# Story Generation Driven by System-Modified Evaluation Validated by Human Judges[*]

Pablo Gervás[1] and Carlos León[2]

[1] Instituto de Tecnología del Conocimiento
`pgervas@sip.ucm.es`
[2] Departamento de Ingeniería del Software e Inteligencia Artificial
`cleon@fdi.ucm.es`

**Abstract.** Building systems which can transform their own generation processes can lead to the creation of novel high quality artefacts. In this paper a solution based on evaluation is proposed. The generation process is driven by evaluation rules which can be modified by the system. A panel of human evaluators provides feedback on the quality of the artifacts resulting after each modification. The system keeps track of which rules have been applied in the selection of each artifact, and learns indirectly from the human judges which modifications to retain based on the relative ratings of the artifacts. Relevant details and difficulties of this approach are discussed.

## 1 Introduction

Societies of human creators are driven by two basic activities: creation of new artifacts (as performed by artists) and evaluation of newly created artifacts (as performed by artists and/or critics). Most of the efforts at modelling human creativity in computational terms in the past have focused on the task of creating artifacts. There are two strong arguments in favour of shifting the focus towards evaluation. First, developing models or algorithms for producing artifacts of a given kind tends to produce good/recognisable/typical artifacts of that type, rather than creative new ones. Innovation requires both departure from established procedures and the means for identifying when new results are good. Second, generate and test approaches constitute a simple computational way of rephrasing the task of creating artifacts in terms of the task of evaluating them. Very simple enumerative procedures for traversing a search space may yield surprisingly good results if driven by an appropriate evaluation function.

If such a shift is taken to an extreme, the enumeration of the valid alternatives would not need to be altered in search for new artifacts, it would be enough to modify the evaluation function to obtain new candidate elements. Under this approach, the task of modifying creative procedures to obtain new artifacts would take the form of modifying the evaluation function.

---

In societies of human creators the development of an evaluation function (usually understood as artistic sensibility or equivalent abilities) is recognised as a fundamental requirement in the learning process of creative individuals. This learning process almost always takes the form of having instances of good artifacts pointed out.

This paper describes a system that outputs new artifacts obtained by exploring a restricted conceptual space under the guidance of a set of evaluation rules. The conceptual space to explore is that of sequences of events that may be understood as stories. The exploration procedure is exhaustive enumeration of the search space. The system starts off from an initial set of evaluation rules for selecting new artifacts as the conceptual space is explored. A method for actively modifying the set of evaluation rules is provided. Modifications of the evaluation rules lead to new artifacts. The system learns which of the modified rules to retain from the responses of a panel of human evaluators that act as audience for its production of new artifacts.

## 2    Previous Work

Boden [1,2] divides creativity into *exploratory creativity* (exploring the common possibilities for creating artefacts) and *transformational creativity* (changing these common rules to find really new and valid objects).

Jennings [3] hypothesizes that societies create the evaluation criteria of creativity in the individual's mind, thus leveraging the concept of creativity to a place beyond pure inner processes. As such, creativity is learned and taught between individuals, and their relationships and the opinions that each one has about another have a strong influence on the ideas about the quality or novelty of artifacts. *Autonomous creativity* is the ability to change one's own standards without explicit direction from the outside. According to Jennings, autonomous creativity in humans is achieved through social interaction

Ritchie [4] identifies the role of humans in Computational Creativity to be still very necessary, given the current state of the art. In his model, it is stated the role of humans must be clearly established before putting them in the generation loop. It is also hypothesized that human actions in the system should never be directly related to the generative objective of the system.

Wiggins [5,6] defines a formalization of Computational Creativity processes in terms of their relation with classic Artificial Intelligence and the characteristics that separate pure exploration processes from those typically and only present in Computational Creativity. In his formalization, several sets are identified: $\mathscr{U}$, the *universe* of concepts, containing the whole set of artefacts; the *conceptual spaces* $\mathscr{C}_0 \cdots \mathscr{C}_n$, which are strict subsets of $\mathscr{U}$, among others. Three functions are also important to mention: $\mathscr{R}$, which establishes the constraints that define the conceptual space of valid results, $\mathscr{T}$, which is the function that transverses this conceptual space and sets an order on the identification of artefacts in the $\mathscr{C}_i$ set constrained by $\mathscr{R}$, and $\mathscr{E}$, the function for evaluating artifacts.

## 3 Story Generation Based on Evaluation

The domain of story generation has been chosen to illustrate the ideas in this paper because it deals with artifacts that are easy to represent symbolically, are linear in nature, and, at a certain level of abstraction, their complete conceptual space may be specified by definition in terms of combinations of their constituent elements. Some of these points are sketched briefly below.

In terms of Wiggins' model, the simplest approach for a generation system that explicitly performs evaluation on the stories it generates could be the definition of the $\mathscr{E}$ function ($\mathscr{E}_g$ at this level) and a basic generative strategy which would generate all possible stories in the conceptual space. The generative strategy, corresponding to Wiggins $\mathscr{T}$ function ($\mathscr{T}_g$ at this level), could be carried out by a simple backtracking generation in which each step adds a new event to the story (which therefore, after several steps in that branch, creates a whole story) and then backtracks to test another generative branch. Given a certain set of terminals like verbs, character names, places and valid time values, for instance, events in the form `subject–verb–arguments` can be easily generated. The stories can be considered to be sequences of events in the form $\{e_1, e_2, \cdots, e_n\}$ (where events would be conceptual statements corresponding to sentences like "Robert went to the park").

The evaluation function ($\mathscr{E}_g$) would output a real value in the interval $[-1, 1]$, $-1$ being a "very bad" story and $1$ being a wonderful one. A value of $0$ would represent a plain, normal story, acceptable but not "good". Thus we could obtain a total order for stories in which any threshold in the $[-1, 1]$ range could be used to differentiate interesting stories (those falling above the threshold) from non-interesting ones. The $\mathscr{E}_g$ function could be composed of rules whose structure could be formed by a set of *preconditions* ($\pi$) considering the current partial evaluation and the current state of the story and a set of *effects* ($\epsilon$) that the application of those rules have on the final evaluation. Then, a very simple evaluator would process the story events iteratively, checking the preconditions and applying the postconditions , in such a way that the state of the evaluator (the partial set of variables that form the evaluation) is progressively updated for each processed event.

### 3.1 Evaluation-Driven Story Generation

The original definition of the $\mathscr{T}$ function given by Wiggins modelled the operation of identifying the next element in the conceptual space to be considered. Under a certain interpretation, this could be understood to refer to the actual construction process followed by the creative system to obtain its next result. In this case, the range of the $\mathscr{T}$ function defines system output. However, under a different interpretation, the $\mathscr{T}$ function would be the procedure for constructing the next element to be considered by the evaluation function $\mathscr{E}$. As some of the candidates proposed by the $\mathscr{T}$ function will be rejected by the evaluation function $\mathscr{E}$, system output in this case is defined by the interaction between the

$\mathscr{T}$ and the $\mathscr{E}$ functions. In this paper we consider this second interpretation. Modifications of the $\mathscr{E}$ function will therefore control system output.

For the purposes of this paper, plain random modification of the rules can be considered. More refined solutions may be considered. However, the system should not rely on the quality of any particular method of transformation. At this new level, we will also shift the responsibility for obtaining acceptable results to the evaluation process, in this case, the evaluation of effects of the modified rules. For this higher level evaluation we resort to a panel of judges.

### 3.2  Social Interaction Between Humans and Computers for Controlling Transformation

The human judges that evaluate stories are asked to produce plain values which are decided when reading stories. For every generated story, a single numeric value in the range $[-1, 1]$ could be received from humans reading a story, as long as the variable to be obtained is clearly defined and it is just dependent on human criteria regarding stories.

The proposed method for evaluating stories involves checking the available set of evaluation rules against each story. Only some of these rules will have their preconditions met, and therefore be applied to contribute to the final rating that the system assigns to the story. For every story $S$ that is finally selected as system output, a record is kept of which evaluation rules contributed to establish its internal rating: the particular subset of the evaluation rules (the $\mathscr{F}_S$ set) that contributed to its being selected as output.

By combining this record with the evaluations obtained from the human judges, each rule in this subset $\mathscr{F}_S$ could be assigned the rating that humans assigned to the story $S$. In this way, rules would receive several ratings coming from humans indirectly. This could be used, for instance, to keep the rules that produce good stories and discard rules creating bad stories according to human evaluation.

## 4  Discussion and Further Development

It is important to consider to what extent the autonomy of such a system would be compromised by the role played by the human judges. Ritchie points out [7] the need to keep humans isolated from the final objective of the system. In the present case, this corresponds to ensuring that the human participants play no direct role in the actual generation of the stories. At a more specific level, since the system is transforming evaluation rules, the human judges must not directly add knowledge concerning transformation of rules. In the described set up, human judges do not at any stage come into contact with the set of evaluation rules or the method used for transforming them. This constitutes a certain safeguard of system autonomy.

Another aspect to take into account is whether the role played by the human judges in the proposed system could be seen to be modelling real phenomena

that occur in human creativity. We believe that it emulates closely the role played by critics and teachers in the formation of the creative capabilities of human creators. Along these lines, improvements to the present proposal could be contemplated. According to Jennings [3], the influence that external individuals have on generators depends on the relation between the generators and the evaluators. Issues like past agreement or mutual admiration may play a significant role in tempering actual feedback. For instance, it might be interesting to consider whether the learning process of the system might be refined by giving priority to the opinions of judges that have awarded good ratings in the past.

The proposed solution would be inefficient. Although the system might explore candidate artifacts at a fast rate, and transform evaluation rules at speed, it relies on a stage of feedback from human judges that would take time (for a number of stories to be read and evaluated by the judges). The system would have to undergo a learning process equivalent to that of human storywriters receiving feedback from knowledgeable mentors.

The current proposal restricts system output to a very specific conceptual space, and all system operations, whatever transformations are applied to the evaluation rules and whatever feedback is received from the judges, cannot lead to outputs beyond that conceptual space. In that sense, it could only aspire to be considered creative in an exploratory manner. Nonetheless, the system is explicitly transforming its own procedures in a search for better valued artifacts. This aspect of creative professions, the continuous search for improvement through modification of the procedures, has yet to be addressed in the computational creativity literature. The present proposal constitutes a first step in this direction.

## References

1. Boden, M.: Creative Mind: Myths and Mechanisms. Routledge, New York, NY, 10001 (2003)
2. Boden, M.: Computational models of creativity. Handbook of Creativity (1999) 351–373
3. Jennings, K.: Developing creativity. Artificial Barriers in Artificial Intelligence. In: Proceedings of International Joint Workshop on Computational Creativity. (2008)
4. Ritchie, G.: Some Empirical Criteria for Attributing Creativity to a Computer Program. Minds & Machines **17** (2007) 67–99
5. Wiggins, G.: Searching for Computational Creativity. New Generation Computing, Computational Paradigms and Computational Intelligence. Special Issue: Computational Creativity **24**(3) (2006) 209–222
6. Wiggins, G.: A preliminary framework for description, analysis and comparison of creative systems. Knowledge-Based Systems **19(7)** (2006)
7. Ritchie, G.: Uninformed Resource Creation for Humour Simluation. In: Proceedings of the 5th International Joint Workshop on Computational Creativity, Madrid (2008) 147–151