Analyzing Meme Propagation in Multimemetic Algorithms: Initial Investigations

Rafael Nogueras, Carlos Cotta
ETSI Informática, Universidad de Málaga
Campus de Teatinos, 29071 Málaga, Spain
Email: ccottap@lcc.uma.es

Abstract—Multimemetic algorithms (MMAs) are a subclass of memetic algorithms in which memes are explicitly attached to genotypes and evolve alongside them. We analyze the propagation of memes in MMAs with spatial structure. For this purpose we propose an idealized selecto-Lamarckian model that only features selection and local improvement, and study under which conditions good, high-potential memes can proliferate. We compare population models with panmictic and toroidal grids topology. We show that the increased takeover time induced by the latter is essential to improve the chances for good memes to express themselves in the population by improving their hosts, hence enhancing their survival rates.

I. INTRODUCTION

F OUR decades ago, Richard Dawkins [3] put forward the definition of meme in analogy to the biological concept of gene. Memes were broadly characterized as units of imitation, that is, ideas or pieces of knowledge that jump from brain to brain, striving and proliferating in some cases and becoming extinct in others. Even more interestingly, memes are not static objects but dynamic entities that mutate during their lifetime; these mutations can make them more strong/interesting/useful/... thus boosting their propagation, or can have the opposite effect, making that particular mutation fade away. This plasticity explains their comparatively faster rate of adaptation with respect to biological genes.

Inspired by this notion of meme, Moscato [13] conceived a new optimization paradigm: memetic algorithms (MAs). MAs are a family of population-based optimization techniques that blend together ideas of different metaheuristics, most notably the orchestrated interplay between global (population-based) search and local (individual-based) search. The most popular incarnation of MAs features an evolutionary search engine endowed with local search add-ons. The notion of memetic evolution is here captured by the Lamarckian lifetime learning. We show that the increased takeover time induced by the latter is essential to improve the chances for good memes to express themselves in the population by improving their hosts, hence enhancing their survival rates.

While memes are typically fixed in classical MAs (i.e., they are given by the particular choice of local search operators), several models trying to make them change during the optimization process have been proposed. This can be accomplished at a variety of levels. A simple possibility is the so-called ‘meta-lamarckian learning’ [18] in which the MA has a collection of local search operators (memes) available a priori, and some mechanism is used to decide which of them is applied on which solution and when (notice the connection with hyperheuristics [2]). A more complex approach features self-adaptation of the memes themselves. An example of this kind of self-adaptation is provided by multi-memetic algorithms (MMAs), in which each solution carries "genes" that indicate which local search has to be applied on it. These can range from simple pointers to existing local search operators to the parametrization of a general local search template [10] or even to the definition of a grammar to specify new complex local search operators [9], [11].

An interesting issue that arises in the context of MMAs is how memes propagate and spread over the population. While population dynamics has been well-studied in the case of evolutionary algorithms – e.g., [1], [6], [19], [20], the scenario is more complex in the case of memes: unlike genotypes (which correspond to solutions and thus can be evaluated according to the problem under consideration), memes can be only indirectly assessed via the effect they have on genotypes. Furthermore, memes evolve in MMAs alongside with solutions by attaching to them. Since this attachment is part of the self-adaptive process, it is up to the algorithm to discover good fits between individual pairs of genotypes and memes, and this is commonly done using only information about the genetic quality of solutions (i.e., fitness information). This work is aimed to study how memes propagate in such an environment driven by genetic selection and spatial structure. To this end, we consider and analyze an idealized model of MMAs. This model is described in next section.

II. MODEL DESCRIPTION

Let us consider an abstract model of MMAs in which each agent is characterized by a pair \((g, m)\) \(\in D^2\), for some \(D \subseteq \mathbb{R}\). The first member of the pair \(-g-\) represents the genotype, which we equate to fitness for simplicity. As to the second member \(-m-\) it represents a meme. More precisely, this value captures the improvement potential of that meme, that is,
a measure of how good solutions can get by applying the meme. We assume there is a monotonically increasing function \( f : D^2 \rightarrow D \) encapsulating the application of a meme to a genotype, i.e., the effect of a single epoch of Lamarckian learning. Thus, an agent \((g, m)\) becomes \((f(g, m), m)\) after the application of the meme, where

\[
\begin{align*}
\lim_{n \to \infty} f^n(g, m) &= m & \text{if } g < m \\
f(g, m) &= g & \text{if } g \geq m
\end{align*}
\] (1)

Here \( f^n(g, m) \) is the \( n \)-th composition of the function on the first argument, i.e., \( f(f(\cdots f(f(g, m), m), \cdots), m) \). It must be noted that while this is very idealized characterization of the potential of a meme (since in general this potential is not going to be absolute but may depend on a complex match between the meme, the genotype and the problem landscape) it serves as an initial approximation to study several issues related to meme propagation in the agent pool.

The population \( P \) of the MMA is thus a collection of \( \mu \) such agents, \( P = \{(g_1, m_1), \cdots, (g_\mu, m_\mu)\} \), endowed with a spatial structure that constrains agent communication. Let this spatial structure be characterized by a \( \mu \times \mu \) Boolean matrix \( S \), where \( S_{ij} \) is true if, and only if, the agent placed in the \( i \)-th location can communicate with the agent placed in the \( j \)-th location. Since we are interested in observing the dynamics of propagation of memes, we consider a selection of the selection-only model of evolutionary algorithms (i.e., using only selection/replacement and no variation operator) in which we add the local improvement stage of memetic algorithms. A scheme of the model is shown below in Algorithm 1.

Algorithm 1: Selecto-Lamarckian Model

```algorithm
for \( i \in [1 \cdots \mu] \) do
    INITIALIZE\((g_i, m_i)\);
end

while \( \neg \) CONVERGED\((P)\) do
    \( i \leftarrow \text{URAND}\(1, \mu\) \) // Pick random location
    \( (g, m) \leftarrow \text{SELECTION}\(P, S, i\); \)
    \( g' \leftarrow f(g, m) \) // Local improvement
    \( P \leftarrow \text{REPLACE}\(P, S, i, g', m)\); \)
end
```

After initializing suitably the contents of the population, the algorithm engages on a cycle of selection plus improvement until the agents converge. Convergence is here approached from a memetic perspective, that is, we terminate the algorithm when the population comprises a homogeneous collection of memes (regardless of whether there is still diversity at the genetic level or not). As to the inner working of the algorithm, it resembles the uniform choice update strategy of cellular automata [21], in which the next location to be activated is selected uniformly at random with replacement.

### III. MEME PROPAGATION

Having defined the general model in the previous section, let us consider some qualitative features of meme propagation that can be extracted from it. Let us assume that selection is done by binary tournament, i.e., once a location \( i \) selected, a neighboring location \( j \) is selected from \( N(i) = \{j \mid S_{ij}\} \), and the agent with the best fitness is retained. As to replacement, let us assume that the improved agent replaces the agent that lost the previous tournament.

We are interested in analyzing the number of copies of each meme in the meme pool, so let us denote by \( N(m, g, t) \) the number of instances of meme \( g \) attached to genotype \( m \) at time \( t \), assuming for simplicity that \( D \) is some discrete domain. If we divide this quantity by the pool size \( \mu \) we obtain \( p(m, g, t) \), the fraction of the population comprising meme \( g \) at genotype \( m \) at time \( t \). In each passing iteration of the system the number of copies can be estimated as

\[
N(m, g, t+1) = N(m, g, t) + C(m, g, t) - D(m, g, t)
\] (3)

where \( C(m, g, t) \) and \( D(m, g, t) \) represent the expected number of copies of meme \( g \) that are created or destroyed at time \( t \). The creation of a new copy can be accomplished by the combined effect of selection of a suitable agent with meme \( m \) and the application of the meme to the corresponding genotype. Let us express this as:

\[
C(m, g, t) = \sum_{g'} \sigma(m, g', t)p(m, g', t) \rightarrow g
\] (4)

where \( \sigma(m, g', t) \) is the probability of selecting an agent carrying meme \( m \) and genotype \( g' \) at time \( t \) and \( p(m, g', t) \rightarrow g \) is the probability that the application of meme \( m \) on genotype \( g' \) results in genotype \( g \). The first quantity can be computed as the probability that the binary tournament picks two agents with meme \( m \) and genotype \( g \) or only one agent with this structure but better fitness than its competitor:

\[
\sigma(m, g, t, p) = p(m, g, t)^2 + 2\left[p(m, g, t)(1 - p(m, g, t))\right] \cdot \sum_{g'} \sum_{g' < g} p(m', g', t) \frac{1 - p(m, g, t)}{1 - p(m, g, t)}
\] (5)

where the last factor is the probability that the fitness of the competitor is worse than \( g \) provided it is not a \((m, g)\) agent. This expression assumes that the global distribution of memes/genotypes across the whole population is the same as for local neighborhoods. Obviously, this holds for the panmictic case in which any two agents are neighbors so we can assume this case initially, and consider it a first approximation to more general situations.

As to the destruction of a copy of a particular pair meme/genotype, it can arise via the selection of such a pair and the subsequent application of local improvement (which will alter the genotype) or via replacement by an agent of higher fitness. The first case also requires that the other agent chosen in the tournament be a copy of the same pair, so that it is later substituted by the improved agent. Thus,

\[
D(m, g, t) = \sum_{g' \neq g} p(m, g, t)^2 \sigma(m, g, t)
\] (6)
The replacement probability $\tilde{\sigma}(m, g, t)$ can be expressed as:

$$\tilde{\sigma}(m, g, t) = 2 \{ p(m, g, t) [1 - p(m, g, t)] \} .$$

Let us consider the evolution of the system in the early-term and mid-term, before a particular meme starts to saturate the population. In this situation memes are widely spread across genotypes, so $p(m, g, t) \ll 1$, so we can take quadratic terms $p(m, g, t)^2$ as approximately 0 and terms $1 - p(m, g, t)$ as approximately 1. We thus have:

$$\sigma(m, g, t) = 2p(m, g, t) \sum_{m'} \sum_{g' < g} p(m', g', t)$$

$$\tilde{\sigma}(m, g, t) = 2p(m, g, t) \sum_{m'} \sum_{g' > g} p(m', g', t)$$

Substituting back into Eqs. (4) and (6) we get:

$$C(m, g, t) = 2 \sum_{g', m'} p(m, g', t) \sum_{g''} p(m', g'', t) p(g' \rightarrow g)$$

$$D(m, g, t) = 2p(m, g, t) \sum_{m'} \sum_{g' > g} p(m', g', t)$$

Since $p(g' \rightarrow g) = 0$ for $g < g'$ or $m < m$, Eq. (10) reduces to:

$$C(m, g, t) = 2 \sum_{m'} \sum_{g' < g} \sum_{g''} p(m', g'', t) p(m, g', t) p(g' \rightarrow g)$$

If $m \leq g$ then $p(g' \rightarrow g) = 1$ if $g' = g$ and 0 otherwise. Subsequently, the difference $\Delta(m, g, t) = C(m, g, t) - D(m, g, t)$ is in this case:

$$\Delta(m, g, t) = 2p(m, g, t) \sum_{m'} \sum_{g'' < g} p(m', g'', t) - 2p(m, g, t) \sum_{m'} \sum_{g'' > g} p(m', g'', t)$$

$$\Delta(m, g, t) = 2p(m, g, t) \cdot \sum_{m'} \left[ \sum_{g'' < g} p(m', g'', t) - \sum_{g'' > g} p(m', g'', t) \right]$$

Focusing on the sign of the difference in the above expression, we essentially obtain that inert memes (i.e., memes that can no longer improve their hosts) can strive by hitchhiking, that is, if they attach to agents above the median of the population.

Let us consider the other side of the population. The replacement probability $\tilde{\sigma}(m, g, t)$ can be expressed as:

$$\tilde{\sigma}(m, g, t) = 2 \{ p(m, g, t) [1 - p(m, g, t)] \} .$$

For $f^{-1}(g, m)$, we essentially obtain that

$$\Delta(m, g, t) = 2p(m, g^{-m}, t) \sum_{m'} \sum_{g'' < g^{-m}} p(m', g''', t) - 2p(m, g, t) \sum_{m'} \sum_{g' > g} p(m', g', t)$$

$$\Delta(m, g, t) = \sum_{m'} \left[ 2p(m, g^{-m}, t) \sum_{g'' < g^{-m}} p(m', g''', t) - 2p(m, g, t) \sum_{g' > g} p(m', g', t) \right] .$$

The sign of this expression depends on the balance between the goodness of genotypes in the basin of attraction of $g$ and the badness of $g$ itself (in both cases goodness/badness relative to the rest of the population). Notice that in general there is a reinforcement between these quantities in the sense that the better a genotype in the basin of attraction of $g$, the better we can expect $g$ to be. This does not just mean that active memes proliferate more and more when they attach to good solutions as one would expect, but also that memes with high potential can find their way to the final stages of the evolution provided they have enough time to improve their hosts (recall that the goodness of solutions evolves with time as an effect of the application of the meme). This suggests that models with slower genetic convergence can have a beneficial effect on the propagation of good memes, allowing the latter enough time to express themselves in the population and overcome the hitchhiking effect of bad memes. Next section provides a more quantitative analysis of this effect via numerical simulations.

### IV. Numerical Simulations

The numerical experimentation is aimed to explore empirically the dynamics of meme propagation and how it is affected by factors such as the population size, the relative improvement potential of memes and the underlying spatial structure of the population. Regarding population sizes, we have considered values $\mu \in \{100, 256, 400, 625\}$. These values cover a broad range of population sizes and also perfect squares, which is important in connection with one of the spatial structures considered, namely a square toroidal grid with von Neumann neighborhood: two locations $(i, j)$ and $(i', j')$ are connected if their Manhattan distance $|i - i'| + |j - j'| \leq r$, where $r$ is the neighborhood radius. We have considered $r = 1$ which leads to the traditional North-South-East-West (plus the current location) neighborhood. The other spatial structure considered is the panmictic model in which all locations are connected. In either case, we have considered the function

$$f(g, m) = \begin{cases} 
    g & \text{if } g \geq m \\
    \frac{(g + m)}{2} & \text{if } g < m
\end{cases}$$

1015

to represent the action of memes. Intuitively, this function provides smaller improvements for increasingly good genotypes much like often happens in practice. All experiments
are averaged over 100 runs in order to obtain representative results. Each run is terminated upon convergence of memes, which for simplicity is determined when all memes are equal to 2 decimal positions.

Let us firstly analyze meme propagation as a function of the relative improvement potential of memes at the beginning of the run. For this purpose, we take $D = [0, 1]$ and consider a scenario in which genotypes and memes are randomly initialized in this range, and another scenario in which genotypes take initial values in $[0, 0.5]$ whereas memes are randomly sampled from $[0, 1]$. Figure 1 shows the distribution of memes at each time step (the lighter the shade of gray, the higher the meme value). Focusing firstly in the upper row (panmictic topology), notice the clearly different behavior depending on genotype initialization. When genes and memes are both initialized in $[0, 1]$ the algorithm does seldom converge to a high-potential meme. Actually, such memes temporarily proliferate in the initial stages of the algorithm but are later driven to extinction by memes hitchhiking on high quality genotypes to which they stucked by chance. The situation is quite different when genotypes are initially drawn from $[0, 0.5]$; in this case the algorithm does gradually converge to the upper part of the meme distribution, with low-potential memes quickly disappearing from the population. A more detailed perspective on this is provided by Figure 2 in which qualified runtime distributions (QRTDs) [8] are provided. These indicate the probability that a certain target (in this case, convergence to a meme in a desired percentile) is reached as a function of the number of iterations. Notice how the probabilities are below 10% for memes above the 95% percentile in the first scenario, whereas this probability is 100% in the second scenario. In the latter a spurious match between a very good genotype and a bad meme cannot happen since these very good solutions do not exist initially. Furthermore, high-potential memes initially
Fig. 2. Qualified runtime distributions for simulations with \( \mu = 625 \). The upper row corresponds to panmictic connectivity and the lower row to von Neumann neighborhood. Similarly, the left column corresponds to genotypes initialized in \([0, 1]\) and the right one to initialization in \([0, 0.5]\) (memes are initialized in \([0, 1]\) in both cases). The curves indicate the probability that the population converges to a meme in the \( i \)-th initial percentile of the population as a function of the number of iterations. Notice the different scale in the X-axis.

At this point, attaching to bad genotypes can highly improve the quality of the latter in the initial steps, thus increasing their chances of survival.

Let us now turn our attention to the effect of the spatial structure. The bottom row of Figure 1 shows the distribution of memes for the case of von Neumann neighborhood. Notice how a similar pattern as in the panmictic case is observed with respect to genotype initialization. A more detailed inspection indicates several differences though. Firstly, notice how the convergence is slower in this case (e.g., the scale in the X-axis is larger). This is a well-known effect of the use of spatial structure and is commonly exploited in the context of evolutionary algorithms for promoting diversity and thus decreasing the chances of getting stuck in local optima [4], [23]. In the case of MMAs this has an additional advantage, namely the fact that a slower convergence increases the lifespan of individual memes, thus giving them more chances to improve their hosts if they have the potential to do so. Hence, the algorithm is more robust and can better cope with hitchhiking inert memes. This can be seen in the meme map in the bottom row of Figure 1 by a larger prevalence of lighter-gray areas, and more clearly in the QRTDs (bottom row of Figure 2), e.g., the 95% percentile is reached with nearly 20% probability in the case of \([0, 1]\)-initialization (cf. below 10% in the panmictic case), and the 99% percentile is reached with nearly 80% probability for \([0, 0.5]\)-initialization (cf. about 65% in the panmictic case). A signrank test [24] indicates that the difference in the final percentile reached is statistically significant in both cases (\( \alpha = .05 \)).

Finally, we consider the takeover time, namely the time required for a meme (not necessarily the best one as shown previously) to dominate the population. Figure 3 shows the growth curves, depicting the percentage of the meme pool occupied by the most repeated meme (notice that the most repeated meme need not be the same throughout a run; we simply count the number of copies of the most repeated one
at each time step). These curves exhibit the typical shape of the well-known logistic model $f(t) = \frac{1}{1 + Ke^{-t/\alpha}}$. Indeed, such a model was proposed early in the literature by Sarma and De Jong [20] in the context of spatially structured evolutionary algorithms. While by no means the unique alternative – e.g., see [5]– it serves as a good starting approximation to quantify the growth of the dominant meme. Qualitatively, we observe as expected the well-known pattern of slower convergence for increasing population sizes and for the von Neumann topology [6] as opposed to the panmictic population. From a quantitative point of view, we have fitted the growth data to a logistic curve to identify the scale factor $\alpha$ that renders the number of iterations dimensionless. The resulting data is shown in Table I. As it can be seen, the fit is quite good, yielding very low mean squared errors. The scale parameters are quite similar for variants with the same topology, and are about 2-5 times larger for the von Neumann topology than for the panmictic case, in correspondence with the relative takeover time which

### Table I

<table>
<thead>
<tr>
<th>Topology</th>
<th>$G$</th>
<th>$\mu = 100$</th>
<th>$\mu = 256$</th>
<th>$\mu = 400$</th>
<th>$\mu = 625$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panmictic</td>
<td>0.5</td>
<td>81.628878</td>
<td>0.000042</td>
<td>210.791430</td>
<td>0.000008</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>77.708284</td>
<td>0.000008</td>
<td>187.904342</td>
<td>0.000030</td>
</tr>
<tr>
<td>Von Neumann</td>
<td>0.5</td>
<td>186.815036</td>
<td>0.000366</td>
<td>578.914610</td>
<td>0.000554</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>168.567695</td>
<td>0.000322</td>
<td>585.804249</td>
<td>0.000532</td>
</tr>
</tbody>
</table>

Fig. 3. Growth curves for different population sizes. The upper row corresponds to panmictic connectivity and the lower row to von Neumann neighborhood. Similarly, the left column corresponds to genotypes initialized in [0, 1] and the right one to initialization in [0, 0.5] (memes are initialized in [0,1] in both cases). Notice the different scale in the X-axis.
can be seen in Figure 3. With respect to the population size, the increase in the scale parameter admits a linear interpolation $\alpha = a + b \gamma$ yielding values of $b = 0.84$ and $b = 0.74$ for the panmictic case and $b = 3.43$ and $b = 3.40$ for grid topology with von Neumann connectivity.

V. CONCLUSIONS

We have presented some initial steps in the line of analyzing meme propagation in MMAs. Using an idealized model of genotypes and memes we have shown that the selection intensity plays a very important role in allowing high-potential memes to proliferate. In a panmictic model, good memes will dominate the final population when the starting solutions have a substantial improvement margin on average. When this margin is smaller, average memes can hitchhike their way to the final stages of the evolution and make other comparatively better memes become extinct. In the presence of a spatial structure inducing longer takeover times (in our case a toroidal square grid with von Neumann topology), this hitchhiking effect is somewhat mitigated, allowing good memes to express themselves and increasing their chances for making it to the final population. An interesting line of future research focuses on the consideration of other topologies and study their effect on meme propagation. Work is in progress in this area. Looking beyond, another topic for further research is the extension of this analysis to coevolutionary memetic algorithms [22] in which memes are detached from genotypes and co-evolve alongside the latter in a separate population.

REFERENCES


