A Genetic Model: Analysis and Application to MAXSAT

Alberto Bertoni
Dipartimento di Scienze dell'Informazione
Università degli Studi di Milano
via Comelico, 39-20135 Milano, Italy
bertoni@dsi.unimi.it

Paola Campadelli
Dipartimento di Scienze dell'Informazione
Università degli Studi di Milano
via Comelico, 39-20135 Milano, Italy
campadelli@dsi.unimi.it

Marco Carpentieri
Dipartimento di Scienze dell'Informazione
Università degli Studi di Milano
via Comelico, 39-20135 Milano, Italy
marcar@udsab.dia.unisa.it

Giuliano Grossi
Dipartimento di Scienze dell'Informazione
Università degli Studi di Milano
via Comelico, 39-20135 Milano, Italy
grossi@dsi.unimi.it

Abstract

In this paper, a genetic model based on the operations of recombination and mutation is studied and applied to combinatorial optimization problems.

Results are:

1. The equations of the deterministic dynamics in the thermodynamic limit (infinite populations) are derived and, for a sufficiently small mutation rate, the attractors are characterized;
2. A general approximation algorithm for combinatorial optimization problems is designed. The algorithm is applied to the Max Ek-Sat problem, and the quality of the solution is analyzed. It is proved to be optimal for $k \geq 3$ with respect to the worst case analysis; for Max E3-Sat the average case performances are experimentally compared with other optimization techniques.

Keywords

Genetic algorithms, Markov chains, dynamical systems, combinatorial optimization, approximation algorithms, Hopfield's networks.

1 Introduction

Genetic algorithms are probabilistic search algorithms inspired by mechanisms of natural selection. They have received considerable attention because of their applications to fields such as optimization, adaptive control, and others (Goldberg, 1989; Holland, 1989, 1992; Koza, 1992). By simplifying natural laws, genetic algorithms simulate reproductive processes over a population of individuals or genotypes, typically represented by binary strings of fixed length l, in an arbitrary environment.

A classic way to describe the behavior of genetic algorithms is obtained by means of homogeneous Markov chains (Eiben et al., 1991; Goldberg and Segrest, 1987; Nix and Vose, 1992) whose states encode populations that are multisets of binary words. General theoretical results are available in the thermodynamic limit (for *infinite populations*) when the systems become deterministic; these results are also useful for the qualitative analysis of the behavior of the finite populations model (Vose, 1996). Unfortunately, simulation of the deterministic system is computationally difficult since the states of the system are represented by vectors in $\mathbb{R}^{2^l}$, where l is the length of the words that represent the individuals.

Our aim is to introduce and analyze a model such that:

1. It preserves most of the properties of the classical genetic systems;
2. The states of the system for infinite populations are vectors in $\mathbb{R}^l$ instead of $\mathbb{R}^{2^l}$;
3. The simulation of the system for infinite populations is a good local optimization algorithm for the fitness function.

In Section 2, we present the model that is based on two stochastic rules: *recombination* and *mutation*.

In recombination, also called "bit-based simulated crossover" in Syswerda (1993), a weighted average of the alleles of the individuals along each bit position is done; these statistics are used to produce offspring, whose alleles in different positions are independently generated.

This rule, like the *gene pool recombination* described in Mühlenbein and Voigt (1996), maintains the population in *linkage equilibrium*: the genotype frequencies are the product of the univariate marginal frequencies. Therefore, the vector of frequencies (element in $\mathbb{R}^{2^l}$) can be reconstructed by means of the vector of marginal frequencies (element in $\mathbb{R}^l$). Other models based on the marginal frequencies are presented in Baluja and Caruana (1995) and Harik et al. (1997). The use of this recombination rule, instead of the biologically inspired one-point crossover, seems to present advantages in the simulation efficiency, and it makes the analysis easier. As far as the fitness optimization is concerned, Mühlenbein (1998) conjectures that the class of fitness functions solvable with models based on recombination is very similar to that based on classical crossover. In the mutation, random changes in the genotypes are introduced (Goldberg, 1989). This rule ensures that Markov chains, which formally describe the model, become ergodic, admitting therefore stationary distributions.

In Section 3, we present a probability concentration result (for large populations) on the Markov chains describing the evolution of the genetic system (Theorem 3.1). This result allows us to derive an iterative deterministic system with states in $[0,1]^l$, which represents the behavior of the genetic system with infinite population (Equations (5)). This system extends the equations derived in Mühlenbein and Voigt (1996) and Bertoni et al. (1996), taking into account the mutation rate.

In Section 4, the attractors of the iterative system are characterized both in the case of a system without mutation (Theorem 4.1) and with small mutation (Theorem 4.4). For quadratic fitness functions, the attractors are interpreted as equilibrium points of the discrete Hopfield's network (Hopfield, 1982) whose energy is the fitness function.

In Section 5, we derive from the genetic model a general algorithm (GENETIC) for finding approximate solutions of a large class of hard combinatorial optimization problems.

In Section 6 we analyze the quality of the solutions of GENETIC on the problem Max Ek-Sat from both theoretical and experimental points of view. In particular, as regarded as the worst case analysis, we prove a bound $1 - 2^{-k}$ to the performance ratio for all k; for $k \geq 3$ the algorithm is optimal because of the result of Håstad (1997).

Finally, the average case performances of GENETIC on Max E3-Sat have been experimentally evaluated and compared with those of other optimization techniques (general or specialized for Max Ek-Sat) such as local search and non-oblivious local search, simulated annealing, some variants of the tabu search, the Johnson algorithms, and a heuristic called Walksat. The comparison shows that GENETIC performs worse than some specialized heuristics, but it is the best among the general techniques.

2 The Genetic Model

Given an arbitrary, fixed integer $n > 0$, a population P is a multiset of n elements of Ω, where $\Omega = \{0,1\}^l = \{\omega_1, \ldots, \omega_{2^l}\}$ is the class of length l binary strings. The population P can be represented by the "frequency vector" $\mathbf{F} = (F_{\omega_1}, \ldots, F_{\omega_{2^l}})$, where $F_{\omega_k} = \frac{n_k}{n}$ and n_k is the number of occurrences of the word ω_k in P. Let Λ_n denote the set of the frequency vectors that represent the populations of n elements. Notice that Λ_n is a finite subset of the *simplex*

$$\Lambda = \left\{ (\pi_1, \ldots, \pi_{2^l}) : \pi_i \geq 0, \sum_{i=1}^{2^l} \pi_i = 1 \right\}$$

and $\bigcup_{n=1}^{\infty} \Lambda_n$ is a dense subset of Λ.

Given a function $X : \Omega \to \mathbb{R}$ and a stochastic vector $\Pi \in \Lambda$, let us denote by $E_\Pi[X] = \sum_{i=1}^{2^l} X(\omega_i)\pi_i$ the expectation of X. Let $f : \Omega \to \mathbb{N}^+$ be the fitness function and $x_k : \Omega \to \{0,1\}$ be defined as

$$x_k(\omega) = \begin{cases} 0, & \text{if the } k\text{-th bit of } \omega \text{ is } 0; \\ 1, & \text{if the } k\text{-th bit of } \omega \text{ is } 1. \end{cases}$$

The evolution of our genetic system can be described by the following steps:

a. At time 0, the state of the system is the initial population P_0;

b. If at time t the state of the system is the population P, represented by the frequency vector $\mathbf{F}$, then the population at time $t+1$ is obtained as follows:

   1. Calculate the ratio

   $$\phi_{k\mathbf{F}} = \frac{E_\mathbf{F}[x_k \cdot f]}{E_\mathbf{F}[f]}, \qquad \text{for } k = 1, \ldots, l;$$

   2. Generate $\{\omega_{s_1}, \ldots, \omega_{s_n}\}$ with probability $\phi_{k\mathbf{F}}$ to obtain 1 in position k, independently from s_i and k, for $1 \leq i \leq n$ and $1 \leq k \leq l$;
   3. Flip k-th bit of word ω_{s_i} with probability $0 < \eta \leq \frac{1}{2}$ for any $1 \leq k \leq l$ and $1 \leq i \leq n$.

We call the stochastic rules described in steps 2 and 3 *recombination* and *mutation*, respectively. By applying recombination and mutation, the probability of generating words with 1 in position *k* is

$$\begin{aligned} g_{k\mathbf{F}} &= \phi_{k\mathbf{F}}(1-\eta) + \eta(1-\phi_{k\mathbf{F}}) = \phi_{k\mathbf{F}}(1-2\eta) + \eta \\ &= \frac{E_{\mathbf{F}}\left[(1-2\eta)(x_k \cdot f) + \eta f\right]}{E_{\mathbf{F}}[f]}. \end{aligned} \tag{1}$$

If P is a population at time t, and $\mathbf{F}$ its frequency vector, then the population at time $t+1$ is obtained by selecting n words with probability distribution $\mathcal{G}(\mathbf{F}) = (\mathcal{G}(\mathbf{F})_{\omega_1}, \ldots, \mathcal{G}(\mathbf{F})_{\omega_{2^l}})$, where the probability $\mathcal{G}(\mathbf{F})_{\omega_j}$ of obtaining the word $\omega_j = x_1 \cdots x_l$ is

$$\mathcal{G}(\mathbf{F})_{\omega_j} = \prod_{k=1}^{l} \left[x_k g_{k\mathbf{F}} + (1-x_k)(1-g_{k\mathbf{F}})\right].$$

Notice that $\mathcal{G} : \Lambda_n \rightarrow \Lambda$ can be extended by continuity to a function $\mathcal{G} : \Lambda \rightarrow \Lambda$, that, with abuse of notation, we denote by the same symbol.

Now, we describe this model by means of a homogeneous Markov chain, following the guidelines presented in Vose (1996). Identifying the populations with the corresponding frequency vectors, the states of the Markov chain are the elements of Λ_n. The probability $Q_{\mathbf{F}\mathbf{F}'}$ that the system evolves from a population $\mathbf{F}$ to a population $\mathbf{F}'$ is given by

$$Q_{\mathbf{F}\mathbf{F}'} = n! \prod_{j=1}^{2^l} \frac{\left[\mathcal{G}(\mathbf{F})_{\omega_j}\right]^{n\mathbf{F}'_{\omega_j}}}{(n\mathbf{F}'_{\omega_j})!}.$$

An important consequence of the mutation is that if $\eta > 0$ then the Markov chain is ergodic, admitting therefore a stationary distribution. For large n, by Stirling formula, the following asymptotic expression holds

$$Q_{\mathbf{F}\mathbf{F}'} = e^{-n \cdot d(\mathbf{F}', \mathcal{G}(\mathbf{F})) + O(\log n)},$$

where d is the Kullback-Leibler distance defined by

$$d(\mathbf{F}, \mathbf{F}') = \sum_{i=1}^{2^l} \mathbf{F}_{\omega_i} \ln \frac{\mathbf{F}_{\omega_i}}{\mathbf{F}'_{\omega_i}}.$$

For $n \rightarrow \infty$ the state space of the Markov chain becomes dense in Λ. Therefore, in case of infinite populations, (*infinite population model*) the state space is Λ, and the system becomes deterministic since $Q_{\mathbf{F}\mathbf{F}'} = \delta(\mathbf{F}' - \mathcal{G}(\mathbf{F}))$. In this case, the dynamic is

$$\mathbf{F}(t+1) = \mathcal{G}(\mathbf{F}(t)). \tag{2}$$

The infinite population model is interesting on its own, and as shown in Vose (1996), it is useful even to derive qualitative properties for the case of finite populations.

3 Infinite Population Model

The main drawback of System (2) is that its states space consists of 2^l-components vectors; for instance, this makes simulations intractable except for very small values of l. In this section, we show that many properties of the behavior of (2) can also be derived by a system whose states space consists of l-components vectors. In this regard, the following property of the Markov chain describing the evolution of the genetic system is useful.

THEOREM 3.1: *Let M be the maximum value of the fitness function $f : \Omega \to \mathbb{N}^+$ and suppose $n \geq \frac{32}{9}\frac{1}{\varepsilon^2}\left[\frac{(1-\eta)M}{E_{\mathcal{G}(\mathbf{F})}[f]}\right]^2 \log\frac{4l}{\delta}$. If, at time t, the system is in the state $\mathbf{F}$, then the state $\mathbf{F}'$ at time $t+1$ is such that for all k*

$$|g_{k\mathbf{F}'} - g_{k\mathcal{G}(\mathbf{F})}| \leq \varepsilon$$

with probability at least $1-\delta$ $(0 < \delta \leq 1,\ 0 < \varepsilon \leq \frac{1}{3})$.

PROOF: Given a state $\mathbf{F}$, consider the probability space $\langle \Omega^n, \mathcal{G}^{(n)}(\mathbf{F})\rangle$, where $\Omega^n = \Omega \times \cdots \times \Omega$ and

$$\mathcal{G}^{(n)}(\mathbf{F})_{\omega_{s_1}\cdots\omega_{s_n}} = \prod_{j=1}^{n} \mathcal{G}(\mathbf{F})_{\omega_{s_j}}.$$

An element of Ω^n selected according to $\mathcal{G}^{(n)}(\mathbf{F})$ is a n-tuple $(\omega_{s_1}, \ldots, \omega_{s_n})$ of words in Ω independently selected according to $\mathcal{G}^{(n)}(\mathbf{F})$. Let X_k, X be the random variables defined as:

$$X_k(\omega_{s_1}, \ldots, \omega_{s_n}) = \frac{1}{n}\sum_{j=1}^{n}\left[(1-2\eta)(x_k \cdot f)(\omega_{s_j}) + \eta f(\omega_{s_j})\right],$$

$$X(\omega_{s_1}, \ldots, \omega_{s_n}) = \frac{1}{n}\sum_{j=1}^{n} f(\omega_{s_j}).$$

Let $(\overline{\omega}_{s_1}, \ldots, \overline{\omega}_{s_n})$ be a n-ple selected according to $\mathcal{G}^{(n)}(\mathbf{F})$ and $\mathbf{F}'$ be the state associate with it; by definition of the random variables X_k, X, and by Equation (1) it can be easily verified that

$$g_{k\mathbf{F}'} = \frac{X_k(\overline{\omega}_{s_1}, \ldots, \overline{\omega}_{s_n})}{X(\overline{\omega}_{s_1}, \ldots, \overline{\omega}_{s_n})} \quad \text{and} \quad g_{k\mathcal{G}(\mathbf{F})} = \frac{E_{\mathcal{G}(\mathbf{F})}\left[(1-2\eta)(x_k \cdot f) + \eta f\right]}{E_{\mathcal{G}(\mathbf{F})}[f]}.$$

In this setting, our problem is to estimate $\mathbf{Pr}\left\{\left|\frac{X_k}{X} - g_{k\mathcal{G}(\mathbf{F})}\right| \leq \varepsilon\right\}$. First of all, we observe that:

$$E_{\mathcal{G}^{(n)}(\mathbf{F})}[X_k] = E_{\mathcal{G}(\mathbf{F})}\left[(1-2\eta)(x_k \cdot f) + \eta f\right],$$

$$E_{\mathcal{G}^{(n)}(\mathbf{F})}[X] = E_{\mathcal{G}(\mathbf{F})}[f].$$

Let us denote $A_k = E_{\mathcal{G}(\mathbf{F})}\left[(1-2\eta)(x_k \cdot f) + \eta f\right]$, $A = E_{\mathcal{G}(\mathbf{F})}[f]$, and $\alpha = 1 - \eta$; observe that:

$$0 \leq X \leq M, \qquad 0 \leq X_k \leq \alpha M, \qquad A_k \leq \alpha A.$$

For all nonnegative numbers $\gamma \leq \frac{1}{4}$,

$$\text{if} \quad |X_k - A_k| \leq \alpha\gamma A \quad \text{and} \quad |X - A| \leq \gamma A, \quad \text{then} \quad \left|\frac{X_k}{X} - \frac{A_k}{A}\right| \leq \frac{8}{3}\alpha\gamma. \tag{3}$$

In fact, since $\frac{A_k}{A} \leq \alpha$ and $\frac{1}{1-\gamma} \leq \frac{4}{3}$, we have

$$\frac{X_k}{X} \leq \frac{A_k + \gamma\alpha A}{A - \gamma A} = \frac{A_k}{A} + \left(\frac{A_k}{A} + \alpha\right)\frac{1}{1-\gamma}\gamma \leq \frac{A_k}{A} + \frac{8}{3}\alpha\gamma.$$

Analogously, it can be shown that $\dfrac{X_k}{X} \geq \dfrac{A_k}{A} - \dfrac{8}{3}\alpha\gamma.$

Let $\varepsilon = \frac{8}{3}\alpha\gamma$. Then, by Equation (3) and by Hoeffding's inequality[1], for $\varepsilon \leq \frac{1}{3}$ it holds that

$$\begin{aligned}\mathbf{Pr}\left\{\left|\frac{X_k}{X} - \frac{A_k}{A}\right| \geq \varepsilon\right\} &\leq \mathbf{Pr}\left\{|X_k - A_k| \geq \frac{3}{8}\varepsilon A\right\} + \mathbf{Pr}\left\{|X - A| \geq \frac{3\varepsilon}{8\alpha}A\right\} \\ &\leq 4e^{-\frac{9}{32}\left(\frac{\varepsilon A}{\alpha M}\right)^2 n}.\end{aligned}$$

Therefore:

$$\mathbf{Pr}\left\{\bigvee_{k=1}^{l}\left|\frac{X_k}{X} - \frac{A_k}{A}\right| \geq \varepsilon\right\} \leq 4le^{-\frac{9}{32}\left(\frac{\varepsilon A}{\alpha M}\right)^2 n}.$$

With δ fixed $(0 < \delta \leq 1)$, we can conclude that if $n \geq \frac{32}{9}\frac{1}{\varepsilon^2}\left[\frac{(1-\eta)M}{A}\right]^2 \log\frac{4l}{\delta}$ then

$$\mathbf{Pr}\left\{\bigwedge_{k=1}^{l}\left|\frac{X_k}{X} - \frac{A_k}{A}\right| \leq \varepsilon\right\} \geq 1 - \delta.$$

□

The previous theorem provides a distribution dependent bound. However, observing that $f(\cdot) \geq 1$ implies $E_{\mathcal{G}(\mathbf{F})}[f] \geq 1$, the following distribution independent bound can be stated:

COROLLARY 3.1: *Let M be the maximum value of the fitness function $f : \Omega \rightarrow \mathbb{N}^+$, and suppose $n \geq \frac{32}{9}\left[\frac{(1-\eta)M}{\varepsilon}\right]^2 \log\frac{4l}{\delta}$. If, at time t, the system is in the state $\mathbf{F}$, then the state $\mathbf{F}'$ at time $t+1$ is such that for all k*

$$|g_{k\mathbf{F}'} - g_{k\mathcal{G}(\mathbf{F})}| \leq \varepsilon$$

with probability at least $1 - \delta$ $(0 < \delta \leq 1,\ 0 < \varepsilon \leq \frac{1}{3})$.

Now consider an arbitrary pseudo-Boolean function $h : \Omega \rightarrow \mathbb{N}$; it can be represented by a multivariate polynomial of at most one degree in each variable defined on $[0,1]^l$ and coincident with h on Ω, that is:

$$h(x_1, \ldots, x_l) = \sum_{y_1,\ldots,y_l \in \{0,1\}} \alpha_{y_1\cdots y_l} x_1^{y_1} \cdots x_l^{y_l}$$

For the sake of simplicity and with abuse of notation, we denote with h both the function and the associated polynomial. Notice that, since h is a polynomial of at most one degree in each variable, global maxima of h are on elements of $\{0,1\}^l$.

[1] Let $X_1, \ldots, X_n$ be independent bounded random variables such that X_i falls in the interval $[a_i, b_i]$ with probability one; let $S_n = \sum_{i=1}^{n} X_i$ and $E[S_n]$ be its expectation. Then, for any $t > 0$, we have that $\mathbf{Pr}\{S_n - E[S_n] \geq t\} \leq e^{-2t^2/\sum_{i=1}^{n}(b_i - a_i)^2}$.

Since in our genetic model x_i and x_j $(i \neq j)$ are independent, the following relation holds:

$$E_{\mathcal{G}(\mathbf{F})}[h] = h(g_{1\mathbf{F}}, \ldots, g_{l\mathbf{F}}). \qquad (4)$$

In fact:

$$\begin{aligned} E_{\mathcal{G}(\mathbf{F})}[h] &= \sum_{y_1,\ldots,y_l \in \{0,1\}} \alpha_{y_1\cdots y_l} \prod_{j=1}^{l} E_{\mathcal{G}(\mathbf{F})}[x_j]^{y_j} \\ &= \sum_{y_1,\ldots,y_l \in \{0,1\}} \alpha_{y_1\cdots y_l} \cdot g_{1\mathbf{F}}^{y_1} \cdots g_{l\mathbf{F}}^{y_l} \\ &= h(g_{1\mathbf{F}}, \ldots, g_{l\mathbf{F}}). \end{aligned}$$

Given the fitness function $f : \Omega \to \mathbb{N}^+$, for every k $(1 \leq k \leq l)$ we can rewrite f as

$$f(\mathbf{x}) = x_k b_k(\mathbf{x}) + a_k(\mathbf{x}),$$

where $b_k(\mathbf{x})$ and $a_k(\mathbf{x})$ are polynomials that do not depend on x_k and whose variables have at most one degree. Observe that $b_k(\mathbf{x}) = \frac{\partial}{\partial x_k} f(\mathbf{x})$ and $a_k(\mathbf{x}) = f(\mathbf{x}) - x_k b_k(\mathbf{x})$.

As a consequence of Theorem 3.1 and by Equation (4), we can conclude that for large n with high probability

$$\begin{aligned} g_{k\mathbf{F}'} \approx g_{k\mathcal{G}(\mathbf{F})} &= \frac{E_{\mathcal{G}(\mathbf{F})}[x_k \cdot f]}{E_{\mathcal{G}(\mathbf{F})}[f]}(1 - 2\eta) + \eta \\ &= \frac{E_{\mathcal{G}(\mathbf{F})}[x_k(b_k + a_k)]}{E_{\mathcal{G}(\mathbf{F})}[f]}(1 - 2\eta) + \eta \\ &= \frac{g_{k\mathbf{F}}[b_k(\mathbf{g_F}) + a_k(\mathbf{g_F})]}{f(\mathbf{g_F})}(1 - 2\eta) + \eta, \end{aligned}$$

where $\mathbf{g_F} = (g_{1\mathbf{F}}, \ldots, g_{l\mathbf{F}})$.

Therefore, in case of infinite populations, the stochastic genetic system becomes the iterative deterministic system whose states are vectors $\mathbf{g} = (g_1, \ldots, g_l) \in [0,1]^l$ and whose dynamics is described by the equations

$$g_k(t+1) = g_k(t)\frac{b_k(\mathbf{g}(t)) + a_k(\mathbf{g}(t))}{g_k(t)b_k(\mathbf{g}(t)) + a_k(\mathbf{g}(t))}(1 - 2\eta) + \eta \qquad k = 1, \ldots, l. \qquad (5)$$

4 Analysis

In this section, the model without mutation $(\eta = 0)$ is analyzed and, for quadratic fitness functions, a relation between the model and the discrete Hopfield's networks is shown. Moreover, for sufficiently small mutation rate, a characterization of the attractors of the dynamical Systems (5) is given.

Case without Mutation $(\eta = 0)$

If $\eta = 0$, then Equations (5) can be rewritten in the form

$$\Delta g_k = g_k(t+1) - g_k(t) = g_k(t)(1 - g_k(t))\frac{b_k(\mathbf{g}(t))}{f(\mathbf{g}(t))} \qquad k = 1, \ldots, l. \qquad (6)$$

Observe that all points in Ω are fixed points for the dynamical Systems (6). Given a fitness function $f : \Omega \to \mathbb{N}^+$, the following results state, under general conditions, which points in Ω are attractors (i.e., asymptotically stable) of (6).

THEOREM 4.1: *Every element* $\mathbf{x} \in \Omega$ *is a fixed point for the iterative System* (6)*; if* $b_k(\mathbf{x}) \neq 0$ *for all* k, *then* $\mathbf{x} \in \Omega$ *is an attractor for System* (6) *if and only if*

$$x_k = HS\,(b_k(\mathbf{x})) \qquad k = 1, \ldots, l,$$

where $HS\,(x) = \begin{cases} 1, & \textit{if } x \geq 0; \\ 0, & \textit{otherwise}. \end{cases}$

PROOF: Let $\mathbf{x} \in \Omega$. It immediately follows from (6) that $\mathbf{x}$ is a fixed point ($\Delta g_k = 0$, for all $k = 1, \ldots, l$). To determine which points in Ω are asymptotically stable we linearize the System (6) in a neighborhood of $\mathbf{x}$ by calculating

$$\frac{\partial}{\partial g_j} \left(g_k(1 - g_k) \frac{b_k(\mathbf{g})}{f(\mathbf{g})} \right) \Bigg|_{\mathbf{g}=\mathbf{x}} = \begin{cases} 0, & \text{if } j \neq k; \\ (1 - 2x_k)\frac{b_k(\mathbf{x})}{f(\mathbf{x})}, & \text{otherwise.} \end{cases}$$

The linearized system in a neighborhood of $\mathbf{x}$ is described by the following set of equations

$$\varepsilon_k(t+1) = \left[1 + (1 - 2x_k) \frac{b_k(\mathbf{x})}{f(\mathbf{x})} \right] \varepsilon_k(t) \qquad k = 1, \ldots, l.$$

The point $\mathbf{x}$ is asymptotically stable iff

$$\left| 1 + (1 - 2x_k) \frac{b_k(\mathbf{x})}{f(\mathbf{x})} \right| < 1 \qquad k = 1, \ldots, l.$$

Since $f(\mathbf{x}) = x_k b_k(\mathbf{x}) + a_k(\mathbf{x}) > 0$ and $a_k(\mathbf{x}), b_k(\mathbf{x}) + a_k(\mathbf{x}) > 0$, it follows that the inequality $1 + (1 - 2x_k)\frac{b_k(\mathbf{x})}{f(\mathbf{x})} > -1$ is alway true. Then $\mathbf{x}$ is asymptotically stable iff $(1 - 2x_k)b_k(\mathbf{x}) < 0$ for all $k \in \{1, \ldots, l\}$, i.e., when $x_k = \text{HS}\,(b_k(\mathbf{x}))$. □

An important consequence is the following:

COROLLARY 4.1: *Let* $\tilde{\mathbf{x}} = \arg\max\limits_{\mathbf{x} \in \Omega} f(\mathbf{x})$*; then* $\tilde{\mathbf{x}}$ *is an attractor of System* (6).

PROOF: Let $M = f(\tilde{\mathbf{x}})$. Since $f(\mathbf{x}) = x_k b_k(\mathbf{x}) + a_k(\mathbf{x})$, and $a_k(\mathbf{x})$ and $b_k(\mathbf{x})$ do not contain the variable x_k, then $0 < a_k(\tilde{\mathbf{x}}) \leq M$, for all $k = 1, \ldots, l$.

Let us suppose that $\tilde{\mathbf{x}}$ is not an attractor of System (6); by Theorem 4.1, there exists k such that $\tilde{x}_k b_k(\tilde{\mathbf{x}}) < 0$. This implies that

$$0 > \tilde{x}_k b_k(\tilde{\mathbf{x}}) = f(\tilde{\mathbf{x}}) - a_k(\tilde{\mathbf{x}}) = M - a_k(\tilde{\mathbf{x}}),$$

contradicting the inequality $a_k(\tilde{\mathbf{x}}) \leq M$. □

Theorem 4.1 gives a characterization of the points in $\{0, 1\}^l$ that are attractors. Since $b_k(\mathbf{g}) = 0$ implies $\Delta g_k = 0$ even if $g_k \notin \{0, 1\}$, the genetic system can have fixed points $\mathbf{p} \notin \Omega$. The following theorem states that these points can not be attractors.

THEOREM 4.2: *Let* $\mathbf{p} \notin \Omega$ *a fixed point of System* (6), *then* $\mathbf{p}$ *is not asymptotically stable.*

PROOF: Let us consider a fixed point $\mathbf{p}$ such that $p_k \notin \{0, 1\}$ for all $k = 1, \ldots, l$ (with minor modifications the proof holds also for the other cases). The linearization of the System (6) gives:

$$\varepsilon_k(t+1) = \sum_{j=1}^{l} J_{kj}\varepsilon_j(t) \qquad\qquad k = 1, \dots, l,$$

where J is the Jacobian matrix $(J_{kj})_{l\times l}$ with entries

$$J_{kj} = \frac{\partial}{\partial g_j}\left(g_k + g_k(1-g_k)\frac{b_k(\mathbf{g})}{f(\mathbf{g})}\right)\Bigg|_{\mathbf{g}=\mathbf{p}}.$$

Since, $\dfrac{\partial b_k}{\partial g_k} = 0$ for all $k = 1, \dots, l$ and $b_k(\mathbf{p}) = 0$, the matrix J has all diagonal entries equal to 1.

Let us denote with $\mu_1, \dots, \mu_l$ the eigenvalues of the matrix J, and let $\mathrm{Tr}(J)$ denote its trace; if $\mathbf{p}$ were an asymptotically stable point, then $|\mu_i| < 1$ for all $i = 1, \dots, l$. This implies the following contradiction:

$$l = \mathrm{Tr}\,(J) = \sum_i \mu_i \leq \sum_i |\mu_i| < l.$$

□

Relation with Hopfield's Neural Networks

Let us now consider quadratic fitness function $f : \Omega \to \mathbb{N}^+$ of the kind

$$f(x_1, \dots, x_l) = \sum_{i<j} w_{ij}x_ix_j + \sum_i \lambda_i x_i + C.$$

In this case, the attractors of the genetic system can be interpreted as equilibrium points of a suitable discrete Hopfield network (Hopfield, 1982).

A Hopfield's network of l neurons with states in $\{0, 1\}$ is described by $\langle Z, \boldsymbol{\gamma}\rangle$, where $Z = (z_{ij})_{l\times l}$ is a symmetric matrix with $z_{ii} = 0$ $(i = 1, \dots, l)$, and $\boldsymbol{\gamma}$ is an l-dimensional vector. The matrix Z is called the *weight matrix* and the vector $\boldsymbol{\gamma}$ is the *threshold vector*. Let $S_i(t)$ be the state of the neuron i at time t, then the asynchronous dynamics is described by the equations

$$S_i(t+1) = \mathrm{HS}\left(\sum_{j=1}^{i-1} z_{ij}S_j(t+1) + \sum_{j=i+1}^{l} z_{ij}S_j(t) - \gamma_i\right) \qquad i = 1, \dots, l. \tag{7}$$

Every network $\langle Z, \boldsymbol{\gamma}\rangle$ can be associated with the following state function called *energy*

$$E(S_1, \dots, S_n) = \frac{1}{2}\sum_{i\neq j} z_{ij}S_iS_j - \sum_i \gamma_i S_i.$$

When the network is made to evolve according to the asynchronous dynamics, the energy function is a Lyapunov function, that is, $E(\mathbf{S}(t+1)) \geq E(\mathbf{S}(t))$ and $E(\mathbf{S}(t+1)) = E(\mathbf{S}(t))$ if and only if $\mathbf{S}(t+1) = \mathbf{S}(t)$. As a consequence, every network trajectory $\{\mathbf{S}(t)\}_{t\geq 0}$ ends up in an equilibrium point, i.e., a state $\mathbf{S}(\bar{t})$ such that $\mathbf{S}(\bar{t}) = \mathbf{S}(\bar{t}+1)$.

A straightforward consequence of Theorem 4.1 is the following:

THEOREM 4.3: *Let $f : \Omega \rightarrow \mathbb{N}^+$ be a quadratic fitness function*

$$f(\mathbf{x}) = \sum_{i<j} w_{ij} x_i x_j + \sum_i \lambda_i x_i + C$$

and suppose $b_k(\mathbf{x}) = \sum_{j=1}^{l} w_{kj} x_j + \lambda_k \neq 0$ for all k, then $\mathbf{x} \in \Omega$ is an attractor for System (6) *if and only if it is an equilibrium point for the Hopfield's neural network with energy function f.*

PROOF: By Theorem 4.1, $\mathbf{x} \in \Omega$ is an attractor of System (6) if and only if

$$x_k = \mathrm{HS}\left(\sum_{j=1}^{l} w_{kj} x_j + \lambda_k\right) \qquad k = 1, \ldots, l.$$

Hopfield's network with energy function f is described by $\langle Z, \gamma \rangle$, where $z_{kj} = w_{kj}$ and $\gamma_k = -\lambda_k$. For this network, $\mathbf{x}$ is an equilibrium point if and only if $\mathbf{x}(t+1) = \mathbf{x}(t)$, i.e.,

$$x_k = \mathrm{HS}\left(\sum_{j=1}^{l} z_{kj} x_j - \gamma_k\right) = \mathrm{HS}\left(\sum_{j=1}^{l} w_{kj} x_j + \lambda_k\right) \qquad k = 1, \ldots, l.$$

□

Case with Mutation ($\eta > 0$)

We now analyze the effect of small mutation on the genetic system described above. For $\varepsilon > 0$, let us denote by $\mathcal{I}(\mathbf{x}, \varepsilon)$ the ε-neighborhood of $\mathbf{x}$ in $[0,1]^l$.

THEOREM 4.4: *There exists $\varepsilon > 0$ such that, for sufficiently small $\eta > 0$:*

1. *if $\mathbf{x} \in \Omega$ is an attractor of System* (6) *and $b_k(\mathbf{x}) \neq 0$ for all k, then in $\mathcal{I}(\mathbf{x}, \varepsilon)$ there is a single fixed point of the genetic System* (5) *and such point is asymptotically stable;*

2. *if $\mathbf{x} \in \Omega$ is not an attractor of System* (6) *and $b_k(\mathbf{x}) \neq 0$ for all k, $\mathcal{I}(\mathbf{x}, \varepsilon)$ does not contain fixed points of the genetic System* (5).

PROOF: By few manipulations of Equations (5), the dynamics of the system with mutation rate $\eta > 0$ is described by the set of equations

$$g_k(t+1) = \eta + (1 - 2\eta)\left(g_k(t)(1 - g_k(t))\frac{b_k(\mathbf{g}(t))}{f(\mathbf{g}(t))} + g_k(t)\right) \quad k = 1, \ldots, l.$$

Then, the equilibrium points satisfy

$$g_k = \eta + (1 - 2\eta)\left(g_k(1 - g_k)\frac{b_k(\mathbf{g})}{f(\mathbf{g})} + g_k\right) \quad k = 1, \ldots, l. \tag{8}$$

Notice that, if $\eta = 0$, Equations (8) become Equations (6), therefore, as stated in Theorem 4.1, any $\mathbf{x} \in \{0,1\}^l$ is an equilibrium point of this system.

Fixed an arbitrary $\mathbf{x} \in \{0,1\}^l$, System (8) implicitly defines, in a suitable neighborhood of $\mathbf{x}$, a function $\mathbf{g}(\eta)$ such that $\mathbf{g}(0) = \mathbf{x}$. In fact, let $F_k(\eta, g_1, \ldots, g_l) = \eta + (1-\eta)\left(g_k(1-g_k)\frac{b_k(\mathbf{g})}{f(\mathbf{g})} + g_k\right) - g_k$. Then:

$$F_k(0, \mathbf{x}) = 0, \quad \frac{\partial F_k}{\partial \eta}(0, \mathbf{x}) = 1 - 2x_k \neq 0 \quad \text{for all } k = 1, \ldots, l,$$

and

$$\det\left(\frac{\partial F_k}{\partial g_j}\right) = \prod_{k=1}^{l} \frac{b_k(\mathbf{x})}{f(\mathbf{x})}(1 - 2x_k) \neq 0.$$

Now consider the following first order Taylor approximation:

$$g_k(\eta) = x_k + g'_k(0)\eta + \mathcal{O}(\eta^2), \quad \text{where} \quad g'_k(0) = \tfrac{d}{d\eta} g_k(\eta)\Big|_{\eta=0}.$$

By System (8), we have

$$g'_k(0) = 1 - 2x_k + g'_k(0) + (1 - 2x_k)\frac{b_k(\mathbf{x})}{f(\mathbf{x})} g'_k(0),$$

from which it follows $g'_k(0) = -\frac{f(\mathbf{x})}{b_k(\mathbf{x})}$ and

$$g_k(\eta) = x_k - \frac{f(\mathbf{x})}{b_k(\mathbf{x})}\eta + \mathcal{O}(\eta^2).$$

Let us consider the two cases.

1. The point $\mathbf{x}$ is an attractor of System (6), that is, $x_k = \text{HS}\ (b_k(\mathbf{x}))$ for all $k = 1, \ldots, l$. If $x_k = 1$, then $b_k(\mathbf{x}) > 0$ and, therefore, for η sufficiently small

$$0 < g_k(\eta) = 1 - \left|\frac{f(\mathbf{x})}{b_k(\mathbf{x})}\right| \eta + \mathcal{O}(\eta^2) < 1.$$

   If $x_k = 0$, then $b_k(\mathbf{x}) < 0$, and for η sufficiently small we have

$$0 < g_k(\eta) = \left|\frac{f(\mathbf{x})}{b_k(\mathbf{x})}\right| \eta + \mathcal{O}(\eta^2) < 1.$$

   Thus, if η is sufficiently small there exists $\varepsilon > 0$ such that $\mathbf{g}(\eta) \in \mathcal{I}(\mathbf{x}, \varepsilon)$, where $\mathcal{I}(\mathbf{x}, \varepsilon)$ is an ε-neighborhood of $\mathbf{x}$ in $[0,1]^l$. Notice that when $\eta = 0$, the point $\mathbf{x}$ is asymptotically stable. Therefore, the module of each eigenvalue of the matrix that defines the linearized System (8) is less than 1. By continuity, if η is sufficiently small, the module of the eigenvalues of the matrix of the linearized system is less than 1, and consequently, the point $\mathbf{g}(\eta)$ is asymptotically stable.

2. The point $\mathbf{x}$ is not an attractor of System (6), that is, $x_k \neq HS(b_k(\mathbf{x}))$ $(1 \leq k \leq l)$. If $x_k = 1$, then $b_k(\mathbf{x}) < 0$ and

$$g_k(\eta) = 1 + \left|\frac{f(\mathbf{x})}{b_k(\mathbf{x})}\right| \eta + \mathcal{O}(\eta^2) > 1.$$

If $x_k = 0$, then $b_k(\mathbf{x}) > 0$, and we have

$$g_k(\eta) = -\left|\frac{f(\mathbf{x})}{b_k(\mathbf{x})}\right|\eta + \mathcal{O}(\eta^2) < 0.$$

We can conclude that $\mathbf{g}(\eta) \notin [0,1]^l$ $(\eta \neq 0)$. □

5 An Approximation Algorithm for NP-hard Problems

In this section, we introduce an asynchronous variant of the genetic System (5), and we suggest that it can be used as a general technique to find approximate solutions of NP-hard problems (Ausiello et al., 1995).

Given a fitness function $f : \Omega \to \mathbb{N}^+$, as before, with abuse of notation, we also denote f as the unique extension of f to $[0,1]^l$ defined by means of a multivariate polynomial of at most one degree in each variable. Remember that $b_k(\mathbf{x}) = \frac{\partial}{\partial x_k} f(\mathbf{x})$ and $a_k(\mathbf{x}) = f(\mathbf{x}) - x_k b_k(\mathbf{x})$.

The asynchronous variant of the genetic System (5) is the dynamical system obtained by updating the component of the state vector one at a time in a predefined order. Formally, for each $k = 1, \ldots, l$:

$$\begin{aligned} g_k(t+1) = (1-2\eta)g_k(t)&\left[\frac{b_k(g_1(t+1),\ldots,g_{k-1}(t+1),g_{k+1}(t),\ldots,g_l(t))}{f(g_1(t+1),\ldots,g_{k-1}(t+1),g_k(t),\ldots,g_l(t))}\right. \\ &\left.+\frac{a_k(g_1(t+1),\ldots,g_{k-1}(t+1),g_{k+1}(t),\ldots,g_l(t))}{f(g_1(t+1),\ldots,g_{k-1}(t+1),g_k(t),\ldots,g_l(t))}\right] + \eta. \end{aligned} \tag{9}$$

Observe that the fixed points of Systems (5) and (9) are coincident. Moreover, if $\eta = 0$, f is a Lyapunov function for the dynamical System (9). In fact:

LEMMA 5.1: *If* $\Delta g_k = g_k(1-g_k)\frac{b_k(\mathbf{g})}{f(\mathbf{g})}$, *then*

$$f(g_1,\ldots,g_k+\Delta g_k,\ldots,g_l) \geq f(g_1,\ldots,g_k,\ldots,g_l).$$

PROOF:

$$\begin{aligned} f(g_1,\ldots,g_k+\Delta g_k,\ldots,g_l) &= (g_k+\Delta g_k)b_k(g_1,\ldots,g_l) + a_k(g_1,\ldots,g_l) \\ &= g_k b_k(\mathbf{g}) + a_k(\mathbf{g}) + g_k(1-g_k)\frac{b_k^2(\mathbf{g})}{f(\mathbf{g})} \\ &\geq f(g_1,\ldots,g_l). \end{aligned}$$

□

As a consequence of the this property, System (9), with the initial condition $\mathbf{g}(0) = (\frac{1}{2},\ldots,\frac{1}{2})$, can be used as a general optimization algorithm. More precisely, the algorithm that we call GENETIC is sketched below.

```
GENETIC
Input:  A fitness function $f : \Omega \rightarrow \mathbb{N}^+$, mutation rate $\eta$;
        for $k = 1,\ l$ do $G_k := \frac{1}{2}$;
        while ($\neg$ end cond.) do
            [ for $k = 1, l$ do $G_k := (1 - 2\eta) G_k \dfrac{b_k(\mathbf{G}) + a_k(\mathbf{G})}{f(\mathbf{G})} + \eta$;
        for $k = 1, l$ do
            [ if $G_k \geq \frac{1}{2}$ then $X_k := 1$, else $X_k := 0$;
Output: $\mathbf{X} \in \Omega$.
```

GENETIC can be used to approximately solve a combinatorial optimization problem of size l if the polynomial of degree one in each variable that extends the objective function to $[0, 1]^l$ can be efficiently evaluated. This can be done for a large class of problems such as Max Sat, Max k-Cut, Max Bisection, Max k-CSP, Min TSP-(1,2), and Max TSP-(1,2) (Grossi, 1999).

With regard to the worst case performance of GENETIC, the following result gives a lower bound to the fitness value of the solution found when $\eta = 0$.

THEOREM 5.1: *If $f : \Omega \rightarrow \mathbb{N}^+$ is a fitness function,* $\mathbf{U}$ *is the uniform probability distribution on* Ω, *and* $(\tilde{x}_1, \dots, \tilde{x}_l)$ *is the solution found by algorithm* GENETIC, *then:*

$$f(\tilde{x}_1, \dots, \tilde{x}_l) \geq E_{\mathbf{U}}[f].$$

PROOF: By properties of the expectation and by Lemma 5.1 it holds that

$$E_{\mathbf{U}}[f] = f\left(E_{\mathbf{U}}[x_1], \dots, E_{\mathbf{U}}[x_l]\right) = f\left(\tfrac{1}{2}, \cdots, \tfrac{1}{2}\right) \leq f(\tilde{x}_1, \dots, \tilde{x}_l).$$

□

We are not able to theoretically evaluate the effect of the mutation rate. However, experimental results show that the introduction of a small mutation rate increases the quality of solutions.

In conclusion, GENETIC is a wide-spectrum heuristic which always behaves better than random choice; on the contrary, some general heuristics as, for instance, local search, cannot guarantee this solution quality (Ausiello and Protasi, 1995).

6 Application to MAXSAT

We focus here on some variants of the Maximum Satisfiability (Max Sat) problem. For this class of problems, we show an optimality result of GENETIC with respect to the worst case analysis; with regard to the average case analysis on random instances, an experimental comparison with other approximation algorithms and heuristics is presented.

The Max Sat problem is the optimization version of the paradigmatic decision problem Sat, the first shown to be NP-complete (Cook, 1971). It is of considerable interest not only from the theoretical side but also from the practical one. For instance, many issues in mathematical logic and artificial intelligence can be expressed in the form of "satisfiability" or some of its variants, like constraints satisfaction. Some exemplary problems are consistency in expert systems knowledge bases, integrity constraints in databases, and asynchronous circuit synthesis.

Let us formally define the variant Max Ek-Sat of Max Sat that we consider here.

Max Ek-Sat

Instance: A set $X = \{x_1, \ldots, x_l\}$ of Boolean variables, a collection $\mathcal{C} = \{c_1, \ldots, c_m\}$ of clauses with exactly k literals.

Solution: A truth assignment for the variables in X.

Measure: Number of clauses satisfied by the truth assignment.

Given an instance $\mathcal{I}$ of Max Ek-Sat, let $\omega^*(\mathcal{I}) \leq |\mathcal{C}|$ be the maximum number of clauses satisfiable by an assignment, and let $\omega(\mathcal{I})$ be the number of clauses satisfied by a given assignment σ on X. The *performance ratio* of $\omega(\mathcal{I})$ is defined as

$$R(\mathcal{I}, \omega(\mathcal{I})) = \frac{\omega(\mathcal{I})}{\omega^*(\mathcal{I})}.$$

A polynomial-time algorithm A for Max Ek-Sat is said to be an *ε-approximation algorithm* if, given any instance $\mathcal{I}$, the performance ratio of the approximate solution $A(\mathcal{I})$ verifies the following relation:

$$R(\mathcal{I}, A(\mathcal{I})) \geq \varepsilon.$$

In other words, the solution provided by the algorithm $A(\mathcal{I})$ must guarantee at least a value $\varepsilon \cdot \omega^*(\mathcal{I})$.

For Max Ek-Sat, Håstad (1997) has shown that, under the conjecture P $\neq$ NP, $(1-2^{-k})$-approximation algorithms are optimal to within an arbitrary positive additive constant α. Stated otherwise, Max Ek-Sat is not approximable within $(1 - 2^{-k}) + \alpha$ for any $\alpha > 0$ and for any $k \geq 3$. Here, we prove that GENETIC is optimal for solving Max Ek-Sat ($k \geq 3$).

THEOREM 6.1: GENETIC *is an* $(1 - 2^{-k})$*-approximation algorithm for* Max Ek-Sat.

PROOF: Let $\mathcal{I}$ be an instance of Max Ek-Sat, i.e., a collection $\mathcal{C} = \{c_1, \ldots, c_m\}$ of clauses on the variables $\{x_1, \ldots, x_l\}$; let C_i be the Boolean function represented by c_i, and $f_{\mathcal{I}}(\tilde{x}_1, \ldots, \tilde{x}_l)$ be the number of clauses satisfied by the assignment $(\tilde{x}_1, \ldots, \tilde{x}_l)$, then

$$f_{\mathcal{I}}(\tilde{x}_1, \ldots, \tilde{x}_l) = \sum_{k=1}^{m} C_k(\tilde{x}_1, \ldots, \tilde{x}_l).$$

Let GENETIC($\mathcal{I}$) be the number of clauses satisfied by the algorithm GENETIC and $\omega^*(\mathcal{I}) = \max_{\mathbf{x} \in \Omega} f_{\mathcal{I}}(x_1, \ldots, x_l)$, it holds:

1. $\omega^*(\mathcal{I}) \leq m$;

2. by Theorem 5.1, GENETIC($\mathcal{I}$) $\geq E_{\mathrm{U}}\left[f_{\mathcal{I}}\right] = \sum_{k=1}^{m} E_{\mathrm{U}}\left[C_k\right] = m\frac{2^k - 1}{2^k}$.

We conclude that

$$R(\mathcal{I}, \textsc{Genetic}(\mathcal{I})) = \frac{\textsc{Genetic}(\mathcal{I})}{\omega^*(\mathcal{I})} \geq 1 - 2^{-k}.$$

□

Experimental results show that the lower bound $1 - 2^{-k}$, obtained in the worse case analysis, is too pessimistic. The actual performances of GENETIC has been experimentally evaluated and compared with those of some widely used approximation algorithms and heuristics for Max Sat such as classical local search, the Johnson algorithms, simulated annealing, tabu search and its variants, non-oblivious local search, and a heuristic called Walksat.

The *local search* (LOCAL SEARCH) performs a search in the space of the truth assignments by starting from a randomly generated one, and then repeatedly changing ("flipping") the assignment of a variable that leads to the largest increase in the total number of satisfied clauses.

The first two approximation algorithms historically known as JOHNSON1 and JOHNSON2 (Johnson, 1974), are based on *greedy* construction strategies. The JOHNSON1 algorithm chooses at each step the literal that occurs in the maximum number of clauses, then setting the correspondent variable to *true* (respectively *false*) if it is positive (respectively negative). The clauses satisfied are deleted, and the algorithm stops when all clauses are satisfied or all variables have been set. According to the given definition of performance ratio, JOHNSON1 is a $\frac{k}{k+1}$-approximation algorithm for Max Ek-Sat. By adding something like a weight to each clause, the JOHNSON2 algorithm obtains a better performance ratio, $1 - 2^{-k}$, shown to be the best possible with respect to the worst case analysis.

The approach based on the *simulated annealing* technique (SIM. ANNEALING) introduces uphill moves into local search by using a noise model based on statistical mechanic (Kirkpatrick et al., 1983). For Max Ek-Sat, we refer to the the algorithm described in Johnson et al. (1991).

The *steepest ascent mildest descent* (SAMD) heuristic presented in Hansen and Jaumard (1990), similar to *tabu search* (Glover, 1989), is based on the temporary prohibitions of recently executed moves.

A recent approach called *non-oblivious local search* (NONOBLIVIOUS- LS), independently introduced in Alimonti (1994) and in Khanna et al. (1994), is based on the use of "non-oblivious" functions to direct the search, i.e., suitable objective functions whose local optima are within a constant factor from the optimum. This mechanism allows the algorithm to escape from a given local optimum of the usual objective functions. The main drawback is the difficulty of finding, for a given problem, a non-oblivious function that really improves the performance. Also, the non-oblivious local search achieves a performance ratio $1 - 2^{-k}$ for Max Ek-Sat ($k \geq 3$).

Finally, we consider the two best heuristics known (to our knowledge) for Max Ek-Sat: WALKSAT (Selman et al., 1995), a stochastic variant of the popular GSAT (Selman et al., 1992) and the *Hamming-Reactive Tabu Search* (H-RTS) designed by Battiti and Protasi (1997). The GSAT procedure performs a greedy local search, while in the stochastic version, WALKSAT's different "noise" strategy to escape from the attraction basin is considered. The H-RTS heuristic conjugate the local search based on temporary prohibitions with a reactive schema (see *Reactive Tabu Search* in Battiti and Tecchioli (1994)) that determines ("learns")

Table 1: Average number (and statistical error in parentheses in some cases) of unsatisfied clauses on Max E3-Sat instances of various sizes obtained by different approximation algorithms and heuristics.

Variables Clauses	**100** **500**	**100** **700**	**300** **1500**
Local Search	15.6	26.6	43.5
Johnson1	14.9	28.2	45.3
Johnson2	13.5	26.9	44.1
Sim. Annealing	8.2	18.1	30.0
SAMD	5.1	14.7	15.3
Genetic ($\eta = 0$)	6.6	17.3	19.5
Genetic ($\eta = 10^{-4}$)	5.5	15.8	19.4
NonOblivious-LS	15.1 (0.3)	27.7 (0.6)	43.9 (0.9)
Walksat (0.5)	2.76 (1.06)	13.40 (1.19)	8.12 (1.89)
H-RTS, init. $T_f = 0.1$	2.76 (1.0)	13.40 (1.1)	7.34 (1.5)

the appropriate values of the prohibition parameter by monitoring the Hamming distance along the search trajectory.

Our experiments have been done on Max E3-Sat instances. To evaluate the performance of Genetic (with and without mutation), we considered random instances obtained by the generator of random 3-sat instances that was written by B. Selman[2]. The benchmark we used consists of 50 random instances for each pair of *variables–clauses* in the set {100–500, 100–700, 300–1500}. This choice is motivated by the analysis presented in Mitchell et al. (1992) in which the dependence of computationally difficult random instances on the ratio *number of clauses/number of variables* is studied.

The results are given in the Table 1. For each technique listed in the table, we report the average number of unsatisfied clauses with, in some cases, the statistical error in parenthesis. The first five entries are derived from Hansen and Jaumard (1990); the last three entries are taken from the experimental results reported in Battiti and Protasi (1997) in which the authors use the same random generator we used to show that their method, H-RTS, is slightly better than Walksat on random instances.

The experimental results evidence that our genetic algorithm performs better than the algorithms with a fixed performance ratio such as local search, the Johnson algorithms, and non-oblivious local search. It performs worse than ad hoc heuristics such as Walksat and some variants of the *Tabu Search* (reactive and SAMD) designed specifically for the maximum satisfiability problem, because they capture some insights from the structure of the problem itself. A partial explanation of this behavior, for a simple genetic algorithm with crossover, can be found in Rana et al. (1998) and Rana and Whitley (1998).

[2] The code is available at the ftp archive: `ftp://ftp.research.att.com/dist/ai`.

The algorithm GENETIC has been implemented in C language, and the simulations have been done on a Pentium 200 PC with Linux operating system. The computation time is high for the largest instances of the benchmark, and it is non-comparable with that of some of the most efficient heuristics recalled here.

Acknowledgments

This work has been supported by MURST project: *Modelli di calcolo innovativi: metodi sintattici e combinatori*. A preliminary version of part of this paper has been published in Bäck, T., editor, *Proceedings of the Seventh International Conference on Genetic Algorithms (ICGA97)*, pages 121–126, Morgan Kaufmann, San Mateo, California, 1997.

References

Alimonti, P. (1996). New local search approximation techniques for maximum generalized satisfiability problems. *Information Processing Letters*, 57(3):151–158.

Ausiello, G. and Protasi, M. (1995). Local search, reducibility and approximability of NP-optimization problems. *Information Processing Letters*, 54(2):73–79.

Ausiello, G., Crescenzi, P., and Protasi, M. (1995). Approximate solution of NP optimization problems. *Theoretical Computer Science*, 150:1–55.

Baluja, S. and Caruana, R. (1995). Removing the genetics from the standard genetic algorithm. In *Proceedings of the Twelfth International Conference on Machine Learning*, pages 38–46, Morgan Kaufmann, San Francisco, California.

Battiti, R. and Protasi, M. (1997). Solving MAX-SAT with non-oblivious functions and history based heuristics. In Du, D., Gu, J. and Pardalos, P. M., editors, *Satisfiability problem: Theory and Applications*, volume 35 of *DIMACS: Series in Discrete Mathematics and Theoretical Computer Science*, pages 649–667, ACM Press, New York, New York.

Battiti, R. and Tecchioli, G. (1994). The reactive tabu search. *ORSA Journal on Computing*, 6(2):126–140.

Bertoni, A., Campadelli, P., Carpentieri, M., and Grossi, G. (1996). A genetic model and Hopfield networks. In von der Malsburg, C., editor, *International Conference on Artificial Neural Networks (ICANN 96)*, pages 463–468, Springer-Verlag, Berlin, Germany.

Cook, S. (1971). The complexity of theorem proving procedures. In *Proceedings of the Third Symposium of the ACM on the Theory of Computing*, pages 151–158, ACM Press, New York, New York.

Eiben, A. E., Aarts, E. H. L., and van Hee, K. M. (1991). Global convergence of genetic algorithms: a Markov chain analysis. In Schwefel, H.-P. and Männer, R., editors, *Parallel Problem Solving from Nature - Proceedings of First Workshop (PPSN 1)*, volume 496, pages 4–12, Springer-Verlag, Berlin, Germany.

Glover, F. (1989). Tabu search - part i. *ORSA Journal on Computing*, 1(3):190–260.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, Massachusetts.

Goldberg, D. E. and Segrest, P. (1987). Finite markov chain analysis of genetic algorithms. In Grefenstette, J., editor, *Proceedings of the Second International Conference on Genetic Algorithms and their Applications*, pages 1–8, Lawrence Erlbaum Associates, Cambridge, Massachusetts.

Grossi, G. (1999). *Expectation-Guaranteed Local Search and related Neural and Genetic Heuristics*. Ph.D. thesis, Department of Computer Science, University of Milan, Milan, Italy.

Hansen, P. and Jaumard, B. (1990). Algorithms for the maximum satisfiability problem. *SIAM Journal of Computing*, 44:279–303.

Harik, G. R., Lobo, F. G., and Goldberg, D. E. (1997). The compact genetic algorithm. Technical Report 97006, Illinois Genetic Algorithms Laboratory (IlliGAL), University of Illinois, Urbana, Illinois.

Håstad, J. (1997). Some optimal inapproximability results. In Sipser, M., editor, *Proceedings of the Twenty-ninth ACM Symposium on the Theory of Computation (STOC 97)*, pages 1–10, ACM Press, New York, New York.

Holland, J. H. (1989). *Induction, Processes of Inference, Learning and Discovery*. MIT Press, Cambridge, Massachusetts.

Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, Massachusetts.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554–2558.

Johnson, D. S. (1974). Approximation algorithm for combinatorial problems. *Journal of Computer and System Science*, 9(1):256–278.

Johnson, D. S., Aragon, C. R., McGeoch, L. A., and Schevon, C. (1991). Optimization by simulated annealing: An experimental evaluation; part ii, graph coloring and number partitioning. *Operation Research*, 39(3):378–406.

Khanna, S., Motwani, R., Sudan, M., and Vazirani, U. (1994). On syntactic versus computational views of approximability. In Goldwasser, S., editor, *Proceedings of the Thirty-fifth Annual IEEE Symposium on the Foundations of Computer Science (FOCS 94)*, pages 819–830, IEEE Computer Society Press, Piscataway, New Jersey.

Kirkpatrick, S., Gelatt, C., and Vecchi, M. (1983). Optimization by simulated annealing. *Science*, 220:671–680.

Koza, J. R. (1992). *Genetic Programming*. MIT Press, Cambridge, Massachusetts.

Mitchell, D. G., Selman, B., and Levesque, H. J. (1992). Hard and easy distributions for SAT problems. In Rosenbloom, P. and Szolovits, P., editors, *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 459–465, AAAI Press, Menlo Park, California.

Mühlenbein, H. (1998). The equation for response to selection and its use for prediction. *Evolutionary Computation*, 5(3):303–346.

Mühlenbein, H. and Voigt, H.-M. (1995). Gene pool recombination in genetic algorithms. In Osman, I. and Kelly, J., editors, *Proceedings of the Metaheuristics International Conference*, Kluwer, Boston, Massachusetts.

Nix, A. and Vose, M. D. (1992). Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence*, 5:79–88.

Rana, S., Heckendorn, R. B., and Whitley, D. (1998). A tractable Walsh analysis of SAT and its implications for genetic algorithms. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98) and the Tenth Conference on Innovative Applications of Artificial Intelligence (IAAI-98)*, pages 392–397, AAAI Press, Menlo Park, California.

Rana, S. and Whitley, L. D. (1999). Search, binary representations and counting optima. In Davis, L., De Jong, K., Vose, M. and Whitley, L. D., editors, *Evolutionary Algorithms*, pages 177–189, Springer, New York, New York.

Selman, B., Kautz, H. A., and Cohen, B. (1995). Local search strategies for satisfiability testing. In Triks, M. and Johnson, D., editors, *Proceedings of the Second DIMACS Algorithm Implementation Challenge on Clique, Coloring, and Satisfiability*, DIMACS Series on Discrete Mathematics and Theoretical Computer Science. American Mathematical Society.

Selman, B., Levesque, H., and Mitchell, D. (1992). A new method for solving hard satisfiability problems. In Rosenbloom, P. and Szolovits, P., editors, *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI-92)*, pages 440–446, AAAI Press, Menlo Park, California.

Syswerda, G. (1993). Simulated crossover in genetics algorithms. In Whitley, D., editor, *Proceedings of the Second Workshop of Foundations of Genetic Algorithms*, pages 239–255, Morgan Kaufmann, San Mateo, California.

Vose, M. D. (1996). Modeling simple genetic algorithms. *Evolutionary Computation*, 3(4):453–472.