A System for Evaluating Novelty in Computer Generated Narratives

Rafael Pérez y Pérez, Otoniel Ortiz, Wulfrano Luna, Santiago Negrete, Vicente Castellanos, Eduardo Peñalosa, Rafael Ávila
División de Ciencias de la Comunicación y Diseño
Universidad Autónoma Metropolitana, Cuajimalpa
Av. Constituyentes 1050, México D. F.
{rperez, oortiz, wluna, snegrete, vcastellanos, epenalosa, ravila}@correo.cua.uam.mx

Abstract

The outputs of computational creativity systems need to be evaluated in order to gain insight into the creative process. Automatic evaluation is a useful technique in this respect because it allows for a large number of tests to be carried out on a system in a uniform and objective way, and produce reports of its behaviour. Furthermore, it provides insights about an essential aspect of the creative process: self-criticism. Novelty, interest and coherence are three main characteristics a creative system must have in order for it to be considered as such. We describe in this paper a system to automatically evaluate novelty in a plot generator for narratives. We discuss its core characteristics and provide some examples.

Introduction

Automatic evaluation is a central topic in computational creativity. Some authors claim that it is impossible to produce computer systems that evaluate their own outputs (Bringsjord and Ferrucci 2000) while others researchers challenge this idea (e.g. Pérez y Pérez & Sharples 2004). Although there have been several discussions and suggestions about how to evaluate the outputs produced by creative computer programs (e.g. Ventura 2008; Colton 2008; Ritchie 2007; Pereira et al. 2005; Pease, Winterstein, and Colton 2001) there is a lack of agreement in the community on how to achieve this goal.

We are currently working in plot-generation and as part of our research project we are interested in developing a computer model that evaluates the stories generated by our automatic storyteller. Pérez y Pérez and Sharples suggest that

> A computer model might be considered as representing a creative process if it generates knowledge that does not explicitly exist in the original knowledge-base of the system and which is relevant to (i.e. is an important element of) the produced output (Pérez y Pérez and Sharples, 2004).

The authors refer to this type of creativity as computerised creativity (c-creativity). They also claim that a computer-based storyteller must generate narratives that are original, interesting and coherent.

Following these authors, in this document we report a system called The Evaluator that automatically evaluates originality aspects of the c-creativity in the narratives produced by our computer model of writing. We assess three characteristics of novelty in the narratives generated by our storyteller: how novel the sequence of actions is; how novel the general structure of the story is; how novel the use of characters and actions in the story is (we refer to this aspect as repetitive patterns; see below for an explanation). In all cases we compare the plot just produced by the storyteller, from now onwards referred to as the new story, against its knowledge-base. Following the definition of c-creativity, a novel narrative must provide the storyteller with new knowledge that can be used in the future for generating original plots. Thus, we also evaluate how many new knowledge structures are created as a result of the plot-generation process. We combine the results of such evaluations to provide an overall assessment. In this document we present our first results. We are aware that human evaluation of narratives is far more complex and involves not just novelty but several other characteristics. Nevertheless, there are few implemented systems for automatic evaluation (e.g. Norton, Heath and Ventura 2010). In the same way, our implemented system innovates by considering different dimensions of novelty (c.f. Peinado et al. 2010).

The paper is organised as follows. Section 2 explains the general characteristics of the knowledge structures employed in our storyteller and how they are used to evaluate novelty. Section 3 describes how our computer model evaluates narratives. Section 4 provides two examples of narrative's evaluation. Section 5 provides the discussion and conclusions of this work.

Knowledge Representation

Our computer-based storyteller employs two files as input to create its knowledge-base: a dictionary of story actions and a set of previous stories. Both files are provided by the user of the system. The dictionary of story-actions includes the names of all actions that can be performed by a character within a narrative together with a list of preconditions and post conditions for each. The Previous Stories are se-

quences of story actions that represent well-formed narratives. With this information the system builds its knowledge base. Such a base is comprised by three knowledge-structures: contextual-structures or atoms; the tensional representation; the concrete representation.

1. Contextual-Structures (also known as atoms). They represent, in terms of emotional links and tensions between characters, potential situations that might happen in the story-world, and have associated a set of possible logical actions to be performed when that situation occurs. For example, an atom might represent the situation or context where a knight meets a hated enemy (the fact that the knight hates the enemy is an example of an emotional link), and it might have associated the action "the knight attacks the enemy" as a possible action to be performed. Contextual-structures represent story-world commonsense knowledge. By an analogy with Case Based Systems, we can think of our storyteller as a Contextual Based System. Thus, Contextual-structures are the core structures that our storyteller employs during plot generation to progress a story.

2. Story-structure Representation. Our plot-generator has as its basis the classical narrative construction: beginning, conflict, development, climax (or conflict resolution) and ending. We represent this structure employing tensions. Emotional tension is a key element in any short story (see Lehnert 1983 for and early work on this subject). In our storyteller it is assumed that a tension in a short story arises when a character is murdered, when the life of a character is at risk, when the health of a character is at risk (e.g. when a character has been wounded), when a character is made a prisoner, and so on. Each tension has associated a value. Thus, each time an action is executed the value of the tension accumulated in the tale is updated; this value is stored in a vector called Tensional Representation. The Tensional Representation records the different values of the tension over time. In this way, the Tensional Representation permits representing graphically the structure of a story in terms of the tension produced in it (see examples below). Each previous story has its own Tensional Representation. The storyteller employs all this information as a guide to develop an adequate story-structure during plot-generation.

3. Concrete-Representation. It is formed by a copy of the dictionary of story-actions and the set of Previous Stories. The system uses this information to break impasses and sometimes to instantiate characters.

In summary, the storyteller uses the following information during plot-generation: a dictionary of story-actions, a set of previous stories (a set of sequences of actions), its corresponding set of story structures (Tensional-representations) and several contextual-structures. We are interested in analysing whether the storyteller is able to produce novel material that increments some of this content with useful information.

As mentioned earlier, the previous stories are written by humans[1]. Previous Stories mirror cultural and social characteristics that end up being encoded within the storyteller knowledge-base. For example, let us imagine that in all previous stories female characters never perform violent actions; or that all previous stories include an important number of violent actions; and so on. We are interested in evaluating if our storyteller is capable of producing stories that somehow move away from some of those recurrent patterns (stereotypes) found in the previous stories.

Thus, the steps to evaluate a narrative are:

1. The storyteller generates a new plot.
2. The Evaluator compares the new plot with all the previous stories. The goal is to see how novel the sequences of actions of the new story are compared to all the previous stories.
3. The Evaluator compares the story structure (the Tensional-Representation) of the new plot with the story structure of all previous stories, to measure how novel it is.
4. The Evaluator verifies if at least one recurrent pattern in the new story is novel compared to those employed in all the previous stories.
5. The new plot is added to the Previous Stories. The Evaluator compares the knowledge-base of the storyteller before and after this operation is performed. The purpose of this is to estimate how many new contextual-structures are added to the knowledge-base as a result of adding the new plot to the set of previous stories.

Description of The Evaluator

The Evaluator is comprised of four modules: 1) Evaluation of Sequences, 2) Evaluation of Story-Structure, 3) Evaluation of repetitive patterns, 4) Evaluation of New Contextual-Structures.

1. Evaluation of sequences. This module analyses how novel is the sequence of actions that encompasses the new story. To analyze its novelty, the new story is split into pairs of actions. For example, let us imagine that the new story 1 is comprised of the following sequence of actions: Action 1, Action 2, Action 3, Action 4, and so on (each action includes the characters that participate in it and the action itself). Thus, the system creates the following pairs: [Action 1, Action 2], [Action 2, Action 3], [Action 3, Action 4], and so on. The program takes each pair and tries to find one alike in the Previous Stories. The system also has the option of searching for what we have called a distance pair. Let us imagine that the first pair of actions in the new story is: [Enemy kidnapped Princess, Jaguar Knight Rescued Princess]. And that in the Previous Stories we have the following sequence: Enemy kidnapped Princess, Princess insulted Enemy, Jaguar Knight Rescued Princess. As we can observe, although in the Previous Stories the insult-

[1] Currently we are testing an Internet application that will allow people around the world to contribute with their own previous stories to feed our plot-generator system

ing action is located between the kidnapped and rescued actions, the essence of this pair of actions is kept (the antagonist kidnaps the princess and then the hero rescues her). In order to detect this kind of situations, The Evaluator is able to find pairs of actions in the previous stories that have one, two, or more in-between actions. We refer to the number of in-between actions that separate a pair of actions as Separation-Distance (SD). That is, in the previous stories there might be a separation distance between the first and the second action that form the pair. For the previous example, the separation distance value is 1.

2. Evaluation of the story-structure. The structures of the new story and the previous stories are represented as graphics of tensions. The Evaluator compares the structure of the novel story against all the previous stories to see how novel it is. The process works as follows. The Evaluator compares point by point (action by action) the difference between the Tensional-representation of the new story and the first of the previous stories. The highest peak in the graphic represents the climax of the story. Because stories might have different lengths, the system shifts horizontally the graphics in such a way that the climaxes of both stories coincide in the same position on the horizontal axis. If the lengths of the new story and the previous story are different, the system eliminates the extra actions of the longest history. In this way both stories have the same length. The process reports how many points are equal (have the same value of tension) and how many points are dissimilar. The system includes a modifiable parameter, known as Tolerance, which defines when two points are considered as equals. Thus, point-A is considered equal to point-B if point-A = point-B ± Tolerance. By default, the tolerance is set to ±20. Then, the system calculates the ratio between the number of dissimilar points and the total number of actions in the story. This number is known as the Story-Structure Novelty (STN). The same process is repeated for all previous stories. Finally, The Evaluator calculates the average of all Story-Structure Novelty values to obtain a final result.

3. Evaluation of repetitive patterns. The Evaluator analyses the previous stories and the new story to obtain information about recurrent patterns. The current version of the system searches for patterns related to: 1) the most regular types of actions within a story; 2) the reincorporation of characters. Regarding the most regular types of actions, we have grouped all items in the dictionary of story-actions in four different categories: helpful actions (e.g. A cured B, A rescued B); harmful actions (e.g. A wounded B, A killed B); passionate actions (e.g. A loves B, A hates B); and change of location actions (e.g. A went to the City). The system calculates what percentage of actions in each story belongs to each category; the highest percentage is employed as reference for comparison. Then, The Evaluator compares the new story against all previous stories to calculate how similar they are. If more than 50% of the previous stories share the same classification, the new story is evaluated as standard; if 25% to 49% of the previous stories share the same classification, the new story is evaluated as innovative; if less of 25% of the previous stories share the same classification, the new story is evaluated as novel. All percentages can be modified by the user of the system. This is our first approach to automatically identify the theme of a story.

Regarding the reincorporation of characters, we are interested in analysing if one or more characters are reincorporated in a story. This is a resource that Johnstone (1999) employs in improvisation and that helps to develop more complex plots (a set of characters are introduced at some point in the story; then, one or more of them are excluded from the plot; later on they reappear without the narrative losing coherence). This is our first approach to measure the complexity of a narrative in terms of the number of reincorporated characters and the number of actions that takes to reincorporated such characters. We refer to this number of actions as the Distance of Reincorporation (DR). So, if a character is introduced in the story, and she reappears again after 5 actions have been performed, the DR is equal to 5. We consider that characters with higher DR are more difficult to reincorporate without losing coherence in the story than those with lower values. In the same way, we consider that the more characters that are reintroduced in a story without losing coherence the more complex the story is. Thus, we want to study how novel the use of reincorporated characters in the new story is. The Evaluator calculates three values: novelty in the use of reincorporated characters, novelty in the number of reincorporated characters and Novelty of DR. The use of reincorporated characters is calculated employing table 1. The first column indicates the percentage of previous stories that reincorporates characters, the second column indicates if the new story reincorporates characters and the third column shows the evaluation assigned to the new story.

Reincorporation of characters in the Previous Stories	Reincorporation of characters in the new story	Evaluation
Less than 25%	No	Standard
Less than 25%	yes	Novel
25%-50%	no	Standard
25%-50%	yes	innovative
More than 50%	no	Below standard
More than 50%	yes	Standard

Table 1 Novelty in the use of reincorporated characters.

Then the system obtains the number of reincorporated characters in the new story and calculates the percentage of previous stories that have the same or higher number of reincorporated characters. We refer to such a percentage as reincorporated percentage. So, the value of the novelty in the number of reincorporated characters = 100 –percentage of reincorporated characters.

The system calculates the percentage of previous stories whose highest value of DR is equal to or higher than the highest DR in the new story. We refer to such a percentage as percentage of DR. So, the Novelty of DR = 100 - percentage of DR.

4. Evaluation of Novel Contextual-Structures. To perform this process the system requires two knowledge bases: KB1 and KB2. KB1 contains the knowledge structures created from the original file of Previous Stories; KB2 contains the knowledge structures created after the new story is incorporated as part of the Previous Stories. Then, The Evaluator compares both knowledge bases to calculate how many new contextual-structures were included in KB2. The system copies the set of new structures into a knowledge base called KB3. That is, KB3 = KB2 – KB1. Then, The Evaluator performs what we refer to as the approximated comparison. Its purpose is to identify and eliminate those structures in KB3 that are alike, in a given percentage (set by the user) to at least one structure in KB1.In this way, KB3 ends up having only new contextual-structures that are not similar (up to a given percentage) to any structure in KB1. We refer to the final number of structures in KB3 (after performing the approximated comparison) as the KB3-value. The Novelty of the Contextual-Structures (NCS) is defined as the relation between the KB3-value and the number of new contextual-structures.

$$NCS = \frac{\text{KB3-value}}{\text{Number of new contextual-structures}}$$

In this way, we can know how many new contextual-structures are created, and how novel they are with respect to the structures in the original knowledge base KB1.

Examples of Evaluation.

We tested our system evaluating two stories: new story 1 and new story 2. In both cases we employed the same set of Previous Stories comprised of seven narratives.

Example 1. The new story 1 is the outcome of two storytellers working together as a team (see Pérez y Pérez et al. 2010). For this evaluation we employ the knowledge base of one of the agents.

New story 1.
jaguar knight is introduced in the story
princess is introduced in the story
hunter is introduced in the story
hunter tried to hug and kiss jaguar knight
jaguar knight decided to exile hunter
hunter went back to Texcoco Lake
hunter wounded jaguar knight
princess cured jaguar knight
enemy kidnapped princess
enemy got intensely jealous of princess
enemy attacked princess
jaguar knight looked for and found enemy
jaguar knight had an accident
enemy decided to sacrifice jaguar knight
hunter found by accident jaguar knight
hunter killed jaguar knight
hunter committed suicide

1. Evaluation of sequences. We compared the new story 1 against all seven previous stories. We ran the process with values for the separation distance ranging from zero to four. In all cases, we did not find any pair of actions repeated in the previous stories. This is part of the report generated by The Evaluator:

Report
Separation-distance = 4
Total of Pairs Found in the File of Previous Stories: 0%
Novelty of the Sequences of Actions: 100%

2. Evaluation of Story-Structure Novelty (STN). The system generated the following report:

Tolerance = 20
Story[1] Coincidences: 6 Differences: 7 STN : 54%
Story[2] Coincidences: 3 Differences: 7 STN : 70%
Story[3] Coincidences: 2 Differences: 9 STN : 82%
Story[4] Coincidences: 0 Differences: 6 STN:100%
Story[5] Coincidences: 2 Differences: 9 STN: 82%
Story[6] Coincidences: 2 Differences: 8 STN: 80%
Story[7] Coincidences: 1 Differences: 8 STN: 89%
Average Story-Structure Novelty: 79%

The structure of the previous story 1 was the most similar to the structure of the new story 1. Therefore, it has the lowest value of the STN = 54%. On the other hand, the structure of the previous story 4 was the most different to the structure of the new story 1. Therefore, it had the highest STN = 100%.

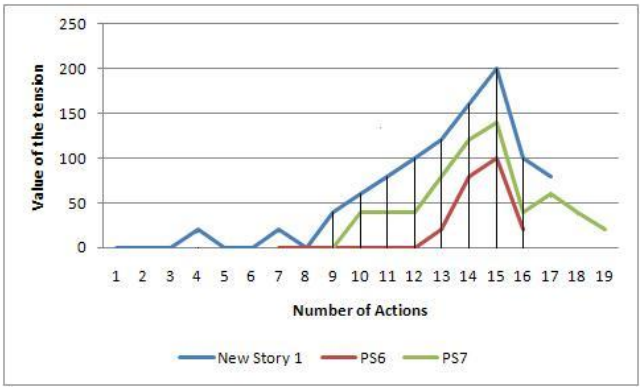

Figure 1. Three story-structures.

Figure 1 shows a comparison of the structures of the previous story 6 (PS6), the previous story 7 (PS7) and the new story 1. The comparison only takes place between actions 9 and 16. The three graphics have been accommodated in such a way that their climaxes are located on action 15. The Evaluator calculated that the average value for the STN was 79%.

3. Evaluation of patterns. Table 2 shows the most regular types of actions employed in each story. For example, 54.55% of actions in the previous story one (PS1) belonged to the classification harmful; 50.00% of actions in the previous story two (PS2) belonged to the classification change of location; and so on. The most regular type of actions employed in the new story 1 (NS1) belonged to the classification harmful. That is, this was a violent story. Four of the seven previous stories shared the same classification and shared similar values of percentage. Therefore, the novelty of the used actions in the new story 1 was classified as standard.

Table 3 shows those characters that were reintroduced at least once in any story and their corresponding distance of reincorporation.

Class of Action	PS1	PS2	PS3	PS4	PS5
Change of location	9.09%	**50.00%**	26.67%	**50.00%**	**44.44%**
Passionate Actions	36.36%	12.50%	26.67%	10.00%	22.22%
Harmful Actions	**54.55%**	25.00%	**40.00%**	20.00%	22.22%
Helpful Actions	0.00%	12.50%	6.67%	20.00%	11.11%

Class of Action	PS6	PS7	NS1	NS2
Change of location	**37.50%**	40.00%	28.57%	14.29%
Passionate Actions	25.00%	0.00%	7.15%	14.29%
Harmful Actions	**37.50%**	**60.00%**	**57.14%**	**71.42%**
Helpful Actions	0.00%	0.00%	7.14%	0.00%

Table 2. Most regular types of actions for each story.

Character	s1	s2	s3	s4	s5	s6	s7	Sto1	Sto2
Eagle Knight	-	-	4	-	-	-	-	-	-
Hunter	-	-	-	-	-	-	-	8	-
Jaguar Knight	-	-	-	-	-	-	-	4	-
Lady	-	-	11	-	5	-	-	-	-
Prince	-	-	-	5	-	-	-	-	-
Princess	-	-	6	-	-	-	7	-	-

Table 3. Reincorporated characters and their DR.

In four of the seven previous stories was possible to find reincorporated characters. However, only one of those stories reincorporated more than one character. The new story 1 reincorporated two characters. Furthermore, this story had the second longest distance of reincorporation. Thus, the novelty in the number of reincorporated characters was set to 86% and the novelty of the DC was set to 86%.
4. Evaluation of novel contextual structures. After comparing KB1 and KB2 the system found ten new contextual-structures. For the purpose of comparison, we ran the approximated-comparison process with 19 different percentage values. For reasons of space we only show eight. This is part of the report produced by The Evaluator:

100% SIMILAR: [0]/[10] 0.00%
75% SIMILAR: [1]/[10] 10.00%
60% SIMILAR: [5]/[10] 50.00%
35% SIMILAR: [7]/[10] 70.00%
25% SIMILAR: [9]/[10] 90.00%
20% SIMILAR: [10]/[10] 100.00%
15% SIMILAR: [10]/[10] 100.00%
10% SIMILAR: [10]/[10] 100.00%

There are no structures 100% equal. If we set the system to find structures that are 75% alike, only one of the ten new contextual-structures is equal or equivalent to at least one structure in KB1. Only when we set the percentage of similarity to 60% half of the new contextual-structures are equal or equivalent to at least one structure in KB1. For this exercise we decided to calculate the value of the Novelty Contextual-structure with the percentage of similarity set to 75%. Thus, NCS = 9/10 = 0.90
In summary, for the new story 1 we got the following values:
Novelty of the Sequences of Actions: 100%
Average Story-Structure Novelty: 79%
Patterns:
Novelty in the use of regular type of actions: Standard
Novelty in the use of reincorporated characters: Standard
Novelty in the number of reincorporated characters: 86%
Novelty of DR: 86%
Novelty of the Contextual-structures: 90%

Example 2. This story was produced by one storyteller.
New story 2.
Jaguar knight is introduced in the story
Enemy is introduced in the story
Enemy got jealous of jaguar knight
Enemy attacked jaguar knight
Jaguar knight fought enemy
Enemy killed jaguar knight
Enemy laugh at enemy
Enemy exiled enemy
Enemy had an accident

1. Evaluation of sequences. As in the case of story 1, we ran the process with values for the Separation-distance ranging from zero to four. In all cases, we did not find any pair of actions repeated in the previous stories. Thus, the novelty of the sequences of Actions is 100%. The report is omitted for space reasons.
2. Evaluation of the story structure. As in the case of the new story 1, we got an average value of 65% of novelty in the story structure. The report is omitted for space reasons.
3. Evaluation of patterns. Table 1 shows that the most regular types of actions employed in the new story 2 (NS2) belonged to the classification harmful. Four of the seven previous stories shared the same classification although, by contrast with all previous stories, the new story 2 had the highest percentage of harmful actions used. Nevertheless, the new story 2 was classified as standard. Table 2 shows that the new story 2 did not reincorporate characters. Therefore, it was evaluated as below standard. As a consequence, the novelty in the number of reincorporated characters and the novelty of the DC were set to 0%.
4. Evaluation of novel contextual structures. After comparing KB1 and KB2 the system found seven new contextual-structures. For the purpose of comparison, we ran the approximated-comparison process with 19 different percentage values. For reasons of space we omit the report. There are no structures 100% equal. If we set the system to find structures that are 55% alike, only one of the seven new contextual-structures is equal or equivalent to at least one structure in KB1. Only when we set the percentage of similarity to 25% more than half of the new contextual-structures are equal or equivalent to at least one structure in KB1. For this exercise we decided to calculate the value of the Novelty Contextual-structure with the percentage of similarity set to 75%. Thus, NCS = 7/7 = 1 for 75%

Thus, for the new story 2 we got the following values:
Novelty of the Sequences of Actions: 100%
Average Story-Structure Novelty: 65%
Patterns:
Novelty in the use of regular type of actions: Standard
Novelty in the use of reincorporated characters: Below Standard
Novelty in the number of reincorporated characters: 0%
Novelty of DR: 0%
Novelty of the Contextual-structures: 100%

Discussion

This paper reports on the implementation of a computer system to automatically evaluate the novelty aspect of c-creativity. Following Pérez y Pérez and Sharples (2004), c-creativity has to do with the generation of material that is novel with respect to the agent's knowledge base and that, as a consequence, generates new knowledge-structures. These authors distinguish two different types of knowledge: knowledge about the story-structure and knowledge about the content (the sequence of actions). In this work we also consider commonsense or contextual knowledge and what we refer to as patterns knowledge.

The sequences of actions in the new stories 1 and 2 are unique with respect to the sequences of actions found in the previous stories. Thus, the storyteller is capable of producing novel sequences of actions. The evaluation of the structures' novelty of both new stories got a value of 65%. That is, the system is able to diverge from the structures found in the previous stories. The results of our tests also show that new contextual knowledge structures, the core information employed during plot generation, are built as a result of adding the new story to the file of previous stories. Thus, The Evaluator shows that our storyteller is able to generate novel knowledge structures in at least three aspects. The results obtained from the analyses of recurrent patterns are not conclusive. We need to make more tests to assess if our system can contribute to the measure of some aspects related to the complexity of a story; something similar happens with the automatic detection of the theme of a story. Nevertheless, the statistical information that The Evaluator generates shows that the storyteller is able to generate narratives that display certain degree of pattern originality.

Automatic evaluation is a key component of the overall assessment of a creative system because it provides unbiased information on the system's behaviour. This feedback also supplies insights that allow improving different aspects of our computer model of creative writing. The system provides an inkling into how novel stories are that help us adjust the various parameters of the system to carry out new experiments. In this way, The Evaluator speeds up the experimentation cycle. Finally, we are also interested in comparing the results that The Evaluator generates against human evaluation. Specific creative patterns could be sought, similar to repetition-break (Loewenstein and Heath 2009), to carry out a more specialized evaluation of the knowledge bases.

References

Bringsjord, S.; Ferrucci, D.A. 2000. *Artificial Intelligence and Literary Creativity. Inside the Mind of BRUTUS, a Storytelling Machine*. Erlbaum (Lawrence), Hillsdale.

Colton, S. 2008. Creativity versus the perception of creativity in computational systems. *Creative Intelligent Systems: Papers from the AAAI Spring Symposium*. 14–20.

Johnstone, K. 1999. *Impro for Stoytellers*. Routledge.

Lehnert, W. 1983. Narrative Complexity Based on Summarization Algorithms. *Proceedings of the Eighth international joint conference on Artificial intelligence* (IJCAI'83) Vol. 2 Morgan Kaufmann Publishers Inc. San Francisco, CA, pp. 713-716.

Loewenstein, J., & Heath, C. 2009. The Repetition-Break plot structure: A cognitive influence on selection in the marketplace of ideas. *Cognitive Science*, 33, 1-19.

Norton, D.; Heath, D.; and Ventura, D. 2010. Establishing Appreciation in a Creative System. *Proceedings of the International Conference on Computational Creativity*. 26-35.

Pease, A.; Winterstein, D.; and Colton, S. 2001. Evaluating machine creativity. In Weber, R. and von Wangenheim, C. G., eds., *Case-based reasoning: Papers from the workshop programme at ICCBR 01*Vancouver. Canada 129–137.

Peinado, F.; Francisco, V.; Hervás R. and Gervás, P. 2010. Assessing the Novelty of Computer-Generated Narratives Using Empirical Metrics. *Mind and Machines*. 20(4):565-588.

Pereira, F. C.; Mendes, M.; Gervás, P., and Cardoso, A. 2005. Experiments with assessment of creative systems: An application of Ritchie's criteria. In Gervaás, P. Veale, T. and Pease, A., eds., *Proceedings of the workshop on computational creativity, 19th international joint conference on artificial intelligence*.

Perez y Perez, R., Negrete, S., Peñaloza, E., Castellanos, V., Ávila, R. and Lemaitre, C. 2010. MEXICA-Impro: A Computational Model for Narrative Improvisation. In *Proceedings of the international conference on computational creativity*, Lisbon, Portugal, pp. 90-99.

Pérez y Pérez, R. and Sharples, M. 2004. Three Computer-Based Models of Storytelling: BRUTUS, MINSTREL and MEXICA. *Knowledge Based Systems Journal*. 17(1):15-2.

Ritchie, G. 2007. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines* 17:76–99.

Ventura, D. 2008. A Reductio Ad Absurdum Experiment in Sufficiency for Evaluating (Computational) Creative Systems. *Proceedings of the International Joint Workshop on Computational Creativity*. 11-19.