ONLINE GENETIC ALGORITHMS

Alfredo Milani

Abstract: This paper present a technique based on genetic algorithms for generating online adaptive services.

Online adaptive systems provide flexible services to a mass of clients/users for maximising some system goals, they dynamically adapt the form and the content of the issued services while the population of clients evolve over time.

The idea of online genetic algorithms (online GAs) is to use the online clients response behaviour as a fitness function in order to produce the next generation of services. The principle implemented in online GAs, "the application environment is the fitness", allow modelling highly evolutionary domains where both services providers and clients change and evolve over time.

The flexibility and the adaptive behaviour of this approach seems to be very relevant and promising for applications characterised by highly dynamical features such as in the web domain (online newspapers, e-markets, websites and advertising engines). Nevertheless the proposed technique has a more general aim for application environments characterised by a massive number of anonymous clients/users which require personalised services, such as in the case of many new IT applications.

Keywords: genetic algorithms, adaptive web, evolutionary computation

1. Introduction

The research on the topic of adaptive systems has mainly focused on architectures based on knowledge representation and reasoning [1], fuzzy reasoning [2][3] and probabilistic models [4]. These approaches are often able to give an adequate account of uncertainty and dynamical aspects of the domain, but they also require a great effort in building a detailed model of the problem. Despite of the good qualitative response, they often reflect too rigidly the domain constraints at modelling time. When the environment, i.e. the constraint of the domain, evolves, the system performance tend to decrease until the model needs to be modified or redesigned.

The increasing diffusion of mass services based on new information technologies (ITs) poses new requirements and goals on adaptive systems which are seemingly contradictory, such as the problem of providing adaptive personalised services to a mass of anonymous users [4]. Sometimes models of user behaviour [1] for the new services does not even exist, and, in addition, services and technologies appear and disappear very quickly thus vanishing the effort of building accurate models.

The growing interest in self adaptive and self modelling systems is partially motivated by these reasons.

The two leading approaches to self adaptation, i.e. genetic algorithms [5][6] and neural networks [7][8] are characterised by somewhat symmetrical features which are worth to be pointed out: *neural networks (NNs) tends to be online systems while GAs operate offline*. GA usually operates offlinein the sense that they can be seen as building a simulated application environment in which they evolve and select the best solution among all the generations, under the well known Darwinian principle of "survival of the fittest".

Some works [9][10][11] have introduced "real world" issues into the GA loop, in the interactive GAs approach [12] the user is inserted in the algorithm with the role of providing the fitness functions by interacting with the GA, in other works still following the offline approach [13][14] about machine learning by GA, historical real data are used as fitness function.

Despite of their offline nature GA are able of a highly dynamical behaviour. The main reason is that the knowledge about "reasoning" structure of GA is embedded in the population chromosomes: when the population evolves the structure evolves as well. GA concepts such as *cross over* and *mutation* have no counterpart in NNs

approach, but they are a powerful tools which can allow a GA to make fast hill climbing of local minimum and plateau in optimisation problems [6].

The idea of bringing these adaptive features in the *online system* scenario is made more challenging from the facts that the population of clients asking for services is evolving over time, then their response to services changes.

In this paper we propose a new approach, *online genetic algorithms* (online GAs) which tries to combine timely responses with the adaptive behaviour of GAs. The basic idea of online GAs is to evolve populations by using the application world as a fitness function, under the principle "the real world is the fitness".

The goal of systems based on online GAs is to give a timely response to a massive set of clients requesting services, and to be able to adapt services to clients, both changing over time in unknown and unpredictable way.

As noted in the beginning, it is not realistic to rely on the hypothesis of detailed user models [1][15]. The increasing consciousness of privacy issues, legal limitations on personal info [16] and the growth of mobile and pervasive interfaces accessible from casual users, often make the user model impossible to collect. The anonymity of users/clients is then a structural constraint in mass adaptive services.

In the next paragraphs we will motivate the online GAs approach by analysing the features of a sample dynamical scenario regarding an online newspaper management system.

The principles and the architectural scheme of the online genetic algorithms approach will be presented, an example application and experimental results will be discussed.

2. The Online Adaptive Scenario: Web Newspaper

Let us consider as a typical scenario for online genetic algorithms: the problem of managing the generation of an online newspaper with the goals of maximising customers, i.e. readers, contacts.

The problem, well known to journal editors, is to build a newspaper in order to publish news according to the newspaper politics and mission, and selling it at its best. Selling news in this contexts means the goal of capturing readers attention for reading the articles, and for, possibly, satisfying the newspaper advertisers. Online readers browse time by time the newspaper web site and read the news which interests them. It is assumed that a good journal will collect a great number of contacts and many users will spend time in reading it. Managing editors of online newspapers have a great advantage with respect to their hardpaper colleagues: while a conventional paper journal is limited and bounded to a single daily edition (except the cases of extraordinary events), an online editors, instead, can make timely adaptation of the newspaper to the latest news, thus maximising the impact of the newspaper on the readers.

Online media have the likely feature that can be produced and delivered instantaneously such that, in principle, each user can read his own single, personalised and different copy of the journal.

The main issues, and source of difficulties, in the newspaper scenario are the lack of information about the users and the unpredictable dynamical evolution of all the elements which characterised it, in particular:

- anonymity of clients
- dynamical evolution of potential services
- dynamical evolution of clients
- dynamical evolution of client goals

these evolutionary features are shared by a wide class of online problems.

2.1. Anonymity of clients

Anonymity of clients means that no hypotheses can be made about profiles of the users of online services. As discussed in the introduction the typical assumption for online newspaper is that the information available to the system comes from anonymous user sessions, where users cannot be identified, nor recognised from previous sessions [16][17].

2.2. Dynamical evolution of potential services

The purpose of online systems is to provide the best of their currently available services for maximising the client impact [18], the situation is made more complex since *the services that are issued by the providers can vary over time in unpredictable way*.

News, seen as services, are characterised by a lifetime cycle (i.e. they appear, disappear and are archived), and the news flow is by its nature unpredictable. Thus the news editor task is to select according to the editorial line, which news best interest and impress their readers, among the available ones.

2.3. Dynamical evolution of clients

The set clients connected with the online system evolves over time in unpredictable way. The set of connected clients are not always the same, since new clients come and previous sessions disconnect.

In the case of online newspaper there can be made some general assumption about the target users. Users are assumed to have somewhat homogeneous features like in the case of readers of newspaper specialised in economics, politics, sports etc. Nevertheless the instantaneous audience profile of online newspaper can vary over time. For instance students can connect mainly in the afternoon, while corporate workers can connect in different time range. In addition, external factors and unpredictable events, such as holidays or exceptional events, can make different classes of readers to connect in unexpected time/dates.

Even assuming that we have a way of determining the ideal journal for the current audience given the currently available news, the newspaper edition will be no more adequate after some time, since the audience will change unpredictably.

2.4. Dynamical evolution of client goals and attitudes

Goals and attitudes of the single clients can vary and depend on time.

As we as pointed out before, external events of general interest can make the journal audience vary, but can also make the interests of the audience to vary. Economical or political events can induce a shift in the typical interest of the readers. Moreover even assuming to have a fixed audience, with fixed goals, is not possible to produce a fixed "ideal newspapers", since people expect that newspaper vary: it would be unlikely to read every day the same identical news; typical users of online newspapers connect to the system many times a day, expecting to read more news on topics of their interest.

2.5. Model of Service Impact Factor

A model of the impact factor of service cannot be easily defined and require classification effort.

The goal of the newspaper editor is to catch the attention of most of its readers by selecting the appropriate news and preparing a suitable edition according to the newspaper editorial line, i.e. mission, policy and cultural goals.

The typical tools available to an editor to maximise the impact of the service he provides (i.e. the news) are: *selections* of the news among the continuous flow (deciding which news are currently published and which news go to archive); *location* of the news in the grid of the newspaper layout (the position of the news usually reflect is evidence or priority in editor's intention); *presentation form* of the news, which regards aspects such are selecting a *title* for the news, and or selecting a possible *picture* accompanying it, and sometime also long or short versions of the article. These tasks are usually regarded to as an "art" which the newspaper editor performs by the help of his/her experience.

It is worth noticing that some factors, such as the news position in the layout, are not necessarily determining the readers' priority. A well prepared journal, for example, usually offers a mix of different news (i.e. not many news on the same topic). The visibility strictly depend not only in the position but also in the context in which news are presented. Sometimes hot emerging topics require breaking these rules and, when it happens most part of the journal news are devoted to a single topic.

The next paragraph will describe a framework based on genetic algorithms for providing adaptive services in highly evolutionary environment to a massive audience of anonymous users, such as in the newspapers scenario.

3. Online GA Schema

GA have been classically proposed for use in an *offline schema*. In the offline approach populations of solutions are evolved offline for a given number of generations in order to produce the best evolved solution (usually determined in the last generation) which given as system output. For instance in classical optimisation problems [6] GA are used for exploring a search space of solutions and the best minimum/maximum value found over all generation is produced. In GA applied to learning problems, such as discovering stock market rules [14], real data about stock market are used to evolve the population, but again, the best solution is computed offline, and it is used in the *real market* afterwards. A different approach is that of Interactive Genetic Algorithms [9] [12] [19] where the real world is included into the GAs loop under the principle that "the user is the fitness", i.e. the user participates to a cooperative optimisation process. In some interactive GAs applications to robot learning [13], the real world is used to evolve the solution, but GAs uses real world in an offline phase of training.

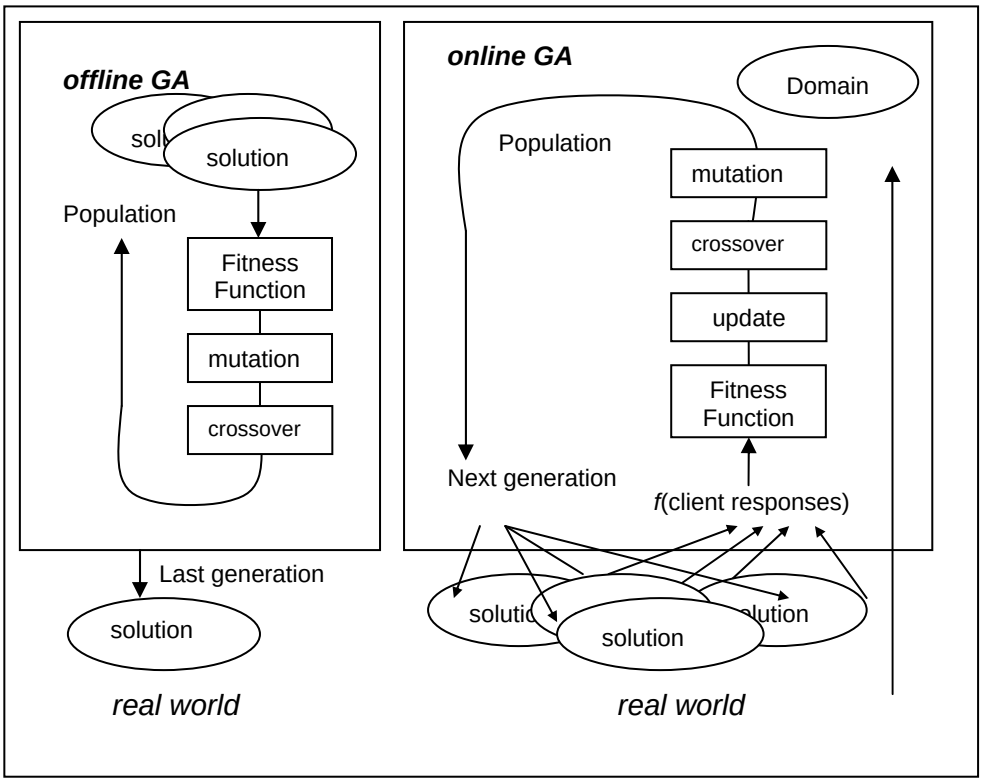

Fig.1 Online GA

In the online GAs approach we propose to literally implement the evolutionary metaphor which originally motivates GAs. In our proposal the basic elements and concepts of GAs such as population, chromosomes, crossover, mutation and selection mechanism still exists but, they are extended with the innovative but simple assumption that "*the real world is the fitness*", i.e. the application world, representing the environment, is used to selects the best surviving fittest solutions, moreover all generated solutions are output to the system and no interactive cooperation is required to users/clients.

The basic scheme of an online GAs versus offline GAs is depicted in the fig.1.

In *online GAs* a population of solutions is evolved with usual genetic mechanisms, with the difference that the solutions in the population are actually "executed", and the client/users behaviours/responses upon solutions are used as a fitness to evolve the next generation of online solutions. In other words the fitness function resides in the real world and it is expected to give timely response to the evolution of clients' population and domain modification.

Updating Phase

Since in online GAs the problem domain also evolves unpredictably (for instance the provided services changes, i.e. the flow of incoming/outcoming news), then there is the need for a novel phase of *updating*, not present in classical GA. The purpose of the *updating* phase is to establish and implementing a policy about how to adapt the current population of solutions to domain changes (such as how to replace an article which is disappeared from the journal, because expired or delete after explicit editor's decision).

Online GAs can be used when it is required to dynamically adapt to evolution and changes in the problem domain, moreover application domains best suitable for the adoption of online GAs are characterised by:

- a solution space with "many" valid solutions to explore, i.e. the solution space with not unique or few valid solutions
- a set of clients which require solutions to be used immediately
- an optimisation function which measures the efficacy of a solution given in output, which can be "sensed" by the system in the external world as a response/result produced by clients

It is worth noticing that online genetic algorithms would not be a real possibility without the new ITs. A massive diffusion of the internet, mobile phones services, on demand phone services has made possible application servers where a huge number of anonymous clients (with no distinction among final users or software agents) are concurrently requiring services in an automated framework which directly connects consumers to service providers. The services providers are usually optimising very simple functions which are completely inside their “sensing” scope such as *time-spent*, *services bought*, *money charged*, *advertising clicked*.

4. An Online GA

The pervasive dynamical and unpredictable evolution of all the key elements in the newspaper scenario represents a difficult challenge for adaptive systems, which should provide adequate services in answers to clients’ requests.

News to be offered in the newspaper is continuously flowing in from news agency and journalists. Different classes of anonymous individual readers continuously connect and disconnect in order to read interesting news The goals and interests of the individuals vary in an unpredictable way (people get bored of old news). The impact of news upon users depends of the form, the position and the context in which the news is given, and it is hard to be deterministically modelled [9][20]. Finally the editor policy represents a pervasive constraint to be respected throughout the journal editing.

In the following we present the architecture of a sample online GAs applied to the newspaper evolutionary scenario.

4.1. Domain and Constraints

A newspaper has a typical layout and structure in term of sections of topics, which are part of the recognizable corporate image. No editors are available to modify it, moreover the editor usually want to have control over the proposed news in order to implement the editorial policy.

In order to reflect these constraints the structure and layout of the journal do not evolve, and the editor decides which news include/exclude in/from the newspaper and how to assign (or remove) them to sections, let the sections, for instance: *TopStories, National, International, Sports, Health and Technology*. A limited set of headlines (for example 4 headlines) is reported in the front page for each section; the sections occupy fixed layout positions; the section headlines are chosen among the articles available inside the sections.

For each single article we will assume that the newspaper editor provide a set of possible alternative formats for each article, i.e. alternative titles, texts and pictures to be used for presentation.

The task of the editor is to decide how to update the set of news and formats, while the online GA actually build the newspapers deciding which articles will be inserted into the sections headlines and which alternative formats will be used in the articles presentation.

4.2. Population and Individuals

The individuals which compose the population of the current generation consist of the different versions of the newspaper which have been issued to the currently requesting online readers.

4.3. Time Intervals

In order to make the online GA having a sufficient number of individuals in the population, and a sufficient time to evaluate user response, i.e. fitness of the individual, it is needed to fix a time interval value, i.e. the duration of the minimal interval of time from one generation to the next one. If, for instance, a newspaper has 6000 contacts per hour, a time granularity of 1 minute guarantee, guarantee an average population of 100 individuals, but doe not allows evaluating responses whose duration is greater than one minute.

4.4. Fitness Function

The fitness function measures the adequacy of the solution in term of client response.

According to the anonymity hypothesis the system is able to "sense" user sessions, but not to recognise user from previously started session. Sensing data are easily collected from the web server log files [Etzioni2000]. In the newspaper problem the fitness of a given solution k (i.e. the individual version of the newspaper) is defined as

$$F(k) = w_s t_s + w_{ca}\, n_{ca} + w_{cn}\, n_{cn} + w_{int}\, (\Sigma_{(i=1..n_{cn})}\, t_{si}/t_r) / n_{cn} - w_{nohl}\, c_{nohl}$$

The listed parameters reflect the general criteria that reward as positive, in particular: t_s is the *total time spen*t on the newspaper (measured as the time between the first and the last browser request); n_{ca} the *number of clicks on* newspaper *advertising*s; n_{cn} *number of clicks on news* (i.e. how many news have been red; the Σ term computes the *average interest of news*, where the interest is measured as the time spent t_{si} on a single news with respect to the time t_r needed to read the news (skipping rapidly a news means little interest versus carefully reading it); the minus terms c_{nohl} in $F(v)$ penalises the situations in which the readers find *no interest in headlines* and go straight to sections to read particular news, i.e. in other words it penalise at a certain extent the journal versions in which the content is interesting while presentation is not. Weights w_s, w_{ca}, w_{cn}, w_{int} and w_{nohl} are used to tune the contributions of the respective terms to the global fitness.

4.5. Chromosomes

The individuals, i.e., the single newspaper versions, are encoded by a set of *sections vectors* each one encoding a section of the newspaper.

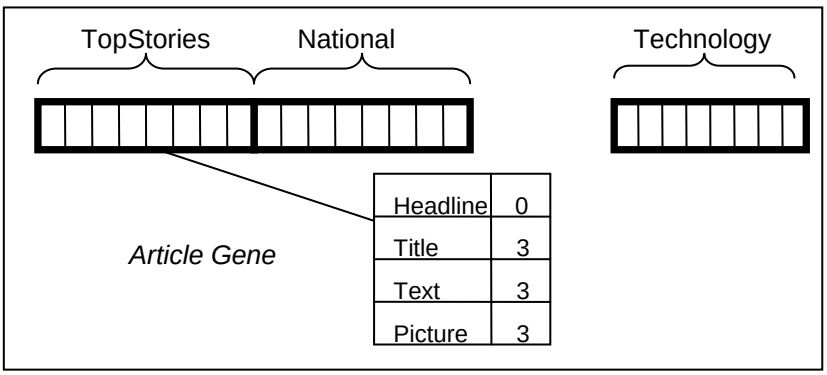

Fig.2 Individual chromosomes

Each element in the journal chromosome specifies a single news in term of its position in the section headlines (0 means not in headlines), and its presentation i.e. values indicating which title, text, and picture, the newspaper edition will contain for the given article among the different available versions.

4.6. Selection

A standard *proportional to fitness* selection method is used in order to determine the intermediate population used for crossover and mutation. The more the fitness is high more chances are given to individuals to survive. On the intermediate population thus determined crossover and mutation are applied.

4.7. Crossover

The purpose of crossover is to generate a new journal version from two individual chromosomes. The two offspring replace the parents. Again a proportional to fitness reproduction criteria is used.

The crossover is operated *section by section* on the whole chromosome. For each section a linear crossover point is determined (see dashed line in the figure 3) for splitting the section subvector. The respective subsection of the two parents is then combined.

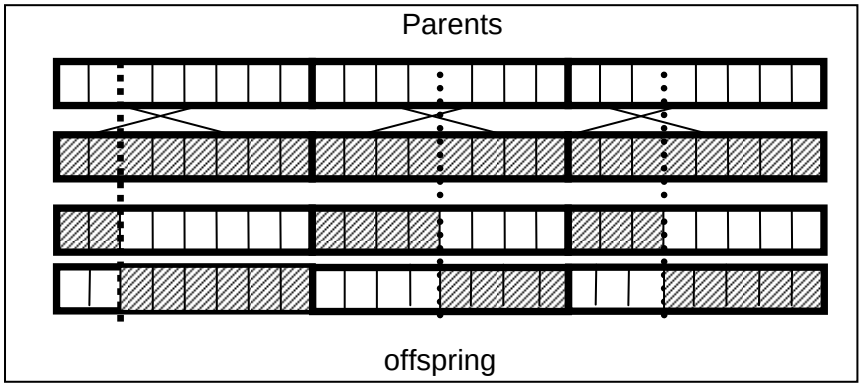

Fig.3 Segmented crossover

Restoring valid solutions can be necessary after crossover reproduction. Suppose that a given section is allowed h headlines; the split point position can divide the section segments such that one offspring segment contains more than h headlines while the other has less than h, i.e. the solution is not valid. In this case in order to restore a valid solution we move headlines from the longest to the shortest one selecting them randomly. Another case of invalid solution is when two headlines in a section points to the same position (another headline position must be empty), in this case the tie is broken randomly. Note this criteria guarantee that all headlines in the parents will be again headlines in the offspring.

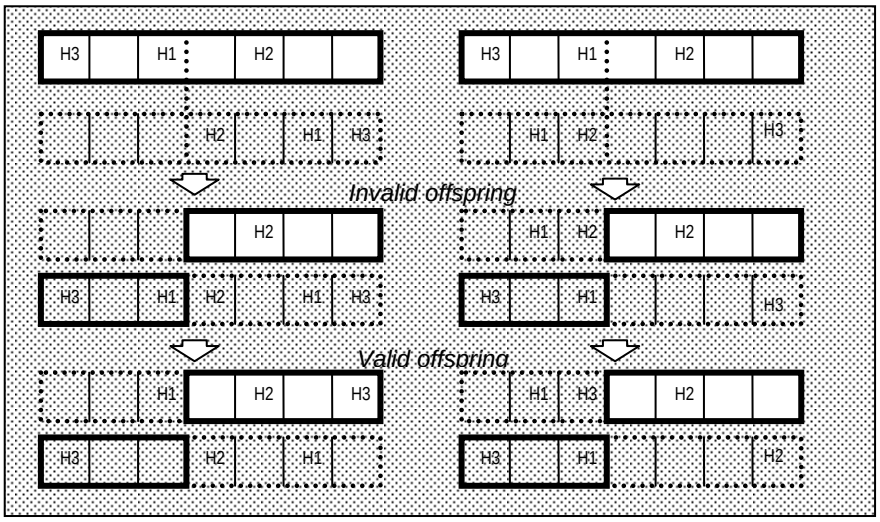

Fig. 4 Restoring Valid Offspring after crossover

4.8. Mutation

Mutations are operated at different levels with different priorities.

- *headline mutation*, is the operation which moves a news from sections into headlines and vice versa, since

an headline mutation is a dramatic change in a newspaper version, the probability P_h of headline mutation is kept relatively low, on the other hand and additional factor P_{new} is considered, P_{new}, is giving more probability to become section headline to new articles versus old ones;

- *format mutation*, this mutation tends to adapt the form in which the single news are given, i.e. order, titles, alternative texts and accompanying pictures, the probability P_f of this mutation is slightly high than the previous one.

A *format mutation* is realised by choosing randomly a format component (*order*, *title*, *text* and *picture*) and a feasible random value in the domain of the format component (e.g. one title over three available candidates). Format mutation of *ordering* is only applied to headline news and consists in swapping an article with another one randomly selected among the headlines.

Headline mutation is realised by randomly selecting the incoming article (taking into account of P_{new} to give priority to new articles), selecting the outcoming article and swapping them among headline and section.

4.9. Update adaptation phase

The adaptation phase concerns the problem of adapting the population of solution which were made invalid by external modifications. For example when the news editor decides that an old article has to be archived and/or a new one has to be inserted into the journal, some individual in the current population could be no more valid. In restoring the validity of the solution we use the following criteria:

- incoming news are added to the respective section with maximum P_{new}
- outcoming news not appearing in section hot headlines are simply deleted from the section
- outcoming news which are on the section headlines are replaced by shifting up the section headlines, and operating a format mutation on the last position where simple insertion replace swapping

The array representation is updated accordingly.

5. Experiments

Experiments to test online GA for newspaper management described in the previous paragraph have been designed and are under implementation. The first stage will take place using the "What's new?" list box in our University home page. The "What's new?" is an approximation of a single section of a newspaper. The "What's new?" area is located right in the center of the home page and it contains a set of headlines which link to announcements of University activities and events. The WebMaster policy limits the number of headlines to a maximum of 8, but much more departments, administrative offices and other institutions are competing for having their announcements on the home page box. There are averages of 30 candidate announcements per day, some 22 of which are forced to reside in the internal "What's new?" section. The system will be implemented in php in connection with an Apache server on a Linux platform. The idea is to provide to about an half of the thousands contacts a day, the GA managed version of "What's new?" and give the rest the usual static version for comparisons. Two further versions are under implementation: a newspaper manager, based on the content management system Nuke [21], and a simulation environment which aims at compare the online GA with respect to simulated user response.

6. Conclusions

Online GAs represent a new approach to systems which provide adaptive services to a large number of anonymous clients/users which evolves over time in unpredictable way.

The basic idea of online genetic algorithms is that “the world is the fitness”, i.e. the fitness function resides in the application environment and it can be evaluated by sensing the environment i.e. by evaluating clients/users response to the current solutions. A phase of adaptation is added to usual GA schema for restoring validity to solutions made invalid by evolution in the problem domain.

Online GAs are related with interactive GAs methods [12][9][10], in which the real world appear in GA in the form of the user cooperation to the selection process, or in the form of environment guided training [13]. The main difference between online GAs and interactive GAs is that in interactive GAs, GAs are used in a sort of offline simulation in order to select a *final* optimised solution or behaviour, used by the application. Instead online GAs based applications made immediate use of the solutions population.

The main issues which motivate the adoption of online GAs have been discussed in the framework of the newspaper scenario. Online GAs represents an answer in all those situation in which adaptation is required, while few o no data are available about users' profiles and attitudes [17].

The increasing diffusion of massive distributed services based on the new ITs, the increasing consciousness and laws about the privacy issues, motivates the apparently contradictory request of providing adaptive services to unknown users in dynamical domains.

Preliminary experimental results on a simplified version of the newspaper application confirm the validity of the online GAs approach.

Open theoretical and practical issues need to be further investigated in the framework of online GAs such as the problem of time granularity with respect to the time needed for fitness evaluation; defining effective methods for tuning GA parameters and weights, discussing typical GA issues such as co-evolution [17] in the context of online Gas, and providing extensive experimental evaluation of the performance.

Moreover the integration between online GA and other non evolutionary techniques such as fuzzy and probabilistic analysis are worth to be investigated.

References

[1] A. Kobsa and W. Wahlster, editors. User Models in Dialog Systems. Springer Verlag, London, 1989.

[2] L. A. Zadeh: Fuzzy Sets Information and Control 8(3): 338-353 (1965)

[3] M.A.S.; Monfared, S.J. Steiner Fuzzy adaptive scheduling and control systems in Fuzzy Sets and Systems Vol.115, n. 2 pp. 231-246, 2000.

[4] J.Binder, D.Koller, S.Russell, K.Kanazawa, Adaptive probabilistic networks with hidden variables. Machine Learning, 1997.

[5] J. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, 1975.

[6] Whitley, D. An overview of evolutionary algorithms. Information and Software Technology, (2001).

[7] J. A. Anderson.An Introduction to Neural Networks,MIT Press Boston 1995

[8] M. T Hagen, H. B. Demuth, M. Beale, Neural Network Design, PWS Publishing Co. Boston 1996.

[9] T. Masui. Graphic object layout with interactive genetic algorithms. Proc. IEEE Visual Languages '92, 1992.

[10] J. G. Peñalver and J. J. Merelo. Optimizing web page layout using an annealed genetic algorithm as client-side script. In Proceedings PPSN, Parallel Problem Solving from Nature V, Lecture Notes in Computer Science. Springer-Verlag, 1998.

[12] H. Takagi, Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation in Proceedings of the IEEE01, IEEE Press, 2001.

[13] M. Dorigo and U. Schnepf, Genetics-based Machine Learning and Behaviour Based Robotics: A New Synthesis, IEEE Transactions on Systems, Man and Cybernetics, vol.23,n.1,pp. 141-154,1993.

[14] L. A. Becker, M. Seshadri, GP-evolved Technical Trading Rules Can Outperform Buy and Hold, in 3rd International Workshop on Computational Intelligence in Economics and Finance. Sept 2003.

[15] J. Kay, B. Kummerfeld, P. Lauder, Managing private user models and shared personas in Proceedings of Workshop on User Modelling for Ubiquitous Computing, User Modelling 2003.

[16] M. K. Reiter and A. D. Rubin, Crowds: anonymity for Web transactions, ACM Transactions on Information and System Security,vol.1, n.1,pp.66-92, 1998.

[17] N. Kushmerick, J. McKee, F. Toolan, Towards zero-input personalization: Referrer-based page prediction, Lecture Notes in Computer Science, vol.1892, Springer-Verlag, 2000.

[18] M. Koutri, S. Daskalaki, N. Avouris, Adaptive Interaction with Web Sites: an Overview of Methods and Techniques, in Proc. of the 4th Int. Workshop on Computer Science and Information technologies CSIT02, Patras Greece,(2002).

[19] A. Oliver, N. Monmarché, G. Venturini, Interactive design of web sites with a genetic algorithm. Proceedings of the IADIS International Conference WWW/Internet, pages 355-362, Lisbon, 2002.

[20] J.González; J.J.Merelo; P.A.Castillo; V.Rivas; G.Romero; A.Prieto. Optimized web newspaper layout using simulated annealing. In Sánchez-Andrés, Mira, editor. IWANN99, LNCS. Springer-Verlag, June 1999.

Author Information

Alfredo Milani – Department of Mathematics and Computer Science, University of Perugia