \tilde{Q}X_{T-1})\|_2$$

$$= \frac{1}{2} \|(I - \tilde{Q})(X_T - X_{T-1})\|_2$$

$$\leq \frac{1}{2} \|(I - \tilde{Q})\|_2 \|(X_T - X_{T-1})\|_2$$

$$\leq (\frac{m}{m-1} - p_m) \cdot \min\{2, O(1/\sqrt{T})\}$$

$$\leq (1 - p_m) \cdot \min\{1, O(1/\sqrt{T})\}$$

$$\leq \min\{1, O(1/\sqrt{T})\}$$

$$\leq 1,$$

where $\tilde{Q} = \frac{1}{m} p_m J + (1 - \frac{m}{m-1} p_m) I$ is a $m \times m$ symmetric, doubly stochastic matrix with $\lambda_{\max}(\tilde{Q}) = p_m + (1 - \frac{m}{m-1})$ that encodes the effect of mutation.

Therefore, by extension of the Azuma-Hoeffding inequality to vector-valued martingales (cf. Theorem 7.4 in Hayes 2005), the probability that the $\ell_2$ distortion induced by mutation is greater than $\varepsilon > 0$ decreases exponentially with $\varepsilon$:

$$\Pr \left[ \|X_T - \hat{X}_T\|_2 \geq \varepsilon \right] < 2 e^{-(2\varepsilon - O(1/\sqrt{T}))^2},$$

which gives the result.

A.1 Proof of Lemma 5

Follows immediately from Lemma 9.

A.2 Proof of Lemma 7

The proof proceeds almost identically to that of Lemma A.2 in Meir & Parkes (2015). To lighten
the notation, I substitute $\alpha_c^{(T)}$ for $e^{-\eta(1-f(\cdot))} p_c^{(T)}$, and $J^l$ for $\{J_1 \cup \{l\}, J_2 \setminus \{l\}\}$.

By Lemmas 4 and 9, MA’s marginal allele densities $\sum_{c \in C} a_i = a p_c^{(T+1)}$ are given by (w.h.p.)

\[
\sum_{c \in C} a_i = a \left( p_m \sum_{d \in C} \frac{1}{|R|} \sum_{J \in R} \frac{\alpha_{c,J_1 \circ d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2} + (1 - p_m) \frac{\alpha_c^{(T)}}{Z_{exp}} \right)
\]

\[
= p_m \left( \frac{1}{|R|} \sum_{J \in R} \sum_{c \in C} a_i d_{J_k} \in J_k \sum_{J \in R} \frac{\alpha_{c,J_1 \circ d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2} \right) + (1 - p_m) \left( \frac{1}{Z^{(T)}_{exp}} \sum_{c \in C} a_i \right)
\]

where the first factor $\frac{1}{|R|} \sum_{J \in R} \sum_{c \in C} a_i d_{J_k} \in J_k \sum_{J \in R} \frac{\alpha_{c,J_1 \circ d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2}$ can be written as

\[
\frac{1}{\left( Z^{(T)} \right)^2} \left| R \right| \sum_{J \in R \setminus \{l\} \notin J} \left( \sum_{c \in C} a_i \sum_{d_{J_k} \in J_k} \frac{\alpha_{c,J_1 \circ d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2} + \sum_{c \in C} a_i \sum_{d_{J_k} \in J_k} \frac{\alpha_{c,J_1 \circ d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2} \right)
\]

\[
= \frac{1}{\left( Z^{(T)} \right)^2} \left| R \right| \sum_{J \in R \setminus \{l\} \notin J} \left( \sum_{c \in C} a_i \sum_{d_{J_2} \in J_2} \frac{\alpha_{d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2} + \sum_{c \in C} a_i \sum_{d_{J_2} \in J_2} \frac{\alpha_{d_{J_2} \circ c_{J_2}}^{(T)}}{\left( Z^{(T)} \right)^2} \right)
\]

\[
= \frac{1}{Z_{exp}^{(T)}} \sum_{J \in R \setminus \{l\} \notin J} \left( \sum_{c \in C} a_i \frac{\alpha_c^{(T)}}{\left( Z^{(T)} \right)^2} + \sum_{c \in C} a_i \frac{\alpha_c^{(T)}}{\left( Z^{(T)} \right)^2} \right)
\]

\[
= \frac{1}{Z_{exp}^{(T)}} \sum_{c \in C} a_i \frac{\alpha_c^{(T)}}{\left( Z^{(T)} \right)^2} \left( \frac{2}{\left| R \right|} \sum_{J \in R \setminus \{l\} \notin J} \frac{1}{\left( Z^{(T)} \right)^2} \right)
\]

\[
= \frac{1}{Z_{exp}^{(T)}} \sum_{c \in C} a_i \frac{\alpha_c^{(T)}}{\left( Z^{(T)} \right)^2}.
\]
Thus we have (w.h.p.)

\[
\sum_{c \in C | a_l = a} p_c^{(T+1)} = p_m \left( \frac{1}{Z^{(T)}_{\text{exp}}} \sum_{c \in C | a_l = a} \alpha_c^{(T)} \right) + \left( 1 - p_m \right) \left( \frac{1}{Z^{(T)}_{\text{exp}}} \sum_{c \in C | a_l = a} \alpha_c^{(T)} \right) \\
= \frac{1}{Z^{(T)}_{\text{exp}}} \sum_{c \in C | a_l = a} e^{-\eta(1-f(c))} p_c^{(T)} .
\]

B Further Experimental Details

B.1 Grammar

The full grammar used \( G = \{V, \Sigma, R, \text{<prog>}\} \) is given in Figure 5 below. It is essentially identical to the one used in Galván-López et al. (2010, Figure 2) except for slight differences in the method calls. The result is a valid JavaScript function with signature \( \text{get\_move}(g) \), where \( g \) is a reference to the main game object (of type \texttt{pacman.game.Game}) of the Java simulator used in Galván-López et al. (2010). The interfacing is implemented via Java 8’s Nashorn JavaScript interpreter, which allows seamless referencing of Java classes from within the interpreter’s environment.

\[
\begin{align*}
\texttt{<prog>} & \rightarrow \text{function get\_move}(g) \{
\text{var next = Java.type(’pacman.game.Constants’).MOVE.NEUTRAL);}
\text{var ctl = Java.type(’main.ge.pacman\_controller’);}
\text{var gt = <gt>, ws = <ws>, agd = <agd>};
\text{if(ctl.edible\_ghosts == 0)\{ <stat> } else\{ <stat> \}
\text{return next; }
\text{\}}
\text{<gt> } \rightarrow \text{1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20}
\text{<ws> } \rightarrow \text{3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19}
\text{<agd> } \rightarrow \text{1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20}
\text{<var> } \rightarrow \text{gt | ws | agd | ctl.adbg | ctl.igd}
\text{<cmp> } \rightarrow \text{< | > | <= | >= | ==}
\text{<le> } \rightarrow \text{< | <=}
\text{<act\_or\_stat> } \rightarrow \text{next = (<act>)(g); | <stat>}
\text{<act> } \rightarrow \text{ctl.npd | ctl.ang | ctl.ngd | (ctl.np>0)?ctl.npd:ctl.ang}
\text{<avc\_or\_ppc> } \rightarrow \text{avc | ppc}
\end{align*}
\]

Figure 5: Full grammar used in the main experiment. Adapted from Galván-López et al. (2010, Figure 2).
Note the resulting code calls upon static methods from the class \texttt{main.ge_pacman_controller}, which implements the high-level function basis \{adbg, ang, avc, igd, npd, ngd, nig, nppd\} as described (where applicable) in Galván-López \textit{et al.} (2010, cf. Table 1). Also note Table 1 in \textit{(ibid)} is missing the functions \texttt{adbg} (which computes the average distance between ghosts), \texttt{avc} (which avoids the closest ghost) and \texttt{igd} (which computes the distance to the closest inedible ghost).

### B.2 Simulator Details

Experiments were run on version 6.2 of the benchmark Java simulator used first in the IEEE CIG 2009 Ms. Pac-Man competition (cf. Lucas 2010; Rohlfshagen \textit{et al.} 2013). Note it differs slightly from the original Ms. Pac-Man in that it has no bonus ‘fruit’ pills and the speed of Ms. Pac-Man and the ghosts (when inedible) are identical.

For a given controller, the function \texttt{Score(\cdot)} described in §4.1 returns the unnormalized game score on a single level (max\{\texttt{Score(\cdot)}\} = 15420) of the controller against the built-in controllers.\texttt{Legacy} ghosts. Note the game is run in ‘asynchronous’ mode —that is, with no time limit for the controllers to make a move.

### B.3 Additional Figures

Figure 6 reports the best fitness of populations evolved under MA (\textit{blue}) and GA (\textit{red}) for the Ms. Pac-Man controller task (\textit{left}) and the sum-of-bits task (\textit{right}). Note, however, the bounds in Theorems 1 and 2 do \textit{not} explicitly control these quantities, so these results should be considered exploratory and heuristic at best.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{best_fitness}
\caption{Best fitnesses of MA (\textit{blue}) vs. GA (\textit{red}) on the Ms. Pac-Man controller task (\textit{left}) and the sum-of-bits task (\textit{right}) (\(N = 100\) replications).}
\end{figure}
References


