An experiment in engineering a synthetic organism-enterprise

Tom Rijntjes
Graduation Thesis, July 2016
Media Technology MSc program, Leiden University
Supervisor: Bas Haring, Matthew Gladden
tomrijntjes@gmail.com

***Abstract*—This research explores the possibility of developing artificial life that functions as an autonomous business in the real world economy. Two major obstacles are explored: completing the business cycle without human supervision, and whether interacting with the real world economy suffices to drive an artificial life based evolving system. Treating business units as life forms allows us to spawn multiple business units and evolve them based on their endogenous fitness properties. If an instantiation is successful as a business, it will be allowed to spawn new units based a recombined version of its genome. The genome is an encoding of the parameters that determine any or all decisions in the business process. A web application was built and tested to explore the practicalities of this model.**

***Index Terms*—Synthetic Organism-Enterprise, artificial life, business process cycle, genetic programming.**

1.Background

In his 2011 research, Gladden introduces the Synthetic Organism-Enterprise (SOE), a model for exploring the possibility of developing artificial life that functions as an autonomous business in the real world economy [1]. Traditional anthropocentric models of the business cycle do not suffice to describe a synthetic counterpart to a human business. An adapted model was created to think constructively about the business cycle without human interference (see figure 1). In his work, Gladden lists two classes of obstacles that must be surpassed for future development of SOE's.

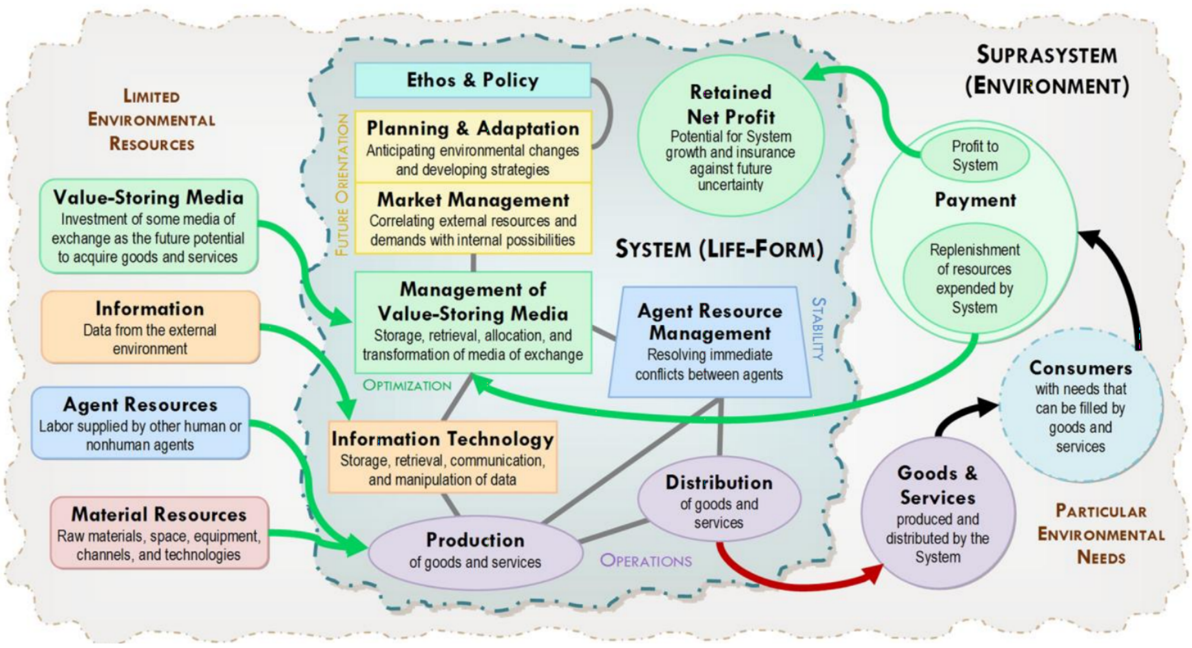


Figure 1. the reconceptualized business cycle applicable to the SOE

The first challenge is to design a business unit of which all the operational processes listed in the model can run independently of human supervision. There must be a means to interface with the real world economy to harvest information, resources and funds. The interface must support delivery of goods and/or services and a means to collect revenue. Secondly, all internal processes for operations, optimization and future development are automated.

Another major question raised by exploratory work is whether the competitive pressure will suffice to drive evolution. Most artificial evolvable systems rely on simulation to quickly assess its capacity to innovate. SOE's are removed from the realm of simulation by definition. It is important to reconsider characteristics of artificial life schemes if evaluation of fitness becomes very expensive in terms of means and time. The current research aims to explore these obstacles by creating a minimal but viable instantiation of the SOE. To operationalize a metric of competition through evolution, it is important to consider the ultimate goal of creating a SOE. Successful SOE's could surpass traditional human businesses in productive and competitive capacity, thus functioning as driver of economic growth. Competing, in this sense, surpassess merely operating in the same market: the SOE is expected to carve out new markets based on its unique characteristics.

1.1Affiliate web shops

The SOE model is potentially applicable to any business of which the business cycle can be completed by machines [1]. One particular business type of this class is the affiliate web shop. The value proposition of affiliate web shops is the aggregation of many web shops in a convenient manner. When a client makes a purchase at one of the connected web shops, the affiliate marketer receives a percentage of the product price. Although margins are small, the affiliate web shop requires no inventory and relies solely on its web infrastructure and marketing pipeline. However, crafting a successful e-commerce platform is a complex task with a myriad of factors that have been hypothesized to predict its survival[2][3]. Affiliate web shops are a subclass of these platforms that rely on aggregation of products to create a highly tailored array to suit the needs of the customer. Primary decisions for crafting such a platform revolve around selecting products from a large array and adapt the marketing strategy accordingly. The selection strategy is often based on market heuristics before being refined by adjusting to market insights of users responding to the platform.

The current research proposes to automate this process by evolving a selection strategy that adapts dynamically to the user’s needs. It is assumed that without human feedback it is impossible to create a ‘pleasing’ selection of products that appeals to consumers because human buying behaviour is difficult to model. Instead, user behaviour can be utilized to evolve the search strategy.

Adapted business cycle	**Affiliate web shop**
Environmental resources	Third party sellers product data
operations	Value creation through aggregation and usability
Future Orientation	-
Optimization	-
Environmental Needs	Influx of means through affiliate fees

1.2 Interactive Evolution

Such an architecture can best be understood within the paradigm of interactive evolution, characterized by hybrid systems of humans directing evolutionary optimization. Interactive evolution is typically used for design problems where the product is difficult to assess for machines, but trivial for human users[4][5]. The current project shares the essential characteristic that the goal is difficult to model; the parameters of a successful affiliate web shop are not merely subtle, they are to be considered unknown and changing. Human users will have no trouble expressing their preference through interaction behaviour.

The major challenge is the fact that fitness evaluation is *very* expensive in terms of both time and means. This drawback can be compensated by the quality of the evolved results [6]. Different types of information are collected from users in typical interactive evolution systems, ranging from categorizing or grading the results to actively selecting individuals for reproduction[4]. However, extensive data collection by surveying the user would seriously constrain future use cases. A basic approach that relies on conversion metrics should suffice to discern successful instances and adapt accordingly.

1.3 Endogenous Fitness

Business units have the property that their fitness is not *immediately apparent.* Intuitive examples are myriad in the form of modern technology enterprises, which might take years before making a profit, whilst still considered successful. To optimize the unit's policy, it must be allowed some time to interact with the market before assessing its quality. Typical evolutionary computing systems rely on generational reproduction. Every iteration the population is renewed based on the calculated fitness of every instance.

This approach is problematic for SOEs for two reasons. The time it takes to bring a business to profitability varies strongly and is not a good predictor for future success. Generational selection after a relatively brief time period would create a bias towards businesses that reach profitability within that particular timeframe, potentially ignoring instances that may create a big impact in the long run. Secondly, if the instances still hold assets, it is inappropriate and expensive to terminate the instance prematurely after one generation. Vice versa, not assessing instances early enough may lead to unnecessarily weak performance of the entire system while enough evidence exists to weed out poor performers.

A number of non-generational selection methods have been described [7], all relying on random selection of pairs. It's informative to look at research in the evolutionary computing field that explores genetic algorithms for artificial life. For example, Ackley and Litman simulated a population of artificial agents in a lattice that seek food and avoid predators[8]. Each agent has a certain 'energy level' that is maintained to survive and reaches a certain threshold before reproduction occurs. There is no *a priori* fitness function, but the fitness emerges from many actions and interactions over the course of the agent's lifetime [9]. The fitness as well as the rate at which the population turns over is "endogenous".

The energy model maps intuitively to a business unit's cash reserve. When the reserves run out, the business perishes. When the business is successful, it is allowed to reinvest its means in franchises. This research considers whether using this metric to drive evolution leads to improved profitability of the entire system. Finally, it is considered whether the SOE model is a practical paradigm for thinking about process optimization for real world applications.

2. Technical details

A web application was created on top of a small database of affiliate marketing products. Users are asked to evaluate the stores by clicking on products that interest them as a proxy for buying behaviour. This conversion data is used to allocate 'energy' to the corresponding instance in the population. A round robin mechanism redirects the users to the different variations of the store. Every fifth session is part of the control group with a random selection of products.

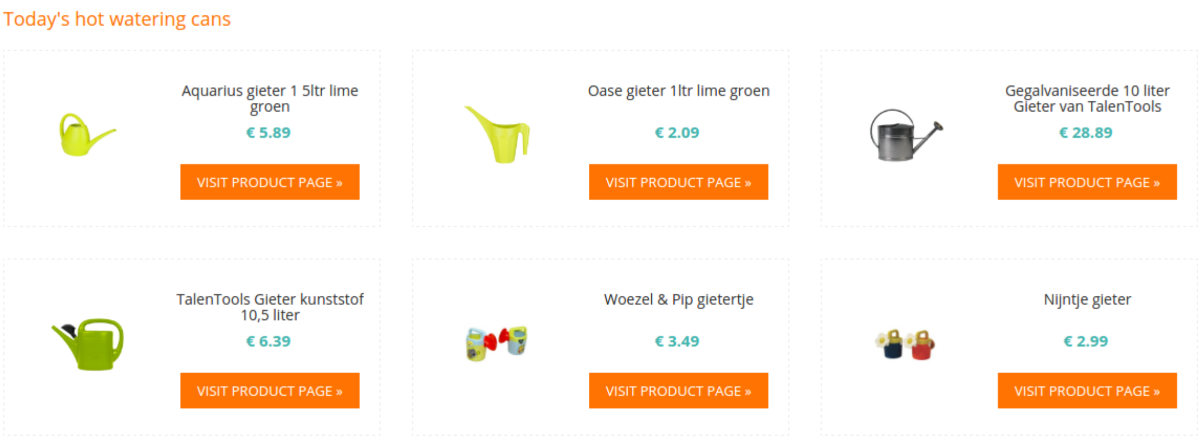


Figure 2. Screenshot of the web application

Between the stores, the product composition varies. A genetic programming engine is used to generate and recombine SQL queries. Every instance is a parse tree grown with the following building blocks:

Primitives	<,AND,OR
Terminals	Price,price_old,volume,colourfulness

Table 1. Primitives and Terminals of the Genetic Programming core

For example, the parse tree in illustration 1 maps to the following SQL query:

SELECT id FROM gieters WHERE (price < 30.76) AND ((price_old < 8.54) OR (price < 28.97))

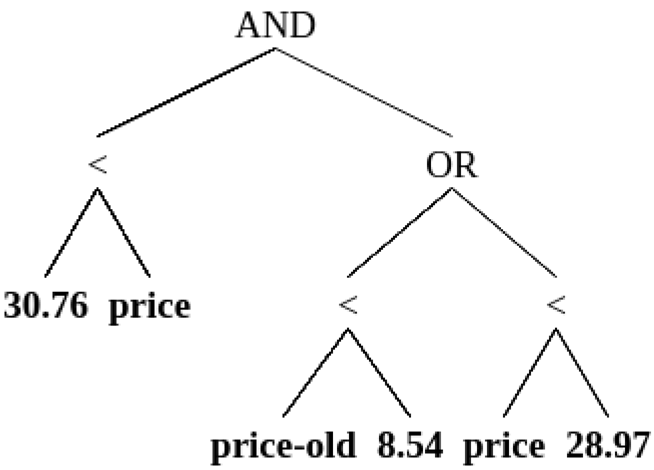


Figure 3. Example parse tree

Four attributes are assumed to be of potential importance when selecting a watering can for private use: price, volume, colourfulness and discount. Colourfulness was modeled using Hasler's metric [10], a method that uses an image as input and returns a colourfulness value based on color variety within the image. The metric is scaled to a range of 0 to 40 for convenience. The volume was retrieved from the product description manually.

A 'ephemeral constant' is added to the primitive set, a random variable that maintains the same value after recombination. The ephemeral constant can take a values between 0.00 and 40.00 with a precision of two decimal places. Given 4 variables, this leads to an effective search space of 4000^4 possible combinations.

Inevitably, the result set of queries varies strongly between 0 and 76 results. To control for simply having the largest offering, a sample of 9 products that's consistent between sessions is taken from result sets that exceed 9 products in size. The full code is available online[11], including the third party library for genetic programming[12].

2.1 Fertility based interaction

GP typically uses fitness proportionate selection or over-selection for large populations as a mechanism for selecting parents [6]. To explore the non-generational nature of business activity, the current research defines a non-generational selection method.

The selection mechanism is inspired by local interactions in artificial life systems. Every instance can operate independently of a supervisor by searching for other fertile partners independently. This design choice is a first step towards fully distributed species of digital SOE, potentially living on different servers that interact asynchronously without a master unit controlling selection.

Each instance internal energy is chosen to resemble its cash flow. When the cash reserves run out, the instance perishes. When enough funds have accumulated to branch out, a new 'start up' is launched, and so on.

When the a priori determined threshold is reached, the unit is added to a reproduction queue. If the reproduction queue holds two units, their credit is reduced by the reproduction cost and a new unit is instantiated based on crossing over the parent's genome. Each visit reduces the visited instance energy by a small amount to account for opportunity costs. Units that fall below zero are removed from the population. Every fifth visitor is directed to a control group of randomly generated phenotypes.

Business concept	Artificial Life concept	SOE	Energy change
Startup cost	Starting energy	Every instance begins with a base energy	0.50
Launching a new branch	Reproduction cost	Two parents pay half the startup cost to create a new instance	-0.25
Sale	Food collection	An energy gain is calculated based on the price of the product	price*0.0025
Opportunity cost	Energy decay	Every new session reduces the instance's energy by a small amount	-0.025

Table 2. Concept mapping

3. Results

3.1 Conversion

The instances subject to the optimization (n=652) performed better than the control group (n=157) with an average clickthrough rate of 49,3% with a standard error of 3,8%. The control group shows an average clickthrough rate of 36,9% with a standard error of 7,5%. In both cases, with a reliability of

95% the true conversion falls within the respective confidence intervals of 45.5%-53.2% and 29.4%-44.5%.

The optimization is hypothesized to perform no better than random before improving gradually by adapting to user behaviour. When comparing the first 322 non-control sessions to the last 329 non-control sessions, no such improvement can be found. The first group shows an average conversion of 48,3% versus 48.6% in the second group. The null hypothesis of equal means cannot be rejected given a student's t-test under the assumption of equal variance with a p-value of 0.97.

The new instances are expected to outperform the base population. The second generation instances 20-27 (n=73) show a higher conversion average of 68.3% (standard deviation of 18.9%) as compared to 0-19 (n=557) with 46.8% (standard deviation 34.2%), but fails to reject the null hypothesis of Student's t-test for equal means with a p-metric of 0.1286.

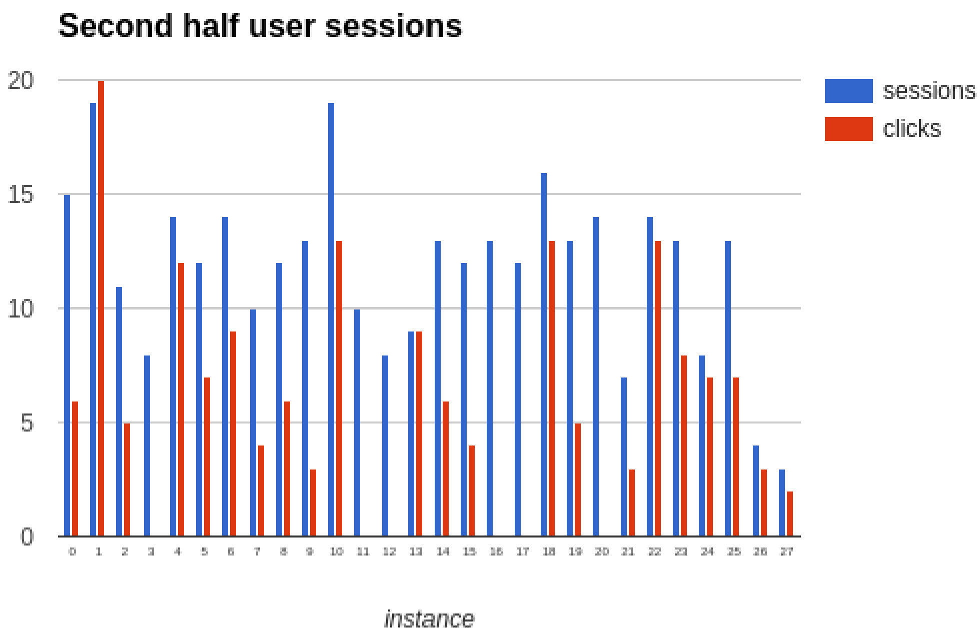


Figure 5. Second half user sessions: sessions and clicks per instance

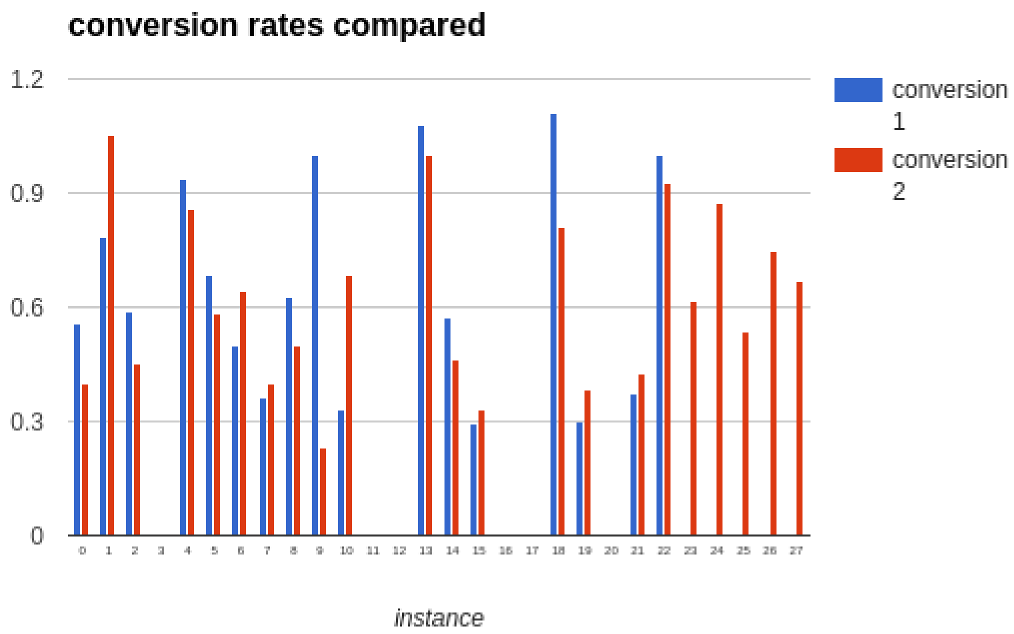


Figure 6. Conversion rates of the first half and second half compared per instance

3.2 Population dynamics

After 652 sessions, 8 new instances were created and 5 instances perished. This population turnover is extremely small compared to typical genetic programming experiments.

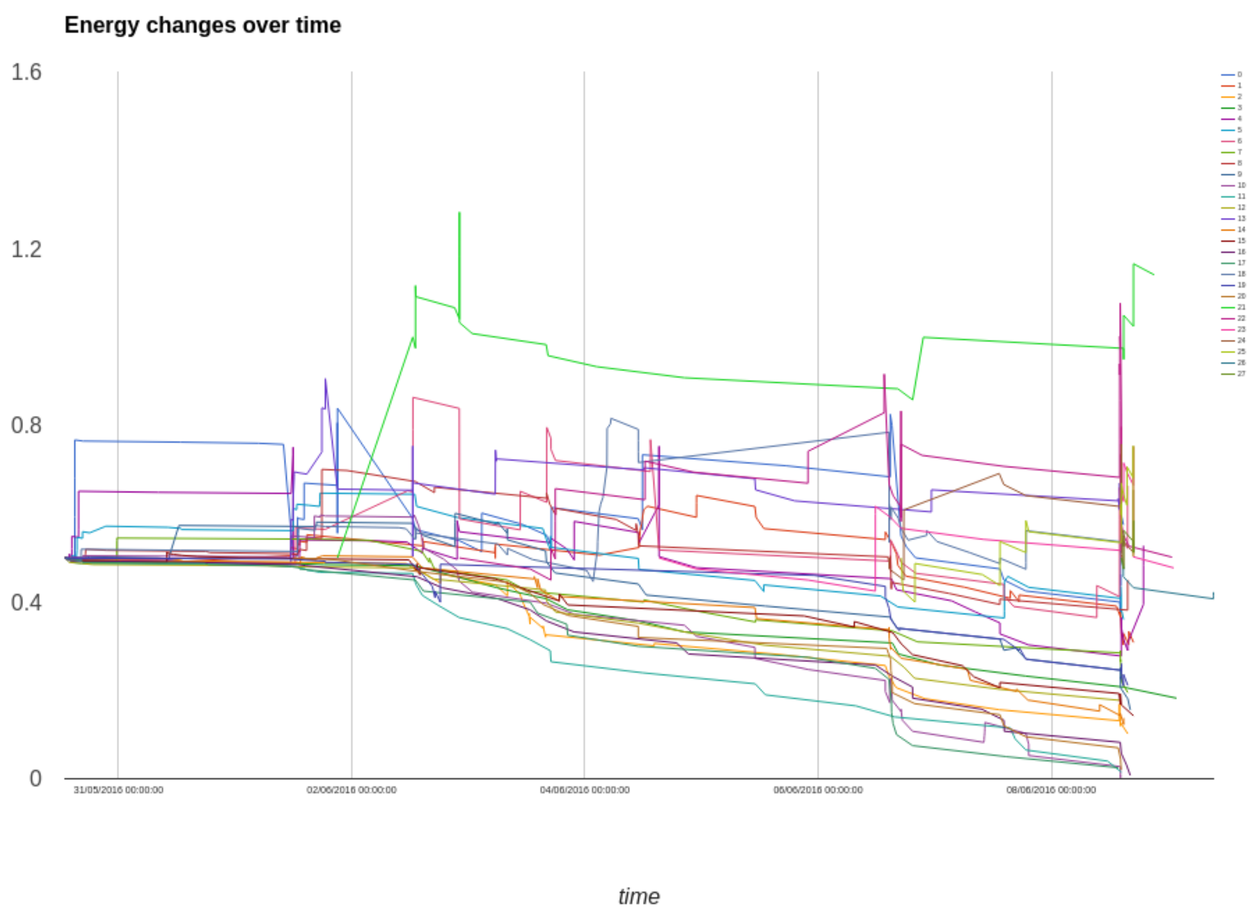
No selection was made on the meaningfulness of the automatically generated queries in the baseline population. For example, instance 1 has genotype '14.84 < 23.95' which is always true and returns the entire range of products. In the experimental setup, sets larger than 9 products were trimmed in a random but fixed manner. The following genotypes are semantically meaningless:

instance	query	results
1	14.84 < 23.95	76
3	21.72 < 17.14	0
11	((6.79 < 38.99) OR (((volume < price) OR (volume < price_old)) OR (volume < price_old))) AND (price < price)	0
16	colourfulness < colourfulness	0
17	price < price	0

Table 3. All semantically meaningless genotypes

Furthermore, instance 5,10 and 12 returned either all results or no results. In total, 40% of the initial set can be considered a dud by contributing no meaningful selection. It was expected these instances would be quickly removed from the population. Appendix 1 shows an overview of the lifetime energy of all instances. The perished instances are 10,11,16,17 and 20. Given a starting energy of 0.50 and an opportunity cost of 0.025, every empty instance will be displayed to a user at maximum 20 times.

Meaninglessly full sets 1 and 10 performed below average with a terminating energy of 0.30 and 0.0075 respectively without reproducing.


Energy changes over time
1.6
1.2
0.8
0.4
0
31/05/2016 00:00:00
02/06/2016 00:00:00
04/06/2016 00:00:00
06/06/2016 00:00:00
08/06/2016 00:00:00
time
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

3.3 Successful instances

Reproduction occurs when two instances have an energy level of 0.75 or higher at the same time. The following instances reproduced at least once.

instance	reproduced	genotype	Selection size
0	3	30.47 < price	11
4	2	((volume < 11.07) OR (colourfulness < price)) AND (colourfulness < 0.17)	7
6	3	28.82 < price	13
13	1	(((price_old < price_old) AND (colourfulness < price)) OR (39.58 < price_old)) OR ((price < colourfulness) OR (colourfulness < colourfulness))	45
18	1	colourfulness < volume	24
22	4	28.82 < price	6
23	1	((volume < 11.07) OR (colourfulness < 28.82)) AND (colourfulness < 0.17)	7
25	1	28.82 < price	13

Table 4. Fertile instances

None of the meaningless genotypes reproduced. Note that 11 out of 16 parents only distinguish products based on their price. This is likely to be a consequence of the reward function based on price; clicking an expensive yields more energy for the instance.

4. Discussion

The current research grants a first look into the practicalities of creating a minimal SOE. First, we consider whether the experimental setup completed the business cycle by reviewing the model in figure 1. By interfacing with human users through web, the experimental web application delivered value by aggregating a selection of products to its users. This value was monetized through the affiliate marketing model by earning a percentage of each sale on third party websites. The system collected information by accessing an externally maintained database. The actual conversion of a page visit to a sale was registered, but not incorporated in the system. A virtual currency that closely resembled actual cash flow was used by accepting an average conversion based on real world observations. Internal processes were highly simplified: each unit created a static offering and did not plan or adapt within its lifetime. A simple rule-based allocation mechanism was used for reproduction. After reaching a fixed threshold, the business unit was allowed to reproduce. Thus, planning and adaptation occurred externally by passing on its genetic blueprint to new instances. However, none of the processes turned out to be an unsurmountable obstacle, even in the light of limited means and an experimental context. It is highly likely more sophisticated systems can emerge that are capable of completing the business cycle autonomously.

Secondly, we consider whether competition within the human economy suffices to drive evolution. During the experiment, the overall performance of the population did not improve. Although the sample size was too small for definitive conclusions, all instances not part of the first generation showed more meaningful genotypes and included the most successful phenotype in terms of fertility. It must be stated that although the quality of intergenerational improvement can potentially outweigh the expensiveness of optimization [6], this analysis is truly microscopic as compared to traditional evolutionary computing experiments. Longer optimization runs can show definitive results.

Moving from a model to a workable design, a number of observations emerged. The model is useful when designing a metasystem with the characteristic that pooled resources are inappropriate. If it is required that every business unit can control its own cash flow, the adapted business cycle is a useful way to govern means across the entire system. Some businesses operate in highly diverse or highly dynamic environments, making it troublesome to create a coherent policy for every environment. The SOE model presents a framework for designing a metapolicy that guides local adaptations in a controllable manner, thus balancing internal consistency and flexibility. Finally, the system is inherently distributed, enabling a robust commercial exploration system.

It is conceivable that a larger scale species or network of affiliate businesses run through the principles listed above would compete with traditional human businesses in the same market. No fundamental objection was encountered during the development of this approach.

5. Future Work

In the present work it is investigated whether thinking about web based enterprises as artificial life forms is of practical and conceptual use. It is shown that a business 'species' can be built with limited means.

5.1 Appropriate Selection Mechanisms

This research took a jab at finding an appropriate selection mechanism that allows for asynchronous fitness evaluation and a distributed architecture. Typical non-generational selection mechanisms are well understood [13], but rely on a supervising master unit, which has undesirable design consequences. The issue of immaturity was addressed by the Avida project. An interaction threshold was introduced where the instances were only recorded after 10.000 interactions [14]. Future work could be directed at finding and investigating a selection mechanism that is appropriate to the requirements of the SOE, whilst having predictable characteristics with regard to turnover rate and convergence for known search spaces.

5.2 Digital enterprises

In this research, an artificial life selection mechanism was employed to continuously and systematically explore a commercial space, focusing on product mix composition. Future work can be directed at modelling larger components of business units by incorporating other data streams. The challenges that apply to artificial life also apply to digital SOEs, with the exception that the environment it operates in is boundlessly complex. It would be highly interesting to investigate possibilities with regard to open-endedness research by developing a robust parsing engine that can leverage new data streams as they emerge. Another direction can be the interaction between evolution and lifetime learning, also known as the Baldwin effect[8][9]. This has been researched in simulations, but never when interacting with more complex environments.

During the experiment, 20 non-identical copies of the same business were launched. To the user, it was irrelevant to think of each business as an individual that competed with the others for means. However, it may be useful to think of franchise models in the same manner. By designing and/or automating the ways in which each franchise adapts to its environment, this approach can lead to highly scalable governance systems. For example, consider the case of networked vending machines. These systems require very little human interference to create value, so a generalized governance system might be applicable. A learning system can be designed centrally to have each instance adapt to its environment in a predictable and controlled manner, thus having it adapt systematically to its local context. There is no upper boundary to the level of autonomy that can be granted to the individual business units: managing its own funds, experimenting with marketing, branching out or salvaging poorly performing instances.

As hypothesized by Gladden, the SOE can be a driver of economic growth by developing new markets by the properties unique to machine entrepreneurs [1]. Entrepreneurship is an inherently creative process. SOE can be approached from the perspective of artificial creativity, a field dedicated to understanding creative processes by modelling and automating these processes. Keane et al. designed an absurd looking but highly effective satellite boom using a genetic algorithm [15]. This research is an example of how evolutionary algorithms are not limited by human preconceptions and implicit assumptions and can thus be truly creative in finding solutions. If innovation is understood as driver of economic growth, employing creative systems like the SOE to systematically explore innovative business opportunities is a path worth pursuing.

6. Acknowledgements

Many thanks to Roald Wegh and Bouke Versteegh for supplying precious data and market insights from their startup company Shopsaloon. I want to thank Sietse Huisman, Judith Westgeest, Han Kortekaas, Justin van der Eerden, Ainara Alonso, Renée Geuze, Derk Overduin, Maarten van Hees, Jeroen van Oorschot, Jules Verdijk, Jos Rijntjes, Patrick Rietschoten, Teun Durlinger, Lieven Slenders, Niels Casteel, Robin Van der Heijden, Krisjan van Laarhoven, Willem van der Pas, Justus Olijslager, Mario Di Florio, Thijs Slijkhuis, Donna Schipper, Eric Trillo, Fabiola Fuzy, Henk Peels, Mark Loonen, Nicole van der Spank, Chetan Mittal, Zoltan Jeszensky, Liduine Knubben and Arthur Poot for kindly participating in the experiment, testing the web application and invaluable feedback. Finally I want to thank my supervisors Bas Haring and Matthew Gladden for helping me turn a big idea into actionable research.

7. References

1. Gladden ME. The Artificial Life-Form as Entrepreneur: Synthetic Organism-Enterprises and the Reconceptualization of Business. ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems. 2011. pp. 417–418.

2. Chen JV, Rungruengsamrit D, Rajkumar TM, Yen DC. Success of electronic commerce Web sites: A comparative study in two countries. Information & Management. 2013/9;50: 344–355.

3. Park EJ, Kim EY, Funches VM, Foxx W. Apparel product attributes, web browsing, and e-impulse buying on shopping websites. J Bus Res. 2012;65: 1583–1589.

4. Breukelaar R, Emmerich M, Bäck T. On Interactive Evolution Strategies. In: Rothlauf F, Branke J, Cagnoni S, Costa E, Cotta C, Drechsler R, et al., editors. Applications of Evolutionary Computing. Springer Berlin Heidelberg; 2006. pp. 530–541.

5. Secretan J, Beato N, D’Ambrosio DB, Rodriguez A, Campbell A, Folsom-Kovarik JT, et al. Picbreeder: a case study in collaborative evolutionary exploration of design space. Evol Comput. 2011;19: 373–403.

6. Eiben AE, Smith JE. Introduction to Evolutionary Computing: Springer Berlin Heidelberg; 2015.

7. Rudolph G. Takeover times and probabilities of non-generational selection rules [Internet]. Unversität Dortmund; 2001. Available: https://eldorado.tu-dortmund.de/handle/2003/5382

8. Ackley D, Littman M. Interactions between learning and evolution. Artificial life II. 1991;10: 487–509.

9. Mitchell M, Forrest S. Genetic Algorithms and Artificial Life. Artif Life. 1994;1: 267–289.

10. Hasler D, Suesstrunk SE. Measuring colorfulness in natural images. Electronic Imaging 2003. International Society for Optics and Photonics; 2003. pp. 87–95.

11. Rijntjes T. Robopreneur [Internet]. 2016. Available: https://github.com/tomrijntjes/robopreneur

12. Fortin F-A, Rainville F-MD, Gardner M-A, Parizeau M, Gagné C. DEAP: Evolutionary Algorithms Made Easy. J Mach Learn Res. 2012;13: 2171–2175.

13. Rudolph G. Takeover Times of Noisy Non-Generational Selection Rules that Undo Extinction. In: Kůrková V, Neruda R, Kárný M, Steele NC, editors. Artificial Neural Nets and Genetic Algorithms. Springer Vienna; 2001. pp. 268–271.

14. Adami C, Brown CT. Evolutionary learning in the 2D artificial life system “Avida.” Artificial life IV. Cambridge, MA: MIT Press; 1994. pp. 377–381.

15. Keane AJ. The design of a satellite boom with enhanced vibration performance using genetic algorithm and other optimization techniques. Parmee and Denham [867], page. 1996;