# Co-operative OuLiPian Generative Literature using Human Based Evolutionary Computing

Michelle Okaley Hammond
London South Bank University
103 Borough Road
London SE1 0AA
+44 (0) 20 7815 7477
michelle.hammond@lsbu.ac.uk

Terence Claus Fogarty
London South Bank University
103 Borough Road
London, SE1 0AA
+44 (0) 20 7815 7482
fogarttc@lsbu.ac.uk

## ABSTRACT
The tools and techniques employed by Interactive Evolutionary Computing [21] offer procedures which put human interaction at the centre of the problem solving process. Human Based GA systems [7], [8] such as the Automatic Concept Evolver (ACE) [6] offer an infrastructure that has been proven to work efficiently in solving problems which involve the evolution of natural language strings. Oulipo is a French literary movement founded in the 1960s which applies mathematical constraints in the creation of literature making them distinct from other experimental literary groups [10]. Their literary creations change the dynamics of the author-reader relationship because they provide a different experience from reading traditional narratives where intuition and emergence are fundamental to creative exploration. This paper proposes the application of the ACE methodology to arbitrate between a group of interacting authors to produce Oulipian literature and specifies experiment to test this approach in practice. It is proposed, that ACE can be likened to an Oulipian constraint and thus the experimental output will allow the reader to choose a path through the narrative, exploring new meaning.

## Categories and Subject Descriptors
J.5 Arts and Humanities; K.3 Computers and Education; or K.3.1 Collaborative learning; 1.2.m Artificial Intelligence

## General Terms
Algorithms; Design and Experimentation

## Keywords
Interactive Genetic Algorithms; Human Based Genetic Algorithms; OuLiPo, Creativity and Generative Literature



## 1. EVOLUTIONARY COMPUTING AND CULTURE
Whilst optimization is a major field of Evolutionary Computing (EC) there is increasing interest in using Evolutionary Computing for more ill-defend problems, such as the artistic exploration. This is evident in the growing application of EC techniques in fields such as art and music generation, analysis and interpretation, architecture and design. In the evolution of pieces of music and art the processes of survival and recombination can be seen and the tools of operation are both appealing and easy to use. In EC, the principle of emergence is fundamental as it is to any kind of creativity. Soddu [20] believes that Genetic Algorithms (GA) are identifiable as one of the most advanced approaches in this field because they perform "the incoming new naturality of the artificial world". What Soddu calls generative art (which can include both music and art works) is the idea realized as genetic code of artificial objects. By transforming generative codes, generative art realizes as natural DNA does, always different and unpredictable series of events, pictures, objects, music, etc.

The approach is interdisciplinary and tends to blur the boundaries between science and art, but has provided tools which give rise to new types of artworks and aesthetics by allowing new forms of interaction between humans and artificial agents, redefining the way art is created. However, with creative systems we are dealing with ill-defined problems, which cannot be easily tackled by conventional GA's due to the subjective nature of their evaluation functions. Instead humans are used to evaluate the potential solutions. This has led to a growing interest in interactive evolutionary systems research, which aims to help humans with creative explorations [7] or "learn" about human creativity and aesthetics from the humans' interaction [19].

## 2. INTERACTIVE EVOLUTIONARY COMPUTING
It is well documented that conventional applications of the GA are not capable of utilizing human impressions, intuition and emotions. The first IEC systems were Interactive Genetic Algorithms (IGA's) and appeared as a result of attempts to apply evolutionary computing techniques to artistic domains where it was unclear how to define fitness functions [7]. The history of research relating to Interactive Evolutionary Computing (IEC) in the main, relates to partial or complete human evaluation of the fitness of solutions generated from an evolutionary search. This has generally been introduced where quantitative evaluation is

difficult if not impossible to achieve [11] as opposed to the optimization performances that are numerically based [21]. IEC is an evolutionary optimization technique in which a human user replaces the fitness function of the system, i.e. the user evaluates the individuals of the population subjectively with each interaction. This fitness is used to select the parents for the next generation.

## 3. HUMAN BASED EVOLUTIONARY ALGORITHMS

The HBGA model provides a communication infrastructure that manages information via its interface [8] and delegates all the usual functions of the GA (selection, fitness evaluation, crossover and mutation) to external human agents. A version of the HBGA was used in the Free Knowledge Exchange (FKE) project for collaborative web-based problem-solving [7]. The FKE project evolves strings of natural language to arrive at better solutions to problems submitted by it participants. The use of natural language as a genotype representation here was inspired by Dawkin's meme theory [4]. The structure of natural language makes the evolutionary method of natural language processing efficient because language strings have evolved to allow a tight linkage of building blocks since in the most frequently used patterns of language, their constituent parts tend to be located close to each other [7].

In the FKE system, text-based questions and answers (chromosomes) of words from a natural language (genes) compose the knowledge database. The system starts with an empty population, to which the first ideas / problems proposed by participants are added. Each answer is then evaluated based on the number of participants interested in a particular answer or idea (human evaluation of fitness). A problem's fitness determines the probability of its selection, so problems with higher fitness (i.e. in which many people are interested) appear more frequently.

Five problems are presented at a time with the most popular solutions to the problem following each offering a system of ranking the population of solutions. Forum participants review a combination of questions and answers and are then invited to either agree with an existing solution (thus increasing the fitness of that solution) or use parts of different solutions in their answer (crossover) or be inspired by previously submitted answers and propose an entirely new answer or idea (mutation). New answers or ideas (offspring) are generated by this process.

For Kosorukoff and Goldberg [7], [8] by fusing these two processes of computational and human innovation, the HBGA model provides a natural method of embedding the competence of human users into an evolutionary procedure; and provides a method of studying innovative behaviour of humans. The HBGA model allows it to be captured and if this can be implemented at the computational level, they believe it will bring us closer to computational intelligence that can be called creative.

## 4. AUTOMATIC CONCEPT EVOLUTION

The Automated Concept Evolver (ACE) originated from an idea of Fogarty when he sought to engineer a GA as a tool that would arbitrate between a group of geographically dispersed individuals seeking to cooperatively solve a problem. Fogarty [6] was particularly interested in using the Genetic Algorithm to explore and support creativity. Fogarty's GA inadvertently presents us with a HBGA model. The automation of Fogarty's proposed GA as an experimental system would help validate the procedures. The first experiment implementation was done by email and the subsequent automation of the process would help develop the idea as a methodology for producing generative literature.

The evolutionary algorithm conceived by Fogarty [6] would moderate the creative input of a distributed group of interacting people to cooperatively produce a solution to a problem. By applying the evolutionary process of "natural selection" the GA would ensure the optimum solution to a given problem from a pool of all possible solutions. Once the problem has been defined, each member of a group suggests solutions to the problem. They then go on to improve the suggestions by "mutation" and "crossover". They can show their preferences for solutions to problems already suggested by "replication" and can "create" new suggestions when they arise. Candidates' solutions for the next population are weighted according to the weights of their parents, if the have any, and a proportion of the new population is probabilistically chosen from them according to their weights.

The problem was to determine the "best" name for the then un-named algorithm. Suggestions were sent by email by participants and operations were performed manually. As with IEC systems, participants in the experiment would subjectively and implicitly evaluate the fitness of individuals in the population by selecting them to work on to produce members for the next generation. An individual in the population is assigned a weight depending on how successful it is in the selection process and displayed.

The individuals in the population are ideas. They were represented as strings. Each participant had to suggest one or more names of the algorithm, in addition to the original ideas, or use genetic operations to create new candidates. Each suggestion would then be weighted according to how popular it has been with the members of the group and those with the highest weight would then be brought forward to the next generational run.

A subsequent automation of ACE [5] was applied to the creative task of story writing and developing an advertising slogan by a group of geographically dispersed authors. The results of the experiment showed that ACE as with the email experiment design, succeeding in regulating - the population size presenting ideas of phenotypic selection as is inspired by nature; a diverse distribution of the GA on all runs; and there was no evidence of propagating an individual solution.

ACE offers an HBGA model of the IEC system because it facilitates the organization of individuals into collaborative communities to explore the evolution of solutions through the subjective evaluation of potential solutions and by performing intelligent crossover, mutation and selection on existing knowledge. As with the FKE experiment [7], the ACE experiment found a spirit of collaboration between users was created because they acknowledged some measure of their contribution in the final result.

In the ACE system, fitness is an implicit consequence of the selection process. In the evolution of natural language strings, the fitness of an individual in the population is manifested though its phenotype – meaning. Genetic operators operate on the structure (or syntax) of a string to evolve new genotype. Because meaning of any language string is usually context dependent – there are numerous representations of phenotype and genotype mappings.

For ACE to be extended to the evolution of more complex language structures and larger group sizes of users, different methods of controlling population explosion and interface design have to be considered if the system is to be effective in avoiding human and IEC fatigue. The different approaches as outlined by Kosorukoff [7] and Takagi [21] have to be carefully considered in order to offer the best balance of human/computer interaction to support the creative writing process within a group of people. However, in the HBGA model using ACE, users need a degree of creative freedom in order to improvise and explore their ideas.

## 5. OULIPO

The Oulipo - an acronym for *Ouvroir de littérature potientielle* - Workshop for Potential Literature was co-founded in Paris in the early 1960's by Raymound Queneau and Francois Le Lionnias. Oulipian writers impose constraints that must be satisfied to complete a text, constraints ranging across all levels of composition, from elements of plot or structure down to rules regarding letters. Oulipo pushes a structuralist conception of language to a level of mathematical precision. The techniques they apply become technical because language as the field of investigation, becomes a complex system made up of a finite number of components. For Mathews and Brotchie [10] the informing idea behind Oulipian work is the writer, who forces a linguistic system and /or literary genre out of its habitual mode of functioning to produce a new literary form.

Oulipo is still an active and significant literary movement and continues to meet once a month in Paris holding monthly public workshops. Oulipo is popular because their approach to literature creation challenges the artists to explore methods for generating novel thoughts by applying processes alien to the human mind's "normal" way of creating ideas through conditioned association. Mathematical equations are usually at the base of their constraints, Oulipian's also pay tribute to literary history by declaring all structures of all various genres of past eras open to innovation. Their algorithmic or rule-based methodology twists language in new directions to reveal unusual and unpredictable expressions and meanings. Oulipians produce works rich in pattern which call the reader to take on a role not usually needed with traditional literary texts. By offering the reader a choice of paths through the narrative, they offer a structure within which the reader can juxtapose literary elements and experiment with emergent meaning.

In 1961, the first Oulipian publication was Queneau's *Cent Mille Milliard de Pèomes* (One Hundred Thousand Billion Poems), which is regarded as exemplifying "combinative literature" [10]. Queneau [15] expounded that computer power was the only thing that could truly explore literatures combinative potential.

Members of the Oulipo literary movement include writers, logicians and mathematicians. The movement arose through discussions of Queneau's *Cent Mille Milliard de Pèomes* which used predetermined rules in its creation opposed to the "ecriture automatique" by the Surrealists [14]. This book of ten sonnets, each fulfilling the formal structure of the sonnet – 14 line piece with a specific rhyming pattern and meter. Each line in a given sonnet can be read with any of the other sonnets, creating a number of possible variations, which Queneau calculated as ten to the power of fourteenth [15]. He constructed the book by "placing the sonnets on top of the other, cut each one into strips so that the reader can open the fourteen lines of each poem and combine the lines of all the poems." [22] Many saw Queneau's *Cent Mille Milliard de Pèomes* as exemplifying the notion of a poetic structure that can be recombined to form new poems. François Lionnais in his post-face to this work in 1961 used the expression "*combinative literature*" to define the method of creation. The discipline of 'combinatorics' in mathematics has been proved to be demonstrably functional in many literary structures from classics to modern literary productions. For Seaman [18] Oulipo's combinatoric methods and experimental concepts in literary and artistic creation continue to be generative and relevant today. Computerised versions of Queneau's *Cent Mille Milliard de Pèomes* improved the effectiveness of the combinative work of the reader and in some cases, widened the field of combinatorial procedures particularly in the realm of 'data-processing' and in the new cultures of generative literature, such as, hypertext, Interactive Fiction, etc.

Oulipian literature offers several different kinds of ways to view and understand writing produced under constraints, which they believe stimulate the creativity and imagination of the writer. Some works provide the opportunity to appreciate the way in which the constraints that govern a text are not noticeable – eg, Perec's *Life, A Users Manual* [12] or Calvino's *If on a winter night a traveler* [3]. Other texts constantly refer to their own constraints – e.g., Perec's *La Disparitio* [13] or Roubaud's *Le grand incendie de Londres* [17]. In Perec's *La Disparition* [13] a lipogramic constraint is applied. In this book, the lipogram is propagated across different linguistic levels: graphical, morphological, lexical, syntactic, and semantic. The book is written without the letter 'e' but constantly refers to the vowel's disappearance. The letter "e" is amputated from the language by a series of word games just as the characters in the novel are eliminated by a series of tragedies that lean towards the absurd [10].

Queneau's *Exercises de Style* [16] is a short story which can be permutated 99 times, from zoological terminology to backslang. The "$S+n$" algorithm is a truly Oulipian invention. The algorithm works by replacing every noun (substantive) in a text with the word that falls $n^{th}$ places ahead of it in a given dictionary ($n$ can be any integer) [10].

This exploration of the original literary material to find a creative re-understanding has led Oulipians to be lovingly dubbed "anticipatory plagiarists" by Mathews and Brotchie [10].

## 6. ACE AS AN OULIPIAN EXPERIMENT

Oulipian methods of combinatorics and constraint have lent themselves to computational techniques and attempts to develop a literary machine that could produce Oulipian text are well documented [1]. Using the ACE system to produce Oulipian literature counters the critique of Oulipo's automation of literary production as proposed by Wolff [23]. The ACE system does not put the humans and the computer on the same ontological footing. Alternatively, ACE offers a system in which humans do not lose control of the creative activity and so maintain the derivative meaning of a piece of text which was notably Wolff's criticism of Queneau's *Cent Mille Milliard de Pèomes*. In the ACE system, the role of the human user is firmly embedded within dual processes: the evolutionary process and the gestation of the text being produced. The aim is not to automate the creative process but to provide tools to support humans engaged in it.

In both ACE and the Oulipian processes, the interaction with the text is dependent on the author's selection and fitness evaluation. In combinatorial literature, in order to combine diverse elements into a whole by arranging the elements into a set, a process of selection and evaluation has to be undertaken. Re-mapping these objects within the set is done at the same time as fulfilling some predetermined rule and is analogous to mutation. The reworking of past literary forms by re-contextualising them is analogous to crossover. Fitness is usually determined by the syntactic structure of the text and when a genetic operator has been applied, represents a candidate phenotype – the new literary form that might or might not have some derivative meaning. One level of Oulipian constraints is at the level of the genetic operators - crossover and mutation.

As with evolution, in the ACE system the main components are population and fitness and this is what allows a group of people to work together. Together the users work on the same pool of ideas – selecting from it and replacing old ideas with new ones. It is this process that allows them to cooperate. The selection process is dynamic but the population is central because it provides the base on which genetic operations can be applied. The evaluation of the fitness of the population is undertaken by the people participating in the process of selection. This requires human discipline on how an operator would influence any selected individual. The effect of an Oulipian constraint in ACE depends on the type of problem that is being solved. If the authors were attempting to produce lipogramic text, for example, the constraint would be at the author level. When performing crossover – the human user evaluates how the operator would influence the offspring's phenotypic meaning. Oulipian authors perform crossover and mutation. Each author interacts with a population of ideas through selection. However, it is this notion of having a population of ideas that enforces cooperation between individual authors. Each author is trying to be supportive whilst at the same time trying to determine their contribution to the final solution. In both ACE and Oulipo, evolution is fundamental regardless of any other techniques are introduced.

The problem domain is defined by the group of people using the system. They agree the context and representation of the problem and create the initial population. Their use of genetic operators is dependent upon the representation of the problem as with conventional GA's. It is this representation of the problem to be solved that determines the context. Textually, it could be sequences of events relating to plot, actions or dialogue. In poetry, it could it could be rhyme patterns, imagery, etc. In music, it could be chords, melodic structure, etc. In all cases, the formal constraint (including syntactic ones such as poetic structure) would be defined as part of the problem representation which authors agree on.

ACE is not intended to replace the humans as authors but to offer support in the writing process. Literary devices, such as iconic features of a particular genre similar to the "betrayal" knowledge base in BRUTUS.1 storytelling computer [2] could be accessible to help spark the imagination of the authors in their creative task. In this way, ACE is supporting the "anticipatory plagiarist" credentials of the authors. Other language support systems (such as a dictionary and grammar checkers) could also be included. It is expected that such interface aids, would have a varying level of significance depending on the problem domain and the user make up. This level of processing is expected to come into play post-selection of population individuals, when the author is rearranging content for genetic operations. The interaction between the authors and the work will include parallel processing activities.

Initial contributions from individuals populate ACE. The individual population is composed of pieces of text contributed by the participating authors. Authors select pieces of text to work on from the population. They do this synchronously or asynchronously. Selection; the strength of members of the population (pieces of text) is increased as the author selects them to work on. The fitness of individuals in the population is the way interaction between the authors is arbitrated.

Once authors have selected pieces of text to work on they are constrained by the processes of mutation and crossover to produce Oulipian text. Mutation allows a change in (or of) one word to create a offspring while crossover allows cutting and pasting from two parents to create one offspring. At this level of processing, grammar checkers, semantic checkers, can be included. You could have a question, vis-a-vis word, where the authors are corrected if their creations meet current criteria but authors should be allowed to override corrections. You could allow here knowledge bases which offer suggestions relating to a particular literary style, etc. Once satisfied, authors replace old members of the population with new creations.

The experiment should involve a larger group of people, ideally 10 or more. Authors define the domain – the literary genre, the direction of plot, development of characters, etc. Oulipo constraints could be applied at this level. The system has to rely on the discipline of individual authors to apply the constraints intelligently. However, the computer could undertake this operation. Large groups of people with differing abilities could be supported with dictionaries / grammar checkers to help ensure that the individual authors' suggestions are valid.

## 7. EXPERIMENTAL OUTPUT

The output of the Oulipian experiment could take different forms. It could be the final population, or all the populations produced during the experiment. In Queneau's *Cent Mille Milliard de Pèomes* the output is all the possible combinations of the fourteen lines of ten sonnets and in this way the final text could be presented as a history of the whole process. The order of each population in the output could be based on ranking or there could be some other Oulipian constraint on the ordering of the population. One could, for instance, order the population based on closeness defined by family relations, i.e. siblings are closer than cousins. The final output could be presented in the form of a family tree. The reader of the text is then given a choice of pathways through the text allowing them to explore meaning as the context changes.

### 7.1 ACE writing experiment

An experiment supporting a group of creative writers producing text is envisioned. Many classes of creative writing undertake writing experiments in groups. For instance, a group will be given a scenario with a description of the different characters in the scenario. Each member of the group will be assigned a character and asked to write part of the scene involving that character. This is a little like the task undertaken by a single script writer for a soap opera in which they will write one episode given an outline of the plot. This is a top down approach. In creative writing the approach is bottom up. ACE will support a group of writers

constructing a scene in a creative way. Each writer will contribute text, characters and bits of dialogue to the initial story. They will then have the opportunity of selecting text, characters and bits of dialogue to work on. Those pieces which are selected more than others will come to the forefront and become dominant. This process will enable the authors to cooperate in the creative endeavor in an interesting, novel and principled way.

## 8. CONCLUSION

The ACE system proposes a HBGA model of innovation [9] managing the interaction of humans in the process of writing literature. The aim is not to automate the creative process but to provide tools to support humans engaged in it. ACE constrains the authors by the processes of selection, mutation and crossover to produce Oulipian text. In this way the human user is firmly rooted within dual processes: the evolutionary process and the gestation of the text being produced. It has been observed that the use of natural language as a genotype representation is possible because the pattern of language and its constituent parts follow a clearly defined syntactic and semantic structure making the evolutionary method of natural language processing within an IEC system efficient [7], [8], [5], 9].

Applying ACE to the production of literature is an Oulipian process. The automation of ACE ensures that the creative process is structured and it is easy to include procedures, which help to avoid one user of the system having too much influence over the final solution. For ACE to be extended to the evolution of more complex language structures and larger group sizes of users, different methods of controlling population explosion and interface design have to be considered if the system is to be effective in avoiding human and IEC fatigue. The experimental output and the different evolutionary representations of it offer an opportunity for the authors to truly explore their creative potential and the reader to choose their own user experience as they project themselves as author in a real Oulipian sense.

## 9. ACKNOWLEDGMENTS

Jennifer Willies, Doug Willies, Chris Osbourne, Mij Kelley, Luis Hercog and Rob Stevens participated in the first ACE experiment which produced the name of the algorithm.

## 10. REFERENCES

[1] Braffort P. and Joncquel J. ALAMO *(Atelier de Littérature Assistée par la Mathématique et les Ordinateurs)*: A Twelve Years Experiment in Literature and Data Processing published by Vuillemin A. and Lenoble M., Artois Université Presses, p.171. 1995 http://indy.culture.fr/docs/alamo/publications/index.html

[2] Bringsjord S. Chess Is Too Easy *Technology Review* 23 1998.

[3] Calvino, I. If On a Winter's Night a Traveller. Trans. W. Weaver. London: Minerva, 1992.

[4] Dawkins R. *The Blind Watchmaker* Essex: Longman 1986

[5] Defaweux A. Grosche T. Karapatsiou M. Moraglio A. Shenfield A. Automated Concept Evolution EvoNet SummerSchool 2003. http://evonet.lri.fr/summerschool2003/resources/solutions/concept_evolution/Paper.ps

[6] Fogarty T. Automatic Concept Evolution *The Second IEEE International Conference on Cognitive Informatics* (ICCI'03) p. 89 August 18 - 20, London, England 2003.

[7] Kosorukoff A. *Human based genetic algorithm*, Illinois, IlliGAL report no. 2001004. 2000.

[8] Kosorukoff A. and Goldberg D. Evolutionary Computation as a Form of Organization, *Proceeding Genetic and Evolutionary Computation Conference* (GECCO), Morgan Kaufmann, pp 965-972, 2002.

[9] Llora X. Kei Ohnishi, Chen Y.P, Goldberg D. and Welge M. Enhanced Innovation: A Fusion of Chance Discovery and Evolutionary Computation to Foster Creative Processes and Decision Making, Illinois, IlliGAL Report 005, 2001.

[10] Mathews H. and Brotchie *The Oulipo Compendium* London: Atlas Press, 1998.

[11] Parmee I. *Interactive Evolutionary Computation* GECCO 2001 and GECCO 2003 Workshops http://www.ad-comtech.co.uk/Workshops.htm

[12] Perec G. *Life, A User's Manual* Boston: D.R. Godine, 1987.

[13] Perec G. *La Disparition*, 1969 (A Void) trans. Gilbert Adair London: Harvill Press 1995.

[14] Quart A. Oulipo Correspondence, *An International Review of Culture and Society* Council on Foreign Relations, USA Issue No. 5, Winter 1999/2000 www.cfr.org/pdf/correspondence/CORR.fall99.pdf

[15] Queneau R. *Cent Mille Milliard de Pèomes*. (Hundred thousand billion poems) Gallimard 1961.

[16] Queneau R. and Wright B. *Exercises de Style* London: J. Calder: Balder and Boyars, 1979.

[17] Roubaud J. *Le grand incendie de Londres*, 1989 (The Great Fire of London: a story with interpolations and bifurcations) trans. Dominic de Bernardi, Dalkey Archive Press, IL. 1991.

[18] Seaman B. Oulipo vs Recombinant Poetics *LEONARDO*, Vol. 34 No. 5, pp 423-430, 2001.

[19] Sims K. Artificial Evolution for Computer Graphics *Computer Graphic*, 25(4), pp 319-328, 1991.

[20] Soddu C. New Naturality: A Generative Approach to Art and Design *Leonardo*, MIT Press Vol. 35, No. 3, pp 291-294, 2002.

[21] Takagi H. Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and Human Evaluation *Proceedings IEEE* vol.89, no.9, pp 1275-1296, 2001.

[22] Thiler, A. *Raymound Queneau*, Twayne Publishers, Boston MA p.51, 1985.

[23] Wolff M. Oulipo, Machines, and Posthuman Literature Presented at the 2003 Twentieth-Century French Studies Colloquium, Urbana-Champaign 2003 http://users.hartwick.edu/wolffm0/20FSCg.html