A Methodology for Designing Emergent Literary Backstories on Non-player Characters using Genetic Algorithms

Rubén Héctor García-Ortega
Fundación I+D del Software Libre
rhgarcia@fidesol.org

Pablo García-Sánchez,
Antonio Miguel Mora and
Juan Julián Merelo*
GeNeura, ETSIIT + CITIC, University of Granada
pgarcia,amorag,jmerelo@geneura.ugr.es

ABSTRACT

The creation of fictional stories is a very complex task that usually implies a creative process where the author has to combine characters, conflicts and backstories to create an engaging narrative. This work presents a general methodology that uses individual based models to generate cohesive and coherent backstories where desired *archetypes* (universally accepted literary symbols) emerge and their life stories are a by-product of the simulation. This methodology includes the modeling and parameterization of the agents, the environment where they will live and the desired literary setting. The use of a genetic algorithm (GA) is proposed to establish the parameter configuration that will lead to backstories that best fit the setting. Information extracted from a simulation can then be used to create the backstories. To demonstrate the adequacy of the methodology, we perform an implementation using a specific multi-agent system and evaluate the results.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Genetic Algorithms; literature; content generation

1. INTRODUCTION

Non-Player Characters (NPCs) in video games are a type of characters that live in the game world to provide a more inmersive experience and, in some cases, present a challenge to the human player. Games can include hundreds or even more NPCs, and their collective interactions improve the gaming experience and lead to a richer and more inmersive world. But inmersion is not guaranteed and can only be achieved if the characters backstory is coherent and also diverse. which motivated us to develop a methodology to generate backstories by modeling the language, agents and a literary setting based in the concept of *archetypes*, behaviors and patterns universally accepted and present in the collective imaginary. This methodology is able to create massive backstories for secondary characters, in order to provide a context for the writer and the player to perceive a virtual world as coherent, detailed and enriched.

To validate our methodology we have developed a multi-agent system called MADE (Massive Artificial Drama Engine) that models a self-organized virtual world where every element influences each other, following cause-effect behaviors in a coherent manner. We have followed the ideas of the work [2], the first widely known multi-agent generative social model. As a step of the methodology, a self-organizing system is defined following the methodology introduced in [3]: a virtual world, agents who are born, grow, interact, reproduce and dead; resources (food), mediators, and relations of rivalry (friction) and cooperation (synergy). In other step of our methodology, the actions of the agents are parametrized according the work [4], based on the use of GAs in order to obtain a plot (solution) where two characters interact in a creative way. The narrative is addressed by our methodology as the final step, giving freedom to creators. This issue has been studied in the systems presented on the review [1].

In this paper we show that GAs, together with a proper design of literary patterns and archetypes, can be used to find the parameters that promote the generation of drama plots and sub-plots in a multi agent based environment.

2. METHODOLOGY AND APPLICATION

In our approach, the mood of the setting will be modelled as a group of abstract archetypes and different conditions over them. Initially, each agent has to be modelled as a Finite State Machine (FSM) whose transitions generate symbols in a language that includes dates, actions, direct and indirect objects.

A Genetic Algorithm is used, mapping the agents parameters (profiles) into the chromosome. A *profile* is a set of properties assigned to an agent. If two agents have different profiles, their behavior in the face of the same inputs and internal states can be different. In the chromosome, the use of n profiles means a chromosome size equal to the number

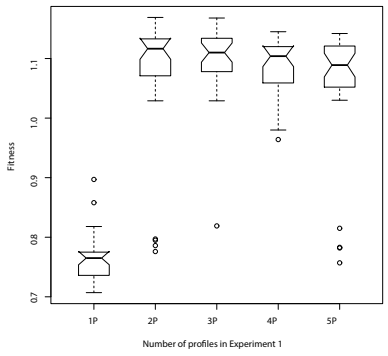
*This work has been developed thanks to the funds provided by the ANYSELF (TIN2011-28627-C04-02) and SIPESCA (G-GI3000/IDIF) projects.

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.
ACM 978-1-4503-2881-4/14/07.
http://dx.doi.org/10.1145/2598394.2598482.

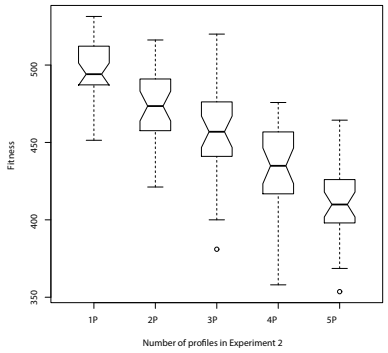
of searchable properties multiplied by n. The fitness function of a chromosome is the result of calculating the average of the application of the setting function over n executions of the environment with defined base properties.

To validate our methodology we apply its steps using a specific scenario and environment: the rats that inhabit in the Invisible University of Ankh-Morpork.

We have developed the MADE environment, a virtual place where different agents play their artificial lives[1]. In this paper we model a virtual rat with four states (be alive, be hungry, look for a mate and be pregnant) and seven actions (move, eat, attack, defend, escape, find mate and have offspring) that lead to a basic instinctive animal behavior. Each agent is created using twelve parameters, that define its base features and probabilities.



(a) "Territorial war" Setting



(b) "Revenge" Setting

Figure 1: Average fitness of the 30 best individuals of the GA for each configuration.

Two different literary settings are going to be tested. The first literary setting, "territorial war", aggregates different sample archetypes where many factors must be taken into account. It tries to find the number of profiles and values that generate at the end of the run an equal distribution of the archetypes *downtrodden* (an agent that has been attacked at least two times and has defended the position), *warrior* (if it has satisfactory attacked at least five times), *helpless* (if it has been defeated 10 times) and *bad warrior* (if it has unsatisfactory attacked at least 10 times). The second literary setting is called "revenge" and its goal is to model an individual complex memory based behavior between two characters: it tries to find the number of profiles and values which are optimal to make *revenge* archetype emerge in as many agents as possible. An *avenger* is an agent (a) that has been attacked by other agent (b) and, after that, it has satisfactory attacked the agent b, in *revenge*. For this setting, every agent whose log matches the archetype adds one point to the fitness.

The experiments performed in this work uses 12 alleles per profile, a fitness function as the average of ten executions, a natural selector with rate 0.9, a crossover operator with a rate of 35%, a mutation probability of 1/12, 100 generations and population size of 30 individuals. These values have been chosen empirically after several test runs. To ensure the representativeness of the results, we have executed 30 times each experiment with different number of profiles, from 1 (P1) to 5 (P5). Boxplots of the best fitness obtained are shown in Figures 1(a) and 1(b). In the setting "Territorial war", we have performed a Kruskal-Wallis test for the best individuals fitness, obtaining differences among all the number of profiles (p-value $<<$ 0.05). It is clear that using one profile is not enough to emerge the desired archetype. However, the pairwise comparison using Wilcoxon does not find significant differences between the results that use more than two profiles. In the setting "revenge", there are significant differences among all configurations (p-value $<<$ 0.05) except between P2 and P3 (p-value=0.3). Therefore, we can conclude that in this kind of global archetype only a profile must be used for obtaining the best results.

3. CONCLUSIONS

This work presents a general methodology to design emergent literary stories in massive virtual worlds. The described steps include the modeling of the agents and literary setting. Then a Genetic Algorithm is used to optimize the parameters of the agent's profiles (sets of parameters assigned to the agents) using as fitness a function that models the literary setting by using the concept of *archetype*. The execution of the MADE Environment using these profiles as input would produce a background (or set of characters' lives) where the archetypes have emerged and have automatically created massive backstories coherent with the settings of the artwork. Results show that the optimal number of profiles depends on the number of desired archetypes and their nature.

4. REFERENCES

[1] M. Arinbjarnar, H. Barber, and D. Kudenko. A critical review of interactive drama systems. In *AISB 2009 Symposium. AI & Games, Edinburgh*. Citeseer, 2009.

[2] J. M. Epstein and R. L. Axtell. *Growing Artificial Societies: Social Science from the Bottom Up (Complex Adaptive Systems)*. The MIT Press, 1996.

[3] C. Gershenson. A general methodology for designing self-organizing systems. *arXiv nlin/0505009*, 2005.

[4] M. Nairat, P. Dahlstedt, and M. G. Nordahl. Character evolution approach to generative storytelling. In *Evolutionary Computation (CEC), 2011 IEEE Congress on*, pages 1258–1263. IEEE, 2011.

[1]The source code of the MADE environment and the algorithms used in this experiment are publicly available in `https://github.com/raiben/made` under a GPL license.