Probabilistic Evolution and the Busy Beaver Problem

Roberto Santana
Alberto Ochoa-Rodriguez
Marta Soto
Center of Mathematics
and Theoretical Physics
ICIMAF. CP 10400. Habana. Cuba
{rsantana,ochoa,mrosa}@cidet.icmf.inf.cu

Francisco B. Pereira
Penousal Machado
CISUC
{xico,machado}@dei.uc.pt

Ernesto Costa
Amilcar Cardoso
Centre for Informatics and Systems
University of Coimbra. CISUC.
Polo II - Pinhal de Marrocos
3030 Coimbra, Portugal
{ernesto,amilcar}@dei.uc.pt

We present an evolutionary algorithm that uses probability distributions for the solution of the Busy Beaver (BB) problem [1]. The BB involves the representation of Turing Machines (TMs). Suppose a TM with a two way infinite tape and a tape alphabet = {blank, **1**}, the goal is to find the N-state halting TM that writes the maximum number of **1**s when starts on a blank tape. This maximum number, which is function of the number of states, is denoted by $\Sigma(N)$. A machine that produces $\Sigma(N)$ non-blank cells is called a Busy Beaver (BB).

In our representation for the BB(N) each potential solution is an integer vector with 4*N variables . Variables that represent transitions to the new state can have N+1 different values (N states plus the halting state). The values for the rest of variables range between 0 and 3, corresponding to the 4 actions that can be performed. Our probabilistic model is one table of 16 entries that keeps the frequencies of all possible states' actions regardless their position in the chromosome. In this partial model approach we only use the probabilistic information of variables representing actions.

We introduce the Probabilistic Oriented Search Algorithm for Turing Machines (POSATM) (see table 1). Experiments were oriented to compare the behavior of POSATM with results presented in [2] for a GA with two point and graph based crossover. Experiments concerned how efficient is the search for the 4-tuple BB(6).

Our bivariate model has been shown to be superior to both, one point crossover and graph based crossover GAs for the experiments conducted. The POSATM was able to find 42 times the maximum value, whereas the best GA only 6 times for the same number of experiments.

The experiments have shown that, at least under certain conditions, an evolutionary approach based on probability distribution seems to outperform recombinative methods in the optimization of problems with graph-like structures. Besides, it is possible to construct useful partial probabilistic models based only on the dependencies displayed for a subset of the variables. This is an important issue because many problems could beneficiate from a preliminary variable selection step. Finally, we believe that our results have a more general scope, they could be applied also to other Turing or Finite State Machine problems.

Table 1: POSATM.

step 0:	$t \leftarrow 1$. Generate T points randomly or by a seeding procedure.
step 1:	Get a selected set S with M points. ($M < T$). Estimate the model $p_M^s(\widehat{x}, t)$.
step 2:	Generate T points making a partial modification of the selected points using p_M^s to update V_a variables.
step 3:	mutate variables
step 4:	$t \leftarrow t + 1$. If the termination criteria are not met, goto to **step 1**.

References

[1] Rado, T. (1962) On non-computable functions, The Bell System Technical Journal, vol. 41, no. 3, pp.877- 884.

[2] Pereira, F. B., Machado, P., Costa E. and Cardoso, A. (1999). Graph Based Crossover A Case Study with the Busy Beaver Problem. Proceedings of the Genetic and Evolutionary Computation Conference GECCO 99. Volume II (pp. 1149-1155) Orlando Fl. Morgan Kauffman Publishers. San Francisco, California.