Analysis of the correlations between the knowledge structures of an automatic storyteller and its literary production

Iván Guerrero Román[1], **Rafael Pérez y Pérez**[2]
[1]Posgrado en ciencia e ingeniería de la computación. Universidad Nacional Autónoma de México, D.F., México
[1,2]División de ciencias de la comunicación y diseño. Universidad Autónoma Metropolitana, Cuajimalpa, D.F., México
[1]cguerreror@uxmcc2.iimas.unam.mx, [2]rperez@correo.cua.uam.mx

Abstract

In this paper, we describe a process to identify relations among the features of the knowledge base of an automatic storyteller and the narratives that it generates. We define structures to analyze the internal composition of the information available for an agent. We also establish a set of metrics to identify diverse story characteristics. Next, we perform experiments utilizing Mexica, an automatic storyteller, to generate narratives and to evaluate them according to our set of metrics. Then, we compare such assessments with the visual structures that we built from the agent's original knowledge base, in order to obtain correlations between them. The results suggest that such correlations are useful to study the links between the agents' knowledge base and the kind of stories they might produce.

Introduction

During the last 15 years, members of our research group have developed a wide variety of models related to storytelling, and we have implemented them in computational programs, or agents. Among these models there is an automatic storyteller, Mexica, and a collaborative story generator, Mexica-Impro; models for evaluating stories and for identifying social norms in the generated outputs. In all of them, the knowledge structures (KS) available in each of the agents have played an essential role. We have utilized emotional and tension links between the characters in a story to represent these KSs, and we have obtained this information from two major sources: a dictionary of story-actions, and a set of previous stories (narratives written by humans that are considered benchmarks four our models). Nevertheless, one pending task, tackled in this work, is the study of how features of the agents' knowledge base influence the narratives that they generate. The direct antecedents of this research arise from a three-fold base: automatic story generation and evaluation, and description of high-level structures emerging from the knowledge bases of our agents.

From the first text generation works in the early 60's (Klein 1965), to the latest storytellers such as Fabulist (Riedl 2004), Mexica (Pérez y Pérez 2001 and 2007) or Minstrel (Turner 1994), automatic narrative generation has intrigued researchers for decades in an attempt to better understand diverse aspects of this process. Despite the fact that they have descriptions of how internally represent their knowledge, it is commonly missing how these structures affect the overall quality of the generated stories. Moreover, they lack of high-level representations of the available knowledge to identify emergent structures, and to analyze how these structures prevent unpleasant behaviors and promote desirable features in their outputs.

Regarding to the evaluation of the generated stories, Pérez y Pérez (2014) proposed a layered model describing how features such as opening, climax, closure... in a story, could be measured to determine how coherent, novel and interesting they are. In our work, we rely on these metrics and extend them to identify additional story features and structural elements of the agent's knowledge bases.

To identify high-level structures of the agent's knowledge, Pérez y Pérez (2015) describes contextual structures maps. They represent how the acquired wisdom of an agent is distributed throughout the space of all the possible structures, and identifies different types of elements according to their number of components. In this work, we build upon this idea to present alternative high-level structures to represent relations according to the similarity among the elements inside the KS of our storyteller.

We claim that if we are able to link previous stories with the agents' KSs, and find out how KSs' features influence the characteristics of the generated plots, we will improve our understanding about the importance of previous experiences for the plot generation process. In this way, our agents will be able to identify for example what type of knowledge is still missing in their repositories, and develop stories to explore specific topics with the purpose of filling these gaps.

In general, we review how Mexica, an automatic storyteller, builds its own knowledge structures, and we present a high-level structure which provides us additional information about the knowledge of the computer agent. Then, we present a set of metrics to describe features of stories generated by an agent implementing Mexica, and we also present features to describe the structure of its knowledge base. Finally, we identify relations between these two different types of features.

In this paper, we describe a methodology to visualize characteristics of the agents' KSs, referred to as connectivity maps (C-maps), to show the similarities among KSs in memory. Then, we illustrate how the topology of such maps affects several features of the computer generated plots.

Next, we present our findings about how these knowledge structures affect diverse features of stories generated by our agents. Finally, we reflect on these results and speculate on possible extensions to this work.

Gathering information

We rely on two main component types to identify the relations that we are looking for: a story generator and story evaluator, a computer program to assess the outputs of the generator. For the first, we utilized Mexica (Pérez y Pérez 1999, 2007, and Pérez y Pérez & Sharples 2001), and we extended it by incorporating Social Mexica (Guerrero and Pérez y Pérez 2014), a computer model for social norms in narratives to provide additional social information to the story generation process. For the second component, we extended a model for evaluating the interestingness of a narrative proposed by Pérez y Pérez (2014). We now describe each of these components.

Generating knowledge structures

Mexica is a storyteller based on the E-R creativity model (Pérez y Pérez 1999), which describes the creative activity of writing as an iterative two-phased process: engagement and reflection. During engagement the agent selects diverse actions to produce a partial story; whereas in reflection, the system evaluates and updates the material previously generated. Additionally, diverse guidelines to constrain the production of material during engagement are set according to the evaluations performed during this stage. This evaluations also serve to determine when a story is considered to be finished. If this is not the case, the system initiates a new engagement stage and the cycle starts all over again until the story is considered to be finished.

Mexica employs two information sources to generate a variety of knowledge structures utilized during the story generation process: a dictionary of story-actions, and a set of previous stories. Actions in the dictionary have associated a name, and a set of preconditions and post-conditions to represent their requirements and consequences when added to a story. These conditions are defined in terms of emotions (such as love or friendship) and tensions (such as life or health at risk, character prisoner...). Every story (either generated or previous) is defined as a sequence of instantiated actions. This occurs when characters (a performer and an optional receiver) are added.

'Virgin fell in love with Warrior', represents a valid instantiated action. Here, Virgin and Warrior represent the characters, and 'fell in love with' corresponds to the action phrase. Some of these actions consist of only one character, like 'Hunter went to the forest'.

We use contextual structures (CS) to represent the knowledge available for our agent. They are built from the previous stories to be further utilized during the generation of new narratives. Mexica internally transforms a story into emotional relations and tensions between characters, and from this representation, called story-context, CSs are extracted. They consist of two elements: a set of relations (emotions or tensions) between characters, and a list of desirable continuations. Figure 1 represents a story-context obtained from the following story: 'Tlatoani (T) was father of the Princess (Ps)', then 'the priest (Pt) made Princess her prisoner'. The link from Tlatoani to Princess, represent a positive friendship relation with high intensity (+3); the link from Princess to Priest, represent a negative friendship relation (representing hate) with high intensity (-3); and the seesaw link from Priest to Princess, represent a tension between them ('Pr' represents the type of tension, prisoner).

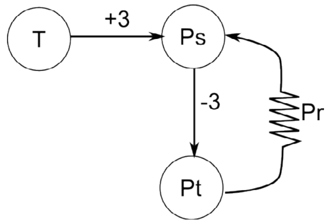

Figure 1: Visual representation of a story-context. Here, nodes represent characters and edges represent relations between characters. The lines with arrow heads represent emotions, whereas the seesaw lines represent tensions.

From the previously described context, Mexica extracts the context of the CS displayed in figure 2. Here, characters are replaced with variables (represented by the letters A, B and C), and the next action in the story is linked to represent a desirable continuation (in our example, the story continued with the action 'T rescued Ps'). A CS can have several actions linked to it. This occurs when identical story-contexts are obtained from different stories, and instead of generating two CS with the same context, we group them into one single CS with multiple actions.

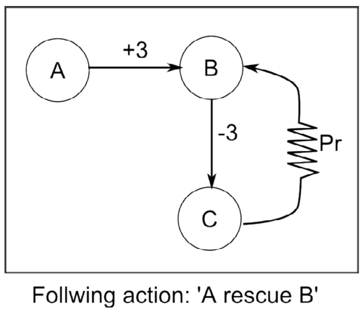

Figure 2: Visual representation of a contextual structure. The rectangle at the top represents a CS-context, and at the bottom is displayed a desirable action for this context.

To identify features related to these knowledge structures, we developed a map to obtain additional information regarding to their similarity. A connectivity map (C-map) represents CSs and relations among them. Every node in this map represents a CS, and two nodes are linked if they are similar enough. The agent determines such similarity by identifying the number of corresponding relations between two structures according to the following rules:

- One emotion is similar to another when they share the same type, valence (positive or negative), and the first has an intensity lower or equal to the second.

- Tensions: Two tensions are similar when they share the same type.
- Once a similar emotion or tension is identified, the character variables of the nodes utilized in the relations are mapped and they cannot be utilized to identify similar relations creating new mappings.

In figure 3, we display a context similar to the one in figure 2. To determine the similarity of the second context with respect to the first, we look for emotions with the same type, valence, and with an intensity lower or equal; we then look for tensions with the same type. In this case, the emotional link between A' and B' in the first context is similar to the emotional link between B and C in the second context. This generates a mapping of the characters A' with B, and B' to C, preventing the generation of new mappings for the variables A', B', B and C. Next, the tension between A' and B' is similar to the tension between B and C, and preserves the original mappings. The only missing element is the emotional link between A and B in the second context. This results in a similarity value of 0.66 (two out of three similar links).

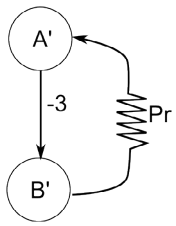

Figure 3: Visual representation of the context in a CS.

From this C-map, nodes are categorized into three different groups according to the number of connections among them. Due to the lack of similar studies, and after analyzing the values obtained, we empirically determined two threshold values of 5% and 10% to create our categories, but we will perform further studies to identify the implications of this values in our study.

- Isolated nodes: Those connected with less than 5% of the total number of nodes
- Regular nodes: Those connected with 5% to 10% of the total number of nodes
- Focal nodes: Those connected with more than 10% of the total number of nodes

When the nodes inside a C-map are linked, they form clusters of similar elements. According to their members, clusters are classified into three categories: islands, towns and cities. After analyzing the number of nodes inside the clusters, we determined two threshold values of 20% and 50% to classify them, but we will develop further studies to determine the implications of this values in our studies.

- Island: Contains less than 20% of the nodes inside the C-map
- Town: Contains between $20-50\%$ of the nodes inside the C-map
- City: Contains more than 50% of the nodes inside the C-map

We present in figures 4 and 5 two samples of C-maps. A gray node represents an isolated node; a red node, a regular; a blue node, a focal. Their size in the picture relies on the number of identical contexts grouped into them. In figure 4, two town-clusters are displayed at the top, and five island-clusters at the bottom of the image. In figure 5, a city-cluster is displayed with an island-cluster at the top.

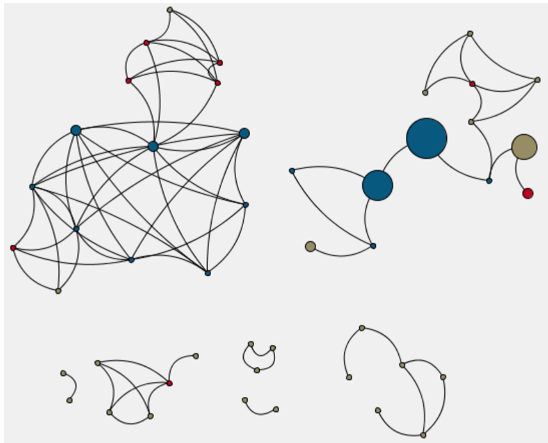

Figure 4: C-map with two town-clusters (top) and five island-clusters (bottom)

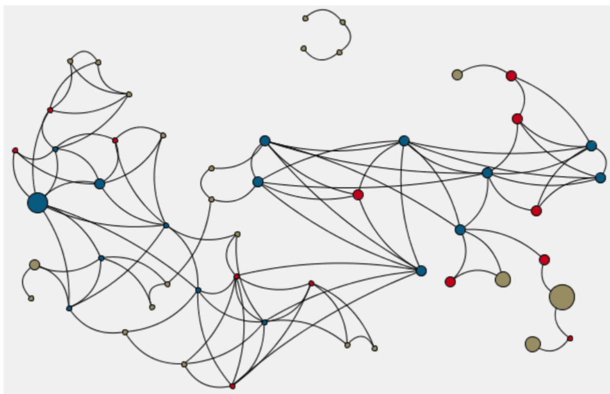

Figure 5: C-map with a city-cluster and an island-cluster

Evaluation process

In grounds of our previous work in this area, we have selected a set of features, known as story-characteristics, for evaluating a plot, and a different set of features, known as structural-characteristics, for evaluating the structures inside the knowledge base of a storyteller.

Evaluating story-characteristics The features utilized to evaluate a story are the following: preconditions fulfilled, novel contextual structures, opening, climax, closure, originality, E-R ratio, number of actions and impasses in a story, character threads and social resolutions. The first eight are part of the feature set described in the evaluation model proposed by Pérez y Pérez (2007 and 2014), and the remaining features are additions to this evaluation model that we considered relevant for this work.

The preconditions fulfilled metric evaluates the number of action requirements satisfied within a story. This value corresponds to a number between 0 and 1 determined by the ratio between the number of preconditions fulfilled for every action versus the total number of preconditions in all the actions.

The novel contextual structures metric determines the amount of new knowledge that a story can generate if it were added to the set of previous stories of an agent. This value is determined by the ratio between the number of new buildable CSs from the story-actions and the total number of CSs that could be generated. We consider a CS new when its context is different from all the existing CSs.

The following metrics are related to the tension curve of a story and to the identification of the three main stages of a story in accordance with the Freytag's pyramid (Freytag 1896). Mexica considers a story to be properly built when it follows this structure. This is why we use it as a reference for these subset of metrics. A story has a correct opening when, at the beginning, there are no tensions and then they begin to grow; it has a correct climax when its highest tension value is similar to a reference value obtained from the set of previous stories; it has a correct closure when all the tensions in the story are solved when the last action is performed.

The originality feature determines the portion of a story that could be generated by the evaluating agent. Mexica is capable of generating a story by itself, from the beginning to a given action, when the following conditions are fulfilled:

1. The story-context associated with the action is similar to the context of one of the available CSs
2. The action is similar to one of the linked actions of the CS with the similar context

The result of this metric is the ratio between the number of actions that could be generated by the evaluating agent, and the total number of actions in a story. If one agent could generate the story on its own, the result is zero; if none of the story contexts are similar to any CS of the agent, the result is one (see figure 6).

$$originality = 1 - \frac{regeneratable\ actions}{total\ number\ of\ actions} \quad (1)$$

Figure 6: Originality

The ER-ratio feature determines the relation between the actions added during the reflection phases versus the actions added during the engagement phases. According to the E-R model, both the engagement and reflective stages should provide a similar number of actions to a story. We claim that a story with engagement actions will be novel, but lack of coherence (since actions requirements are not validated at this stage). On the other hand, a story with reflective actions will be coherent, but lack of novelty (causal constraints are validated during reflection). In general, the result for this metric corresponds to one minus the absolute value of the difference between the engagement ($actions_E$) and reflection ($actions_R$) actions divided by the total ($actions$) number of actions (see figure 7). When the actions added in engagement and reflection are the same, the result is one. When the actions added in engagement or in reflection is zero, the result is zero.

$$ER - ratio = 1 - |\frac{actions_E - actions_R}{actions}| \quad (2)$$

Figure 7: ER-ratio

An impasse occurs when, during engagement, the context of a story is not similar to any context from the available CSs, and the stage finishes. We claim that this behavior occurs when the current story is interpreted as an unknown context for the agent. This feature determines the number of times this situation occurs during the generation of a story.

The character threads feature determines the number of groups (threads) of characters inside the story. We state that two characters belong to a thread when they have a significant relationship inside a story. This condition is fulfilled when two characters participate together in an action that generates or removes a tension between them. For this work, we narrow the number of groups in a story to maintain it simpler and to prevent the existence of parallel stories. The result of this evaluation is a number between 0 and 1 calculated as one divided by the number of character threads inside a story.

The social resolutions feature determines the number of social tensions that remain unsolved by the end of a story. These tensions are added by the Social Mexica component every time a social norm is broken inside a story. We are interested in determining how accurately Mexica finishes these additional tensions within a story in order to fit into the Freytag's pyramidal model. The result of this feature corresponds to the ratio between the number of social tensions solved versus the total number of these tensions in a story. When every social tension was solved, the result is one. When none of the social tensions were solved, the result is zero.

Evaluating structural-characteristics With regards to the knowledge structures, we analyze the C-maps defined to obtain the following set of metrics:

- Percentage of clusters of each type (cities, towns, and islands)
- Percentage of nodes of each type (focal, regular, and isolated)

The percentage of city-clusters describes the ratio between the number of them contrasted against the total number of clusters inside the C-map. Similar calculations are performed to determine the percentage of town-clusters and island-clusters. For the percentage of focal-nodes we count the number of such nodes inside any cluster of the C-map and divide this by the total amount of nodes. We obtain the percentage of regular-nodes and isolated-nodes in a similar way.

Identifying relations

Here, we describe a process to identify relations between the story-characteristics and the structural-characteristics described above.

The relations identified in this work are classified into two categories: cluster ratios and node ratios. In the following paragraphs, we explain the process to identify such relations.

The first step consisted in gathering 40 previous stories and partitioning them into sets. With this, ten stories were located into each set, conforming four story-sets (SS). Then, we split them into two story-banks (SB) with two story-sets each. Next, we recombined the stories on every bank to generate two additional sets, each with 70% of the stories of one story-set and 30% of the stories of the other (see figure 8). We performed the same process in both of the SBs.

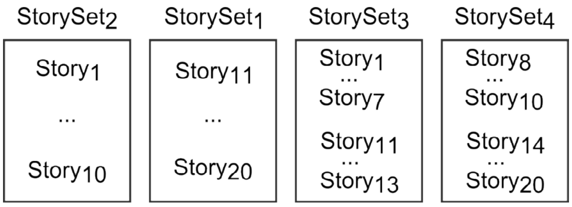

Figure 8: Visual representation of the first SB consisting of four SS. 70% of the stories in SS_3 came from SS_1, and 30% from SS_2. This proportions were inverted to generate SS_4.

After these, we obtained four story-sets on each bank (eight story-sets in total divided into two banks) with the following characteristics:

- Every story-set has totally different stories from one of the story-sets in its story-bank
- Every story-set has 70% similar stories from one of the story-sets in its story-bank
- Every story-set has 30% of stories from one of the story-sets in its story-bank

We utilized each story-set as input for each of the eight different story-generation agents. We let each one of them to generate thirty stories, and we repeated this process three times. By the end of this process, we collected 90 stories per agent. The next step consisted on evaluating every generated story. Each of them was evaluated by every agent in the same bank of the generator, obtaining four evaluations per story. Each evaluation comprised the metrics previously described.

Once these evaluations were completed, we removed those outputs that we did not considered as valid stories according to the following criteria: its preconditions are fulfilled in at least 75%, it has only one character thread, and it contains at least four actions. We collected the evaluations of the remaining stories, obtained the averages for each metric (considering the four evaluations), and we validated if there were differences among them for each of the agents. For this task, we performed an analysis of variance preceded by a K-S test -Kolgomorov-Smirnof test (Massey 1951)- to validate that the data was normally distributed (a request for the variance analysis).

Once we obtained the average values for every metric for every agent, we analyzed the knowledge utilized during the story generation process. The first step consisted in generating the corresponding C-map for every agent to obtain ratios between the different types of nodes and clusters.

We calculated the coefficient of determination (R^2) and the Pearson correlation for every metric utilized during the evaluation process against every metric utilized to describe the knowledge structure. These values leaded us to identify relations between the story-characteristics and the structural-characteristics.

In general, the Pearson-correlation coefficient is a decimal value between -1 and 1. A positive value represents a direct relation between two data sets (when one grows the other does it too), whereas a negative value represents an inverse relation (when one grows, the other decreases), and a value close to 0 represents no linear relation between them. The R^2 value represents how close a data set behaves according to a polynomial of degree n. When $n = 1$, it represents how close is the data to a linear behavior. A value of one for this metric corresponds to a perfect match with a linear behavior, whereas a value of zero represents the absence of a linear correspondence. We now present the relations between every pair of metrics whose values were close to one, which identifies highly related data sets.

Results

Now we present only the results obtained for those relations found between story-characteristics and structural-characteristics with a strong Pearson correlation value (greater than 0.5 or lower than -0.5). The rest of the possible pairings were removed since their Pearson correlation values were not significant. Further studies will determine whether exist additional nonlinear relations among these banned pairings.

In figure 9, we present the novel contextual structures evaluation averages contrasted against the percentage of focal and isolated nodes for each agent. The Pearson correlation values obtained were 0.8 for focal nodes and -0.71 for isolated nodes, and the R^2 values for $n = 1$ were 0.51 and 0.64 respectively. The first values represent a positive linear relation between the novelty of a story and the number of focal nodes inside the story generator, and the second values represent a negative linear relation between the novelty and the number of isolated nodes.

In figure 10, we present the opening averages against the percentage of city and island clusters for each agent. The Pearson correlation values obtained were 0.78 for city clusters and -0.86 for island clusters, and the R^2 values for $n = 1$ were 0.60 and 0.74 respectively. The first values represent a positive linear relation between the opening of a story and the number of city clusters inside the story generator, and the second values represent a negative linear relation

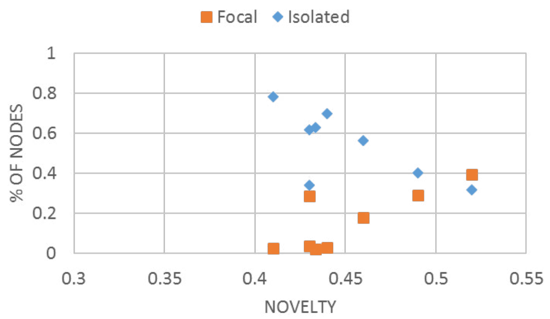

Figure 9: Novel contextual structure averages versus percentage of focal and isolated nodes

between the opening and the number of island clusters.

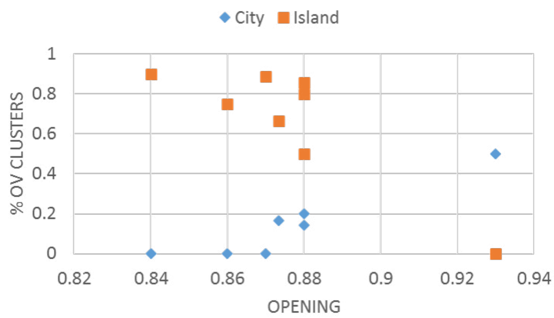

Figure 10: Opening averages versus percentages of city and island clusters

In figure 11, we present the climax averages against the percentage of focal and isolated nodes for each agent. The Pearson correlation values obtained were 0.83 for focal nodes and -0.85 for isolated nodes, and the R^2 values for $n = 1$ were 0.69 and 0.72 respectively. The first values represent a positive linear relation between the climax of a story and the number of focal nodes inside the generator knowledge base, and the second values represent a negative linear relation between the climax and the number of isolated nodes.

In figure 12, we present the closure averages against the percentage of city and island clusters for each agent. The Pearson correlation values obtained were -0.66 for city clusters and 0.67 for island clusters, and the R^2 values for $n = 1$ were 0.44 and 0.45 respectively. The first values represent a negative linear relation between the closure of a story and the number of city clusters inside the story generator, and the second values represent a positive linear relation between the closure and the number of island clusters.

In figure 13, we present the character threads' averages against the percentage of focal and isolated nodes for each agent. The Pearson correlation values obtained were 0.79 for focal nodes and -0.86 for isolated nodes, and the R^2 values for $n = 1$ were 0.62 and 0.74 respectively. The first

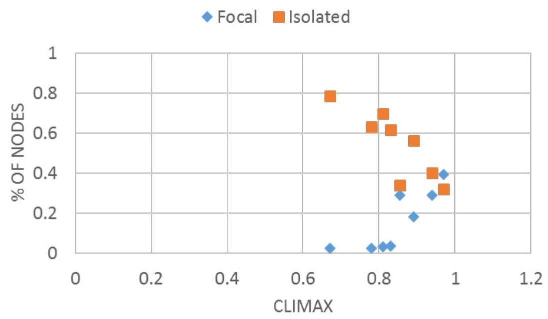

Figure 11: Climax averages versus percentages of focal and isolated nodes

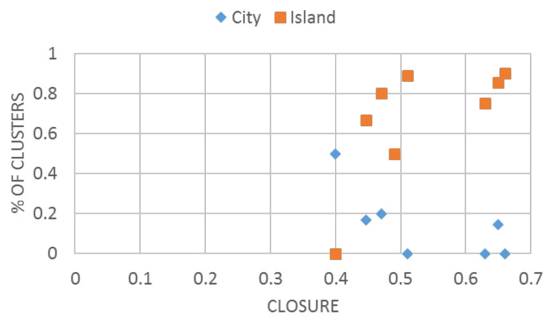

Figure 12: Closure averages versus percentages of city and island clusters

values represent a positive linear relation between the character threads of a story and the number of focal nodes inside the generator knowledge base, and the second values represent a negative linear relation between the character threads and the number of isolated nodes.

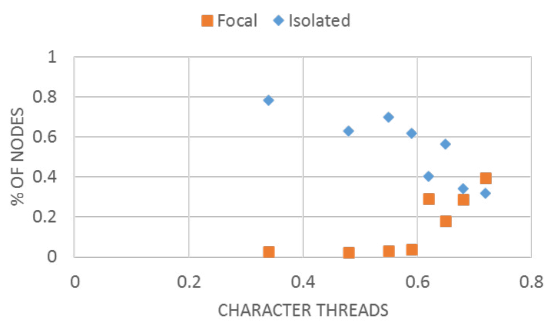

Figure 13: Character threads averages versus percentages of focal and isolated nodes

In figure 14, we present the social resolution averages against the percentage of city and island clusters for each agent. The Pearson correlation values obtained were -0.87 for city-clusters and 0.67 for island-clusters, and the R^2 values for $n = 1$ were 0.75 and 0.45 respectively. The first values represent a negative linear relation between the social resolutions in a story and the number of city-clusters inside the generator knowledge base, and the second values represent a positive linear relation between social resolutions and the number of island-clusters.

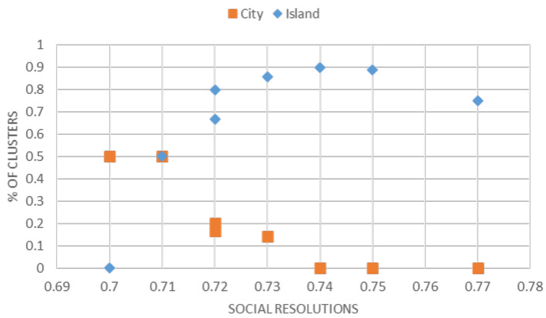

Figure 14: Social resolution averages versus percentages of city and island clusters

In figure 15, we present the originality evaluation averages contrasted against the percentage of town and island clusters for each agent. The Pearson correlation values obtained were -0.75 for town clusters and 0.61 for island clusters, and the R^2 values for $n = 1$ were 0.56 and 0.38 respectively. The first values represent a negative linear relation between the originality of a story and the number of town clusters inside the story generator, and the second values represent a weak positive linear relation (since values are not close to 1) between the originality and the number of island clusters.

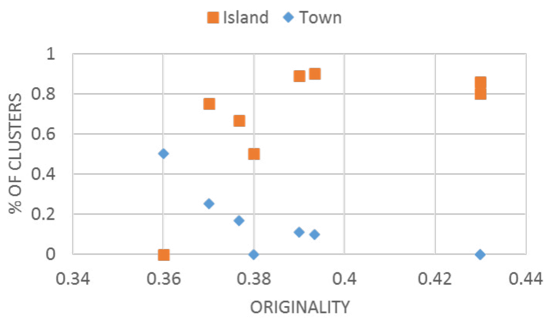

Figure 15: Originality averages versus percentages of town and island clusters

In figure 16, we present the E-R ratio averages against the percentage of focal and isolated nodes for each agent. The Pearson correlation values obtained were 0.87 for focal nodes and -0.86 for isolated nodes, and the R^2 values for $n = 1$ were 0.75 and 0.73 respectively. The first values represent a positive linear relation between the E-R ratio and the number of focal nodes inside the generator knowledge base, and the second values represent a negative linear relation between the E-R ratio and the number of isolated nodes.

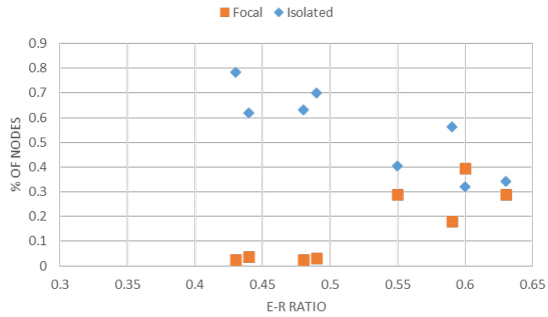

Figure 16: E-R ratio averages versus percentages of focal and isolated nodes

In figure 17, we present the average story size (in actions) contrasted against the percentage of focal and isolated nodes for each agent. We also present the average number of impasses against the percentage of regular nodes for each agent. The Pearson correlation values obtained were 0.87 for focal nodes, -0.86 for isolated nodes, -0.74 for regular nodes, and the R^2 values for $n = 1$ were 0.75, 0.74 and 0.55 respectively. The first values represent a positive linear relation between the story size and the number of focal nodes inside the story generator, the second values represent a negative linear relation between the story size and the number of isolated nodes, and the third values represent a negative linear relation between the number of impasses and the number of regular nodes.

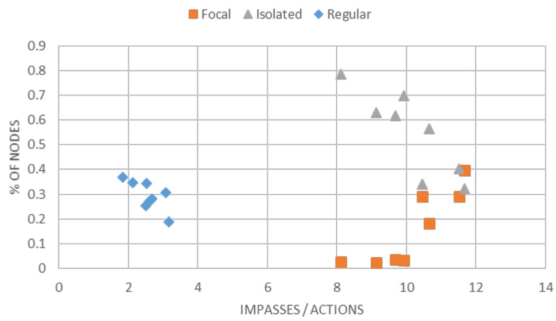

Figure 17: Story size averages versus percentages of focal and isolated nodes, and impasse averages versus percentages of regular nodes

Discussion

The main goal of this project was to identify how the knowledge structures of an automatic generator of narratives influence the presence of diverse features of its generated stories.

In table 1, we present a summary of the linear relations found between the story-characteristics and the structural-characteristics described on this paper. These results are divided into two sections: node relations and cluster relations.

Element	Positive relations	Negative relations
Focal nodes	novelty, climax story size character threads E-R ratio	
Regular nodes		impasses
Isolated nodes		novelty, climax story size character threads E-R ratio
City clusters	opening	closure soc. resolutions
Town clusters		originality
Island clusters	originality closure soc. resolutions	opening

Table 1: Summary of the obtained relations

We utilized as story-characteristics opening, climax, closure, originality, novel contextual structures, impasses, E-R ratio, story size, character threads and social resolutions, in grounds of a previous work on evaluation of stories to determine its interestingness, novelty and coherence (Pérez y Pérez 2014), features considered relevant for a story to be considered creative.

We made use as structural-characteristics the percentage of nodes and clusters inside the knowledge base of the analyzed agents. Nodes represent CSs obtained after interpreting the previous stories of the agents, and clusters represent groups of similar nodes. We defined the concept of similarity between nodes in terms of the similarity between the relations of the CS-contexts. Representing the internal knowledge of an agent as CSs let us qualitatively describe it, which lead into the creation of structures, called C-maps, to visualize the similarities among its information.

Our findings let us now formulate questions about the process of incorporating new stories into the agent's knowledge base. Before this research, we envisioned to have an agent with as many previous stories as possible, but know we have evidence that this is not always the best scenario. For instance, this agent would be a deficient evaluator since its evaluations, in particular for novel CSs and originality, will often be low. This assumption lead us to redefine our definition of novelty. Now, we perceive diverse scenarios where novel CSs emerge: when a new story originates different context from those in the evaluating agent; when a new story utilizes the existing contexts but in different ways; when a new story utilizes rare contexts. With the categorization presented for nodes and clusters, we are able to identify these new context types, to measure its presence, and to validate how it affects the story generation process.

These results lead us to think on the optimal number of previous stories that an agent should have to generate higher-evaluated stories, and to become an accurate story evaluator. If an agent had enough stories to cover all the possible story-contexts, its evaluations of novelty and originality will always be zero, and the number of possible continuations for every story would be so vast that unusual and even incoherent stories could be generated. Its C-map would consist of focal nodes galore, and a big city-cluster. In general, as we develop a better understanding of the implications of diverse knowledge arrangements for the story generation process, we will be able to progress in the construction of more accurate ways of generation and evaluation of such outputs.

It is worth to mention that our final averages does not consider all the 90 stories generated by every agent, since some of these outputs lacked of what we considered as basic characteristics to be considered stories (a minimal number of actions, preconditions satisfied and one character thread). We also measured the ratio of these valid stories against the invalid stories and we looked up for relations with our structural metrics, but we did not find any linear relation. These results give us an inkling of the complexity of generating valid stories. In further research, we will look for non-linear relations and multifactorial relations to cast light on which structures might diminish the generation of invalid stories.

In grounds of our presented results, we showed that, in general, focal nodes improve the novelty of the generated stories because of its conception process. These nodes are built from similar inspiring stories when their CS are extracted and incorporated to the repository. In fact, these nodes provide a wide variety of continuations for a single context since every connection to a focal node comes from a similar CS that can be employed to progress a new story. Moreover, the size of the generated stories is bigger when focal nodes come into play because of this higher number of possibilities, and becomes easier to reach an appropriate number of tensions during the story climax, and to maintain a unique character thread. On top of that, the number of $actions_E$ increases, and is closer to the number of $actions_R$, resulting in a higher E-R ratio.

We also found that, in general, isolated nodes play the opposite role of focal nodes. For instance, they diminish the novelty, climax and the size of the generated stories. Nevertheless, an isolated node can be perceived as a focal node in an early developmental stage, so they are required for the focal nodes to come into play.

Regarding to clusters, cities provide a solid ground for the stories to initiate, but as the process continues, cities widen the number of possible continuations and the stories tend to have closures with multiple unsolved tensions. Contrasting with our initial assumptions, we did not find any evidence of a strong negative relation between cities and originality nor a strong positive relation between them and valid stories.

On the other hand, the presence of islands in the early stages of a story originates multiple impasses, but they incorporate original paths and bounded closures. Finally, towns diminish the originality of the stories since they provide solid structures with multiple similar contexts, but they still lack of focal nodes so the continuations are still not too different.

These results support our claim about the existence of linear relations between structural elements in the knowledge base of our storyteller and features of its generated stories. In our model, these elements are obtained from a set of previous stories, which shows how previous experiences affect the generation of new narratives. Nevertheless, we still need to do additional research efforts to validate if the obtained relations are causal (i.e. the structural-characteristics are the origin of the story-characteristics), or circumstantial (i.e. the structural and the story characteristics are both generated by additional factors). This research has widened our scope to identify the existence of these additional factors, to progress in our understanding of how the structural elements inside the knowledge base of any agent affects the characteristics of its generated narratives.

Conclusions

We showed in this paper relations among structural settings of the knowledge base of an automatic storyteller (Mexica) and features of its generated stories.

We introduced the concept of nodes and clusters built upon CSs inside the agents' knowledge bases. We classified nodes into three different categories: focal, regular and isolated, and also classified clusters of these nodes into three different sets: cities, towns and islands. We have described connectivity maps (C-maps), which reflect how similar the nodes inside the knowledge base of a storyteller are.

We described a set of metrics to identify story features such as preconditions fulfilled, novel contextual structures, opening, climax, closure, character threads, social resolutions, originality, E-R ratio, and number of impasses, and a set of metrics to describe knowledge structures inside the agents based on the nodes and clusters they contain.

We hypothesized how nodes and clusters, when present in the knowledge structure of an automatic storyteller, affect diverse story features. Next, we validated these claims by implementing our model utilizing Mexica and Social Mexica, evaluating each of the generated stories, and then contrasted the evaluations against each of the metrics describing the internal structure of the knowledge bases utilized during the generation process.

Acknowledgments

This research was sponsored by the National Council of Science and Technology in Mexico (CONACyT), project number: 181561.

References

Freytag, G. 1896. *Technique of the drama. An exposition of dramatic composition and art*. S. C. Griggs and company.

Guerrero, I., and Pérez y Pérez, R. 2014. Social mexica: A computer model for social norms in narratives. In *Proceedings of the Fifth International Conference on Computational Creativity*.

Klein, S. 1965. Control of style with a generative grammar. *Language* 41:619–631.

Massey, F. 1951. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association* 46:68–78.

Pérez y Pérez, R., and Sharples, M. 2001. Mexica: a computer model of a cognitive account of creative writing. *Journal of Experimental and Theoretical Artificial Intelligence* 13(2):119–139.

Pérez y Pérez, R. 1999. *Mexica, A computer model of creativity in writing.* Ph.D. Dissertation, University of Sussex, England.

Pérez y Pérez, R. 2007. Employing emotions to drive plot generation in a computer-based storyteller. *Cognitive Systems Research* 8:89–109.

Pérez y Pérez, R. 2014. The three layers evaluation model for computer-generated plots. In *Proceedings of the Fifth International Conference on Computational Creativity*.

Pérez y Pérez, R. 2015. A computer-based model for collaborative narrative generation. *Cognitive Systems Research* 36-37:30–48.

Riedl, M. 2004. *Narrative Planning: Balancing Plot and Character*. Ph.D. Dissertation, Department of Computer Science, North Carolina State University, Raleigh, NC.

Turner, M. 1994. *The Creative Process: A Computer Model of Storytelling*. Hillsdale: Erlbaum (Lawrence).