

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA

**Departamento de Software e Inteligencia Artificial- Lenguajes y
Sistemas Informáticos**



**A COMPUTATIONAL MODEL FOR AUTOMATED
EXTRACTION OF STRUCTURAL SCHEMAS
FROM SIMPLE NARRATIVE PLOTS.**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Carlos León Aznar

Bajo la dirección del doctor

Pablo Gervás Gómez-Navarro

Madrid, 2011

ISBN: 978-84-694-2084-3

© Carlos León Aznar, 2010

are removed from the non-coherent set and the next one is obtained as the definitive set.

$$\begin{aligned}
 & \text{breakup}(z?, x?) \wedge \text{love}(x?, y?) \diamond \text{die}(x?) \\
 & \text{despise}(x?, z?) \wedge \text{forces}(z?, x?) \diamond \text{die}(x?) \\
 & \text{despise}(y?, x?) \wedge \text{love}(x?, y?) \diamond \text{breakup}(y?, x?) \\
 & \text{help}(z?, y?) \diamond \text{breakup}(z?, x?) \\
 & \text{love}(x?, y?) \diamond \text{despise}(x?, z?) \\
 & \text{ill}(x?) \wedge \text{kill}(z?, y?) \diamond \text{forces}(z?, x?) \\
 & \text{kill}(y?, z?) \diamond \text{help}(y?, x?) \\
 & \text{want}(z?, x?) \diamond \text{kill}(z?, y?) \\
 & \text{ill}(x?) \diamond \text{love}(y?, x?) \\
 & \text{together}(y?, x?) \diamond \text{want}(z?, x?)
 \end{aligned}$$

The whole experiment took approximately 8 minutes and a half. This means an average time of 25 seconds per story, approximately. Although the creation of the rules by hand from the same domain has not been carried out, it is claimed that this pseudo-automatic process is much faster, based on experience. Therefore, the solution is promising according to its objectives. At least, for simplistic domains as the one that has been tested. It is planned to study more complex domains as part of the future work.

7.5.1. Saturation

A limit for the experiment has been imposed: once the user reaches a point where most stories are rated as coherent, the experiment stops. It can be seen in the previous section that this limit has been set to 5. This is because during the early tests a *saturation level* was discovered.

After some classifications, the current procedure is unable to extract more rules in a robust way. While it is possible to find more rules, this does not happen as fast as in the first stage (*non-saturation*) of the execution. It was empirically found that once 5 or 6 stories in a row are rated as coherent, the probability of being in the saturation stage is quite high, so it was decided to stop the experiment once this happens.

Saturation happens because of the experimentation system itself, the abstract model for gathering rules is not creating the saturation. The way in which modifications are made to generated stories to create new ones has a limit. In order to keep resemblance with the stories that can be generated by only applying rules, some constraints are cutting too many possibilities. The amount of different patterns it can create is small, therefore eventually exhausting this set of patterns. The proposals for future

work, in Chapter 9, examine other possible approaches to story generation in order to overcome this limitation.

7.6. Overall Results

In order to evidence that useful domain information can be extracted in general, and not only in one single case, the test explained in the previous section was repeated for other subjects. In total, the test was run for 14 people (including the subject from the previous example). Their age ranged between 25 and 60 years old, and all of them are native Spanish speakers, so the surface realization for stories was only implemented for Spanish. It is considered that the messages are so simple that the particular natural language is not affecting the identification of coherence, although discussion is pertinent about this matter (Section 8.1.6).

The test was run once per each human evaluator. Each experiment finished when the saturation zone was reached. The experiment was stopped at that point and the evaluator was not asked to run the test again. Therefore, 14 experiments have been executed. The statistical data graphically shown in 7.3 is acquired from that number of experiments, then.

Elapsed time to reach saturation was also measured. If acquiring rules pseudo-automatically for some domain is slower than doing it by hand, the use of the solution becomes discussable. The mean elapsed time was 8.78 minutes, with a standard deviation of 2.752. The fastest classification took 4.224 minutes and the slowest one took 13.407 minutes.

The average time per story was 26.20 seconds with a standard deviation of 7.02.

Regarding the comparison against crafting the rules by hand, it has to be taken into account that the proposed system is able to generate a set of rules from some input corpus without human intervention, if it is desired. While it has been shown how this approach would be incomplete, it is clear that the benefit in terms of time is noticeable.

Figure 7.3 shows the average structural coherence refinement process from all the evaluators. It can be seen how the proportion between coherent and non-coherent stories is almost lineally raised during the execution of the tests.

It is assumed that the content of the stories is affecting the execution, that is, not every type of plot or any input set would yield these results. The reasons for this are explained in this chapter, mainly in Section 7.1. While the results are promising and the empirical results shows that pseudo-automatically acquiring rules is possible, the application of this approach



Figure 7.3: Average proportion of coherent vs. non-coherent stories during the refinement process. The relation between the last 5 stories is shown.

to other domains (mainly more complex ones) is planned as part of the future work.

For instance, if the tests are run using just one story as corpus, the learning process will be much faster, as it can be seen in Figure 7.4. While it is just an example and, depending on the evaluator and on the input story it does not have to happen in all cases, the learning process is quite straightforward because all the generated stories are coherent according to human criteria. This is just because acquiring rules from one single story leads to a rule-set that only creates stories very similar or even only equal to the source story (which is coherent a priori). This is one of the cases in which this approach is not really useful: more stories are needed.

7.7. Chapter Summary

This chapter details the implementation of the theoretical model described in the previous chapters. Additionally, it shows the experiments that have been run with that implementation to prove, to the possible extent, that the main hypotheses are plausible. Graphical depiction of some executions are shown as well.

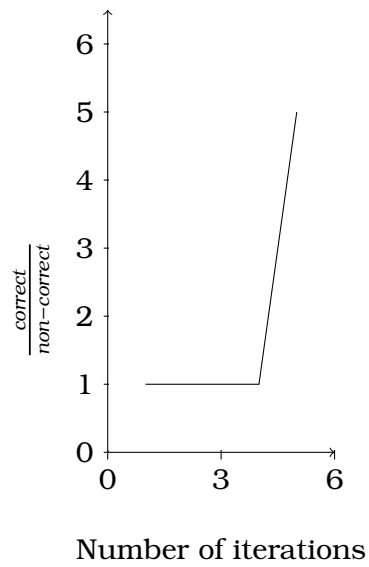


Figure 7.4: Example result of the rule acquisition process when using one single story as input corpus.

Chapter 8

Discussion

THIS CHAPTER DISCUSSES ALL THOSE DECISIONS that have been taken during the development of this research, trying to show the justification of having chosen them. Given the characteristics of the current research, some assumptions are subject to discussion because they are outside the pure formal or computational domain. While some of them are hard or impossible to demonstrate with absolute certainty, this part of the dissertation tries to show that, at least to some extent, the taken approach is useful according to the current focus of the development.

This chapter have been thematically organized, instead of keeping the order followed in the previous chapters. It has been considered that a chapter devoted to discussion is better understandable with such a layout.

8.1. Conceptual Aspects of the Current Approach

This research is based on an adoption of a structural focus to computational narrative processing. Therefore, the extent to which this is really useful and plausible from a conceptual point of view must be discussed.

The main assumption in this research is that empirical results are valid, and the fact that the modelled method does not mimic human psychological processes is not influential. As introduced in Chapter 1, the reduced knowledge available about underlying human mental processes on narrative makes it very difficult to implement systems that *really* replicate humans.

8.1.1. Number of Used Stories in the Rule-Extraction Process

Algorithm 3 only learns from one single story. While results show useful outputs (Chapter 7), the system could learn content from more than one story. This would be beneficial in principle because there is implicit content in the *set* of stories that could be extracted, not only in the stories individually. For instance, the system could extract a rule r_a from a story s_a and a rule r_b from the story s_b . Taken independently, r_a and r_b could be useful, but they could be incompatible together.

Since the rule extraction algorithm must return a single set of rules, these must be coherent as a set, and not only independently. So far, the rule extraction process only ensures coherence for single stories.

While developing the modification of the algorithm to make it able to process more than one story, the search space became much bigger. Because of the exponential nature of the solution based on search, adding more candidates made the problem intractable again. Additionally, extracting rules from the whole corpus at once turned the system into a non-scalable one, because adding more stories to the corpus would make the search space grow exponentially instead of growing lineally. This is why it was decided to process one story at a time. However, apart from the execution cost, nothing prevents from analysing more than one story.

8.1.2. Impact of the Input Corpus in the Final Results

The chosen set of input stories affects the overall results of the experiments shown in Chapter 7 because they are the base for the generated stories. Different corpora would create different stories with different characteristics. While it is assumed that the resulting rules will depend on the input content, it is important to discuss the way in which the *quality* of the results is affected by the process of creation of the corpus.

In the particular implementation that has been shown, operas have been summarized and translated to a formal version by hand, which has introduced the author's criteria in the execution. It makes sense then to question if the author's knowledge has influenced the results.

According to this, the summarization and translation of human-written stories to formal representation processable by the algorithm is considered to be one of the main drawbacks detected in the development of this research. In this stage, author's knowledge and intuition have admittedly influenced the whole process. In the instantiation of the model for operas, the real *libretto* is very different from the outline, and even the outline of the

events taking place in the opera are inherently different from their formal version.

While other summaries for the operas and other possible formalizations would have been possible and they would have possibly led to different values, it is claimed that, since the influence of the author's knowledge takes place in the first stage only, the quality of the results is not really affected because it is driven by external humans' opinion. That is, if the stories coming from the author's translation were totally non-coherent, the rules acquired by the algorithm would create non-coherent stories as well. On the other hand, if coherence in the input stories led to coherent results, all generated stories in the rule-gathering process would lead to acceptable stories *always*, which is not the case, as it has been shown in Chapter 7.

This does not mean that this translation is the best method, though. The option of translating the stories by hand was chosen to reduce the needed effort to get the system running. Training other humans to create short summaries of operas and then having them formalize this outlines to the required formal representation was not always possible, specially when the formal model and the computational requirements of the implementation have been updated several times during the development. If other individuals had created the processable versions of the stories, they would have been forced to do it several times and with several formalisms.

External influence on translations, or even pseudo-automatic translation, which is hypothetically true, would be, in the opinion of the author, better options for future versions of the proposed system.

8.1.3. Influence of Aspects Different from Coherence in Classification

As shown in Section 7.5, humans were required to focus on narrative coherence when rating stories in the experiments. This explicit restriction was clearly stated in the questionnaire to reduce the influence of other aspects of stories on the rating. Experience from the work presented in Chapter 4 showed that the opinion about the overall quality that readers perceive when they read a story influences the rating of particular variables. That is, if a reader finds the story funny or interesting, she or he will consider that the story is good, therefore it is possible to find coherence in a story that is actually lacking good coherence.

However, it is difficult to measure the extent to which this has been correctly reached. Since psychological aspects are unavoidably involved when measuring coherence, and human psychological processes for iden-

tification of coherence are not completely known, no formal study about this is possible.

In order to partially address this issue, the model assumes that validation against human criteria is the best way to evaluation. That is, if human evaluators rate a story as coherent, the rules are correct. This decision is a consequence of the structural definition of this work, which tries to avoid cognitive models. If the stories are not coherent from a narratological or psychological point of view but evaluators rate it as coherent, it is valid for the system.

8.1.4. Influence of Human *Opinion* in the Preconditional Rule Extraction Process

Establishing a general definition of coherence in stories valid for every possible story is not possible because particular human opinion is involved in the process. There is no way of formalizing human opinion and therefore it does not make sense to give more importance to one opinion over another in all cases. Therefore, the particular set of human evaluators driving the rule gathering process explained in Section 6.3 affects the final set of acquired rules. That is, the final set can be non-valid for other humans.

The results collected in Section 7.6 suggest that the evaluation for coherence is reasonably uniform for a simple domain under somewhat restricted conditions, but this does not have to be the case in all scenarios. It is hypothesized that more complex domains in which the evaluation of coherence in stories will not be uniform can be found. The study of this hypothesis and its consequences is targeted as future work.

8.1.5. Good and Bad Stories

Narrations have been divided in the model in two partitions: *good* and *bad*. That is, according to the model, two good stories are not comparable to each other to test which one is better, they are just *good*. An analogous method is followed for bad stories.

From a cognitive point of view, this is far from complete. Not only stories can be “good”, “normal” or “very good”, for instance, but also “good” stories can be considered “bad” as well if several human judgements are taken into account.

This common aspect of evaluation is only partially addressed as a possible side effect of the classifications obtained from different human evaluators, but the theoretical model does not take this into account. This is

a limitation of the model, which should be improved in future versions to take this into account.

On a different tack, the set of “good” stories was defined in terms of coherence: structural coherence from the point of view of the computational system, and coherence as understood by humans in the experiments. It is accepted that *quality* is not totally covered by coherence, and both short and complex plots, even on its high level representation (leaving out literary aspects), are rated differently about its quality taking into account many more aspects than only coherence. While it is assumed that coherence is a basic starting point for describing quality in narrations, the model accepts this restriction in order to offer a working solution, instead of a broader but incomplete one. Whether coherence is one of parts of basic quality is discussable as well, but that discussion is considered to be outside the scope of this dissertation.

8.1.6. Language for Text Realization

Experiments have been carried out by presenting the stories in natural language to human evaluators. As all the evaluators were native Spanish speakers and this was known before implementing the text realization, simple patterns, the experiments were therefore carried out with texts in Spanish.

The quality of the texts is low, but the patterns and their joint as a full story is grammatically correct. No special check about this correctness was carried out because basic knowledge of Spanish was enough to ensure it.

Both aspects could be discussed regarding text realization: the influence that the particular way of telling the plots exerts on evaluators, and the way in which the particular language (Spanish, English or whatever other language) affects human criteria.

Regarding the first aspect, the influence of the quality of the texts, it is assumed that the patterns are so simple that no real influence is being exerted. As a loose test, it was checked during the execution of the tests that evaluator’s criteria were very similar to the author’s criteria for every story, with only slight differences. This suggests that the text patterns created by the author, which had defined the meaning for the kernels of the actions, expressed what he meant, and that the evaluator was approximately interpreting the same thing. This is obviously not a formal check of the assumption, but text realization is not an objective of this work and it is therefore accepted that the results are valid.

The example of realized story shown in Formal Story 11 shows the text in English, and the one shown in Formal Story 10 shows the Spanish

version. It is hard to evaluate their similarities without speaking both languages, but these examples try to show an evidence (for those who speak English and Spanish) that the meaning of the realized actions is very similar in both version.

8.2. Comparison Against Other Approaches

The proposed method is a very particular approach to story generation. There are, as it has been shown in Chapter 2, many other computational approaches to narrative. This section compares and discusses those points in which the comparison is interesting.

8.2.1. Classic Machine Learning and Rule-Gathering Algorithm

To be able to gather rules, the algorithm needs, as input, a set of stories for use it as a “training corpus”. This corpus, however, must satisfy some constraints. Any set of texts put together is not enough to make the system work. The approach is based on narrative structural knowledge that short, simple narrations contain, not on statistical properties of elements.

Most machine learning algorithms can learn most of the times no matter the quality of the corpus, the accuracy of the learning depends on the quality of the corpus. However, the proposed algorithm lacks this ability: if the input stories are not processable by the evaluation function, nothing will be learned from them.

On the other hand, the results from this learning process are totally usable after the acquisition: the set of rules is perfectly readable and modifiable by humans. Additionally, certain approaches to Machine Learning have already developed useful techniques for some aspects related with the current work, like, for instance, acquisition of time relations [Mani et al., 2006, Bramsen et al., 2006].

The other approach considered worth the comparison is the acquisition of causal rules from an external knowledge base. The overall benefit is clear: there is no need of creating a knowledge base, and this saves time and effort. However, possible options like OPENCYC [CyCorp, 2010] or CONCEPTNET [Liu and Singh, 2004] have several characteristics which make it difficult to apply the current approach to evaluation of narrative content, namely:

- Nothing ensures that information about the working domain is avail-

able on these databases. While they capture (in different ways) common sense knowledge, some structure may well not be present.

- Causality is considered, but not as much developed as in the proposed learning algorithm. Full focus on causality is put in the presented system (because it is the structure on which the definition of preconditional link relies), but acquiring robust link information from these knowledge bases is not simple because that information is not as abundant as some other type of relations (as existential properties, for instance).
- Multiple links between concepts are hard to find. Cases in which some fact is the result of the *combination* of two different facts are seldom found, and that construct is quite frequent, at least, in the domain used for the implemented prototype.
- Improving the knowledge base for adapting it to new circumstances (new domains) requires, in the case in which the domain knowledge is not present, hand-writing rules, which implies the problems previously analysed.

It is finally concluded that the option of learning the rules is the most profitable one. Its cost is quite low compared to the option of hand-writing rules, and it provides a more robust coverage inside concrete domains. It comes at the cost of developing the learning algorithm (which is not low), but once this has been created, the adaptation to new domains is, in principle, straightforward.

But that approach is not perfect, since the results are not optimal and it is quite sensible to errors. Putting out human supervision carries the risk of letting the system infer wrong rules, although this supervision can be obviously done *post hoc*. Another benefit of this approach is that any improvement on the learning algorithm can directly benefit all working domains. The current approach to learning is not the optimal one, but it can probably be improved for future versions of the evaluation function.

8.3. Discussion related to the Implementation

The implementation of the theoretical model presented in chapters 3, 5 and 6, as detailed in Chapter 7, has involved additional assumptions and restrictions that have realised the algorithms in a real program. How these adaptations have affected the collected results is discussed in this section.

8.3.1. Appropriate Number of Input Stories for the Corpus

For carrying out the presented results, 9 adaptations of operas have been used. As discussed in Section 8.1.2, it is assumed that the way in which the input corpus is used does only partially affect the results because human supervision is what really drives the rule acquisition process.

However, the amount of input stories affects not only the particular set of rules that is extracted after the execution of the process, but also the time needed to reach the saturation point. If only *one* story was used and this story was coherent, the acquired set of rules used to generate stories would lead to only one single structure of narrations, very similar to the original story. Correcting this set of rules through classification of generated stories would only require to check a few examples (the exact amount would depend on the length of the story). So the amount of different stories that can be created is directly proportional to the set of preconditional rules, which is directly proportional to the number of input stories in the corpus, *a priori*.

This does not have to happen always because the actions in the stories are very influential in the rules. In the most extreme case, the system could be fed with many copies of the same story, this would be exactly the same as feeding it with one instance of that story. In the opposite case each story has a set of actions totally different from the actions in the rest of input stories. This would lead to many rules, but only one rule for each action, which would make the iterative process of refining the set of rules quite slow.

As it has been intentionally carried out in the experimentation, the pseudo “optimal” case takes place when the availability of several stories in which the set of actions is not the same, but every actions appears in several stories with different previous actions. This leads to several rules for each action, which helps to create instances and to iteratively create a useful set of “good” and “bad” preconditional rules.

8.3.2. Application of the Structural Information Extraction Algorithm to Other Domains

Results obtained after this study prove that the gathering of structural relations in simple stories can be pseudo-automatically acquired. This is considered an important merit of the research, but further studies must be carried out in order to prove that the proposed method is valid for other domains.

It is well known, as it has been presented in Chapter 7, that the concrete set of actions in the input story and the proportion in which they appear against the amount of input stories in the corpus determine the final quality of the rule-refinement process, which affects itself the generative capabilities.

On the one hand, the imposed restrictions about what a *simple* story is reduces the scope of the approach. That is, these structural approaches are not necessarily valid for complex narrative forms like novels, which are not addressed. Although the study of the inherent differences between complex narrations and simple stories has not been carried out, many assumptions that have been made for this work to be feasible are not satisfied in complex narratives, so the proposed algorithm is quite unlikely to work in those cases.

On the other hand, short versions of opera plots are not the only possible input domain. While reduced operas are useful for the particularities explained in Section 7.1.3, many other domains fulfill the same requirements, so they could be used as well, and nothing prevents, in principle, a successful execution.

The main point, then, is that the definition of “simple story” is quite restrictive, but that many forms of narrative fall in this category. Fables (if a good subset of them could be found), tales or short descriptions of travels. Although future work must address this to prove it, it is hypothesized in advance that it makes sense to keep on studying the structural analysis of narrative as a useful computational method.

8.4. Chapter Summary

This chapter discussed the most interesting and conflicting points of the research. Both the strong parts and the weaknesses of the current approach are addressed, and future lines of discussion, outside the scope of this thesis, are proposed.

Chapter 9

Conclusions and Future Work

THIS LAST CHAPTER OF THE DISSERTATION summarizes the presented work and proposes different possible further ways in which the explained research could be improved.

While there is still much to be done in this field (structural approaches to computational narrative), this dissertation has tried to create a plausible and useful proposal that can be used, if not for solving problems in the industry, at least for triggering new research in these terms.

9.1. Summary

All along the previous chapters, these have been the subjects that have been studied:

- The objectives and hypotheses that motivated this work were presented in the first chapter.
- The previous research setting the base for this research was studied in detail in Chapter 2.
- The formal and conceptual definition of *narration* used in this research has been defined in Chapter 3.
- The preliminary work on a computational system for processing narrations based on a semantic approach is then detailed.
- A structural approach to computational processing of narrative has been modelled and described in Chapters 5 and 6.
- Then, the implementation, experiments and obtained results are shown in Chapter 7.

- Finally, the most relevant points are discussed and summarized.

All this dissertation has been written trying to focus on what were considered the most important ideas and contributions of the research, without leaving out all the techniques, algorithms and references that base the work.

9.1.1. Scientific Contributions

The main scientific contributions are summarized in this list:

- A formalism for representing narrations, and a definition of *simple narration* that is adopted in this work. This is described in Chapter 3.
- The development of a knowledge intensive system in which an explicit evaluation function drives a generative process, and its partial validation through human supervision (Chapter 4).
- The definition of an artificial relation used as schematic unit to perform schema acquisition and rule learning, the *preconditional link* (Chapter 5).
- A set of algorithms for evaluating a story in terms of its structural properties and for acquiring instances of schemas and rules (Chapters 5 and 6).
- A framework for human-supervised acquisition of generation rules and narrative generation (Chapter 6).

9.2. Conclusions of this Research

This section tries to resume the conclusions that have been gathered after the obtained results were analysed. The most important part of these conclusions is the validation –or rejection– of the main hypothesis according to the results of the empirical tests.

The hypotheses, defined in Section 1.3, state that a structural approach to narrative not taking into account cognitive models can be implemented for some domains in such a way that stories recognizable as such by humans are generated. It was also hypothesized that the defined structural properties could be pseudo-automatically acquired.

After the analysis of the results shown in Chapter 7 it can be concluded that the hypotheses have been *partially* validated. Through the implementation and the tests it has been shown how *it is possible* to build a system performing structural processing on narrations that creates stories that are considered *coherent* by humans.

The other part of the model, the proposed rule-acquisition algorithm, has been implemented as well and tested as part of the experiments, so the second part of the hypotheses, the plausibility of acquiring rules pseudo-automatically (in the proposed way) has also been *partially* demonstrated.

A special emphasis is put on the fact that the hypotheses have only been *partially* proven. This is considered so because:

- The experiments involve human criteria. This means that the evaluation of coherence is only partial, and that many more human evaluators could be used to perform a more specific study. However, due to the prototype characteristics of the implementation, this was considered unnecessary at this point.
- The inherent focus of this thesis is related to narrative, which is a very complex concept which does not have a fully accepted definition. Thus, stating that the stories *are coherent*, while plausible, cannot be formally proven.

Although only partial validation has been carried out in this research, the overall conclusion is that the hypotheses, to the possible extent, are plausible enough to keep on researching on the field of structural processing of narrations in these terms.

9.3. Benefits and Drawbacks

No research project is perfect, and this one is not an exception. This sections briefly summarizes the main benefits and drawbacks that have been presented along the exposition of the previous chapters.

It has been concluded that the current contribution reduces, to some extent, the human effort required to improve the domain of generation of narratives for an application. While this is considered an obvious benefit, it must also be made clear that the proposed system, so far, is only applicable to simple domains and short and plain narrations. While more work could lead to an improved model providing coverage to more sophisticated forms of stories, this has not been carried out yet.

The supervision system is very simple and based on a straightforward boolean classification. It is claimed that this is a clear benefit because it is very easy for humans to classify content in this way. Additionally, such a simple boolean system does not necessarily have to be explicit, it is hypothesized that this can be queried by other means (although this study is outside the scope of this research). But this simple classification comes at a price: only a basic conception of coherence can be gathered using this method. Interest, for example, is not boolean, there are virtually infinite levels of interest.

Theoretically, the proposed model covers the full space of solutions for all the proposed algorithms. This permits a theoretical full coverage of the solutions, which is good. However, this process has not been feasible in practice because the spaces are so big that traversing them completely is not a tractable problem. Therefore, the theoretical advantage is hindered by the practical limitations of current computational power.

Finally, it is considered that an important part of the contribution is the proposal of a totally structural paradigm for computationally management of narrative content. It has been shown how this can be positive with an original model. But it is hypothesized that, while this could open new paths for research, it is quite likely that only structural processing will not be enough by itself, and semantic computation will be required to achieve really good quality, at least when comparing the stories generated by computers to those generated by humans.

9.4. Future Work

The previous chapters have defined the bound of this research, leaving several points as part of the future work. This section summarizes those improvements on the system or related research lines that are being currently considered. While there could be many others, these are directly related and planned as the next issues to study.

9.4.1. Improving the Model for Structural Definition and Processing of Narrations

It has been empirically shown that structural processing of stories makes sense, at least for some domains. However, the application of the proposed approach is not limited to the presented system, and it is hypothesized that more sophisticated and powerful solutions can be developed.

For instance, the classification of the stories could be improved so that the evaluation function and the human criteria are not boolean, but range on some real interval. In this way, plain separation between *coherent* and *non-coherent* stories could be improved and a *ranking* of stories could be performed. This would lead to the concept of a story which is “more coherent” than some other. This was the approach followed in the cognitive system developed and presented in Chapter 4.

Figure 6.1 shows a diagram of the proposed preconditional rule acquisition method. While it has been proven that it makes sense to build rule-sets in this way, some improvements can be applied to the algorithm so that the rule gathering approach is made more powerful. For instance, the first accepted candidate set of rules is taken as the solution. This is not necessarily optimal, and more options could be considered. For instance, it would be possible to select the *best* candidate based on a non-boolean classification of story according to their coherence.

The model could be also improved by adding more complex time relations to the definition of *simple stories*. The assumptions that were made are considered to be extremely restrictive. This was done to keep the presentation of the model simple and focused, but it is hypothesized that the model in its current form can be applied to stories in which the duration of each action is higher than a single unit of time, for instance. More complex time relations would broaden the range of application of the model.

In general, this expansion of the range of stories for which the structural model is able to give a correct solution must be improved as the next step in this research for the system to be really useful. It is assumed that humans handle stories that are much more complex than those presented in this dissertation, and reaching, at least partially, that expressive power is a very important –and, perhaps, ambitious– objective that, in the opinion of the author, should be targeted.

9.4.2. Improved Story Generation

The rule-acquisition method is based on creating candidate stories that are classified under supervision. The quality of the story generation system affects both the required time to reach the saturation point in the experiments and the capabilities of finding new rules.

As it has been designed, the acquisition process can only collect the preconditional rules that are implicitly present in the input stories because no new possibilities are included. The saturation point is reached when there is nothing new to “learn”, that is, a local maximum regarding the preconditional rules has been hit. In order to avoid this, *noise* could be

introduced in the model.

That is, to be able to gather new preconditional rules, relatively original stories, slightly different from those which can be generated with the rules from the original stories, could be created. In this way, the rule-acquisition process could expand its capabilities by changing the way in which stories are built.

This opens new issues to solve, as for instance the amount of noise to add to the generated stories. Too much noise will generate an excessive amount of non-coherent stories, and too little noise will not be useful at all.

It would be interesting as well to use more sophisticated Natural Language Generation systems for text realization. The one that has been shown here is too poor, and the human criteria about stories is probably being affected by the realization and not only the underlying content. This text realization is, however, supposed to happen at a different level than the rest of the system. That is, improving the text realization, while important, should not influence the rest of the algorithm, at least in principle.

9.4.3. Improving the Implementation

The implementation of the theoretical model has been carried out for testing purposes, and it has not been adapted for use beyond this. To make it possible further research of the model, and given that all the demonstration of the validity of such a kind of model is carried out empirically, it makes sense to create a version of the implementation that can be used in a more robust way.

On the one hand, the implementation could be executed faster. During the middle stages of the development of the prototype, parallelization of the code was taken into account. While it was finally discarded in order to keep a fast development process, due to the type of algorithms, parallelization can make it possible to explore bigger subsets when searching candidate preconditional rules and links and when performing story generation.

On the other hand, the development of such a model is intended to be useful to the research community, so modifying the prototype following a software design approach that makes it possible to implement and release the system as a library –or whatever other form of distributable and usable system– makes sense.

9.5. Applications of the Structural Processing of Narratives

The overall objective of this research problem is not only the development of a theoretical model for managing narrative texts, but also the application of the results to real applications. While the current version of the implementation is still far from being applicable in a real scenario, it is considered that the ideas, once built together in a working system, could improve or ease certain human tasks.

The first application that could benefit from the use of this model is plain story generation. While, as it has been shown, the solution is yet unable to produce high-quality stories, nothing prevents from applying the concepts developed in this research to create a simple story generator. This hypothetical generation system would probably reach a higher quality of generation by creating a mixed approach in which both structural and cognitive methods are used.

Not only classic story generation could be the objective of this model. One of the main benefits of the proposed algorithms is that rule acquisition can be carried out pseudo-automatically, which leverages the possible range of applications to any scenario in which human criteria can be gathered. Since the ratings that are required for the system to work are boolean, one can imagine several situations in which this is easily doable.

For instance, in narrative videogames, the player could play an automatically generated story and then rate it as coherent or non-coherent. Then, rules could be extracted from this rated story, and they could therefore, as the model proposes, be included in supervised rule-sets. This would provoke an iteratively refined generation system which would lead to the generation of coherent stories without any additional human intervention in principle.

The synthetic proposed implemented scenario for demonstrating that the preconditional rule acquisition is possible does not represent well real scenarios in general. While a commercial system could require preliminary training of the generation process, a more user-friendly system would perform this rule extraction in a more indirect way.

For instance, in an even more sophisticated approach, the opinion about coherence could be extracted in a more automatic way using non-explicit methods. Instead, the way in which the player plays the game could be analysed. For a trivial solution, non-coherence stories could be directly rated as such if the player stops playing it without having finished the story, for instance. This is just a hypothetical scenario and no real

solution is being proposed, but it is considered that this kind of systems could be built as an application of this model. Of course, rating coherence in this way would be polluted with the implicit rating of many other aspects as interest or engagement. This would have to be addressed in the creation of a system like this one.

Many other uses of this research could be found, and this section has only intended to propose a brief perspective of how the structural processing of narrative can be applied. While some of this applications are in general solvable by any approach to story generation or understanding, doing it with algorithms based on structure can offer new things that, at least partially, can make the solution of these long term problems more reachable.

9.6. Chapter Summary

This chapter concludes the thesis by examining the main outcome, and proposes further lines of investigation. Both the benefits and the drawbacks of the research process are highlighted, and the extent to which the hypotheses suggested in the first chapter have been demonstrated is discussed.

Parte II

Resumen de la tesis en español

Capítulo 10

Introducción

EL TRATAMIENTO SEMÁNTICO de los textos narrativos trata de alcanzar, desde uno u otro punto de vista, la interpretación de cómo los humanos procesan historias. En general, las aproximaciones semánticas han predominado en los sistemas de generación [Meehan, 1976, Turner, 1992, Bringsjord and Ferrucci, 1999, Riedl, 2004]. Mientras que nada evita que la investigación en este tipo de sistemas pueda llegar a tener éxito, no es una tarea fácil. Los problemas de la IA en general (el cuello de botella de la adquisición del conocimiento, y otros) también aparecen en la Narrativa Computacional. Como una cantidad importante de información se requiere para implementar este tipo de sistemas, el coste de construir estas bases de conocimiento ha sido una barrera en el desarrollo de programas de este tipo.

De acuerdo con estas características de la Narrativa Computacional, una importante desventaja de los sistemas de generación es que, mientras que están contruidos sobre modelos genéricos, el tamaño de historias que pueden procesar es reducido en comparación al coste de inserción de conocimiento. Esta es una de las razones que ha evitado que la generación de historias se aplique en la industria. En general, esto ha sido así por el *cuello de botella de adquisición del conocimiento*.

En particular, el foco científico de esta investigación se ha puesto en intentar mejorar la escala a la que se genera en relación al esfuerzo de añadir conocimiento a mano. El proyecto propone un sistema de adquisición semiautomática de esquemas narrativos. El objetivo es ampliar la escala de generación mediante la reducción de esfuerzo para incluir nuevas reglas.

10.1. Motivación de la investigación

Varios sistemas de generación ya existen, pero la cantidad de historias que pueden generar es muy reducida. Para generar nuevas historias, nueva información del dominio tiene que ser incluida a mano, lo cual hace aumentar el coste de mejora y escalado de las capacidades de proceso. Esto es costoso y la experiencia prueba que impide crear sistemas muy complejos.

En vez de tratar de resolver el problema de adquisición de conocimiento el bloque, tiene sentido restringir el alcance del problema a un subdominio. En concreto, los textos narrativos tienen ciertas propiedades que pueden ser usadas para crear modelos restringidos de adquisición de esquemas.

Tiene sentido, por lo tanto, crear un modelo que, si bien no sustituye totalmente la intervención manual, al menos puede reducir la cantidad de trabajo que los humanos deben desarrollar para crear nuevo conocimiento. Por lo tanto, la principal motivación de esta investigación es el beneficio que el estudio de estas posibilidades puede darse. Como se hace la hipótesis de que es posible reducir la cantidad de trabajo humano necesario para extraer estructuras narrativas, esto es considerado una motivación justificada.

10.2. Objetivos

- Crear un sistema computacional que reciba historias escritas por humanos como entrada.
- De ellas, extraer instancias de algún tipo de estructura, minimizando hasta el límite posible la necesidad de intervención humana. Esta estructura será definida como parte de la investigación.
- Usar esas instancias para generar automáticamente historias.
- Validar el proceso completo a través de evaluación humana, intentando restringir, en la medida de lo posible, el interfaz hombre-máquina a texto en lenguaje natural.

Las narraciones objetivo tienen que cumplir estos requisitos:

- Las narraciones sólo contendrán un sólo hilo narrativo. Complejas y múltiples líneas causales no están permitidas.
- Las narraciones no estarán ordenadas por relaciones temporales complejas. Se imponen estas restricciones:

- Cada evento en las narraciones durará una sola unidad de tiempo.
- Los eventos estarán ordenados por tiempo en las narraciones.

10.3. Hipótesis

Inicialmente, el foco fue puesto en la generación de historias. Como la intención era romper parcialmente el límite que la adquisición de conocimiento pone en este campo, se formuló la hipótesis de que una función de evaluación explícita podía ser suficiente [León and Gervás, 2010].

Sin embargo, esta hipótesis no fue correcta y la función de evaluación no sirvió para mucho. Sin embargo, el proceso llevó a la formulación de una nueva hipótesis (que resultó definitiva), basada en el tratamiento estructural:

Un cierto conjunto de relaciones *estructurales* en una narración es suficiente para crear historias evaluadas como *coherentes* por humanos. Es decir, una vez que estas relaciones son conocidas, la coherencia puede ser analizada en estos términos.

Adicionalmente, se completó la hipótesis con esta sub-hipótesis:

Las instancias de estas relaciones estructurales hipotéticas pueden ser adquiridas semiautomáticamente mediante un algoritmo computacional. Por lo tanto, los patrones extraídos desde historias coherentes y no coherentes pueden ser recogidos como contenido para la generación.

10.4. Metodología

El desarrollo de este trabajo ha sido dirigido por la aplicación de metodología de investigación basada en el estudio del trabajo previo en el campo que nos ocupa y la aplicación de las tecnologías disponibles para probar, hasta el punto posible, la validez de las hipótesis previamente enunciadas.

Una aproximación secuencial para la propuesta científica ha sido llevada a cabo: la identificación del problema, el enunciado de la hipótesis, la sugerencia de una solución, y la experimentación y el análisis de los resultados.

La primera etapa de la investigación consistió en la lectura de literatura científica y no científica en el área de la Narrativa Computacional. Por lo

tanto, los avances en Narratología e Inteligencia Artificial fueron tenidos en cuenta para aprender de un sustrato rico antes de afrontar la identificación de objetivos útiles.

Aparte de un ordenador personal (AMD Phenom™9550 Quad-Core Processor), la implementación ha sido llevada a cabo usando el sistema SWI-PROLOG [Wielemaker, 2010]. Todo el sistema de documentación ha sido maquetado con L^AT_EX [The LaTeX Project, 2010], incluyendo este documento. Para el análisis estadístico, el programa y lenguaje R ha sido usado [R Development Core Team, 2010]. Scripts de post-procesado de datos han sido creados con Lua [Lerusalimschy et al., 2006] y el lenguaje de programación Ruby [Thomas and Chad Fowler, 2005]. Todo el desarrollo, pruebas y análisis han sido ejecutados en GNU/Linux.

Capítulo 11

Trabajo previo

ESTE CAPÍTULO ESTÁ DEDICADO al estudio de la investigación científica previa, cuyo desarrollo influencia este trabajo. Principalmente trata de analizar las técnicas teóricas y prácticas relacionadas con el diseño y el desarrollo de esta tesis.

11.1. Narratología

La Narratología es la ciencia dedicada al estudio estructuralista de las *narraciones* y de la manera en la cual los humanos las entendemos y las usamos. Dado que la aproximación a las narraciones que propone la Narratología está estrechamente relacionado con la Narratología Computacional, es útil comentar algunas de sus más importantes características.

La Narratología se centra en las características internas de la narrativa y en las similitudes y diferencias con otros tipos de comunicación. La Narratología estudia estas características, intentando quedar parcialmente aislada de otros fenómenos como la lingüística o la semiótica. De este modo, la Narratología permanece como una disciplina independiente.

Algunas descripciones de *narración* la definen como un cuento, algunas otras como la acción de contar algo, y otras se centran en la descripción de las partes estructurales del discurso narrativo propiamente dicho [Real Academia Española, 2010]. Este enfoque estructuralista a la descripción de las narraciones establece la base de la Narratología moderna, aunque la ciencia de la Narratología se considera de un modo retrospectivo, quizá enraizando en Aristóteles *Poetica* [Aristotle, 1974]. Aristóteles postuló que la imitación del mundo real crea argumentos de los cuales las unidades más importantes son escogidas y ordenadas en una trama. La imitación de las acciones en el mundo real (*praxis*) crea un argumento

(*logos*) del cual las unidades (*mythos*) son escogidas.

La Narratología moderna tiene sus orígenes en los análisis estructuralistas de los textos, basados en el trabajo de Barthes, Genette, Todorov y otros [Barthes and Duisit, 1975, Todorov, 1977, Genette, 1979]. La Narratología más moderna, post-estructuralista, crea perspectivas en las cuales la ciencia cognitiva hace que se consideren las narraciones como fenómenos psicológicos, y propone un estudio de la narrativa en términos de una perspectiva cognitiva [Herman, 2000].

11.2. Narrativa Computacional

La Narrativa Computacional se dedica a la representación y procesado de las narraciones por ordenadores. Uno de los subcampos más estudiados de la Narrativa Computacional es la generación de historias. Aparece en los años 70 del siglo XX con el interés humano de entender y procesar historias.

Roger Schank fue el pionero del estudio del impacto de la narrativa en humanos desde un punto de vista formal y computable. Schank postuló que la manera de la que la memoria funciona no está sólo basada en procesos que manipulan datos mentales, sino que es un proceso continuo de recuerdo y adaptación de historias previas que definen nuestro mundo, como en el Razonamiento Basado en Casos (que se origina en el trabajo de Schank) [Schank and Abelson, 1977, Schank, 1982].

Schank propone el término *script*, que consiste en unidades cortas de conocimiento secuencial sobre los pasos típicos que han de ser tomados en una cierta situación. Por ejemplo, Schank propone el clásico ejemplo del restaurante: alguien entra en un restaurante, pide comida, come, etcétera. Este conocimiento común, según Schank, es muy importante en el conocimiento humano.

Para implementar este tipo de conocimiento semántico en ordenadores, Schank propone la *dependencia conceptual*. La dependencia conceptual es un método formal basado en acciones primitivas y relaciones entre ellas y sus objetos [Lytinen, 1992].

Programas como SAM [Cullingford, 1981] or PAM [Wilensky, 1981] comenzaron el desarrollo de una teoría de narrativa computacional basada en personajes y en la manera de la que intentaban alcanzar sus objetivos. Algunos modelos del modo en el que los procesos de la memoria pueden ser implementados han sido también estudiados [Kolodner, 1980].

Estos estudios provocaron la aproximación inversa: en vez de intentar entender los métodos humanos, tenía sentido intentar generar narrati-

va a través de modelos de métodos humanos, como se hizo en TALESPIN, [Meehan, 1976, Meehan, 1981] y otros.

Tras esta etapa inicial, la Inteligencia Artificial aplicada a la narrativa en estos términos fue abandonada, con las excepciones de algunos sistemas [Turner, 1992, Mueller, 1987].

La Inteligencia Narrativa [Mateas and Sengers, 1999] y otros proyectos multidisciplinares en Narrativa Computacional renuevan el interés del campo en el siglo XXI. La formalización teórica de la narrativa, el desarrollo de los sistemas de generación de historias, sistemas de extracción de conocimiento humano narrativo, y otros proyectos, están, de nuevo, acrecentando el interés en el campo.

11.3. Evaluación de Narrativa

La evaluación y la comprensión de la narrativa han sido estudiadas principalmente por la Psicología. Por ejemplo, Kelly analiza cómo la personalidad se desarrolla de acuerdo con las construcciones estructurales que los humanos creamos a través de las historias que leemos u oímos [Kelly, 1955]. Applebee estudió cómo las capacidades evaluativas en narrativa de los niños crecen en paralelo con sus capacidades de contar historias [Applebee, 1978].

En relación con la evaluación computacional de la narrativa, hay interés en encontrar algún tipo de métrica para comprar la calidad de los sistemas de generación de historias. Rowe et al. proponen STORYEVAL, un armazón para medir y comparar sistemas de generación de historias. STORYEVAL sugiere evaluar estos sistemas teniendo en cuenta *métricas narrativas, estudios cognitivo-afectivos, estudios centrados en el director y evaluaciones narrativas características* [Rowe et al., 2009].

Capítulo 12

Definición de narraciones

Teniendo en cuenta la representación de historias que se ha hecho en investigaciones previas, una definición formal de narración ha sido creada y refinada iterativamente para encajar con los requisitos de la aproximación computacional que se trata en esta tesis. Durante el proceso de creación de este modelo, diferentes formalizaciones fueron consideradas [León et al., 2007b, León and Gervás, 2008, León et al., 2008].

12.1. Definición de narración en su relación con modelos conceptuales

Esta investigación define *narración* o *historia* como un conjunto ordenado de hechos por orden cronológico. Por lo tanto se igualan, hasta cierto punto, los términos de *fábula* y *narración* de acuerdo con la definición de algunos formalistas rusos [Propp, 1928]. Es importante notar que esto es una definición *ad-hoc* usada como nomenclatura para esta tesis: no se asume ninguna similitud psicológica, formal o narratológica. De aquí en adelante, cualquier referencia a *narración* o *historia* se refiere al mismo concepto, excepto donde se especifique lo contrario.

Esta investigación se centra en el estudio formal de tramas de historias, y no en características narrativas o lingüísticas más complejas. Este foco en las tramas ha sido también seguido en la mayor parte de sistemas de generación narrativa que se presentan en la literatura [Meehan, 1976, Bringsjord and Ferrucci, 1999, Riedl and Young, 2006]. Por lo tanto, no se hace ningún estudio sobre la realización superficial como texto o las propiedades artísticas. Aunque una generación simple de texto ha sido llevada a cabo por motivos de experimentación, se asume que no se ha conseguido una calidad suficiente.

12.2. Definición formal de narración en este trabajo

La representación formal se centra en la capa de discurso, y, en particular, la definición formal de narraciones no se enmarca en ningún dominio semántico en concreto. Permite, así, la representación de conocimiento del mundo [León and Gervás, 2010]. La semántica que el nivel cognitivo asigna no se tiene en cuenta.

El modelo formal está presentado en orden ascendente en relación con sus partes constituyentes. Primero, los elementos más básicos, los *átomos*, se muestran. Después, las *acciones*, y después las *narraciones* propiamente dichas. Estos nombres de las partes han sido creados para tener una palabra de referencia, y no se asume ninguna similitud con los conceptos que representan en el idioma común. Aunque la selección de los términos está claramente influenciada por conceptos similares de Narratología, sólo el significado formal que se asigna en esta investigación debe ser interpretado cuando se haga referencia a ellos.

No se hace ninguna asunción sobre la aplicabilidad posterior de este modelo de narraciones en otros sistemas. Ha sido creada para este proyecto de investigación y, aunque podría servir para basar otros trabajos, éste no ha sido el objetivo.

12.3. Constituyentes de las narraciones formales

12.3.1. Átomos

Los *átomos* son elementos atómicos que definen una sola cosa, carácter, idea o cualquier instancia simple de cualquier concepto. Pueden representar cualquier entidad en un dominio particular, pero están restringidos de tal modo que no pueden hacer referencia a ninguna parte de la narración. Por tanto, los átomos no pueden representar *acciones* ni *narraciones* (definidas a continuación).

Los *átomos* están unívocamente especificados por su nombre. Esto significa que dos átomos con el mismo nombre representan el mismo concepto. Ejemplos de átomos son *john*, *bird*, *house*, *hope* o *sad*. En este sentido, los *átomos* son similares a los de la lógica de primer orden.

12.3.2. Variables

Más adelante se verá que la evaluación y el algoritmo para la extracción de estructuras para la creación de nuevas historias necesitan *variables* para representar reglas abstractas. Las variables que se proponen en este modelo son similares a las variables lógicas, y pueden representar cualquier átomo. Las variables están representadas por un símbolo seguido de un signo de interrogación. Por ejemplo $x?$ ó *token_variable?* son representaciones válidas de variables.

Una vez que una variable está ligada a un átomo en una ejecución de los algoritmos (descritos más adelante), todas las instancias de esa variable están ligadas a ese átomo. Sólo hay un espacio de variables en el modelo, de tal modo que *todas* las acciones que contienen la variable $x?$ verán ligado su valor una vez se ligue en un punto de la ejecución.

12.3.3. Acciones

Las *acciones* son los constituyentes básicos de las narraciones, y representan eventos en la historia. Las acciones están definidas por estos eventos, propiedades o relaciones (de aquí en adelante, el *núcleo*) y una secuencia ordenada de átomos. La acción, por tanto, enlaza un núcleo con átomos, significando que la acción, propiedad o relación es cierta para ellos.

Un núcleo está representado en forma de una palabra, que puede ser compuesta. Por ejemplo *love* o *take_to* son núcleos. Aunque esto podría dar lugar a confusión porque los núcleos están formados de la misma manera que los átomos, en la práctica esto no ocurre porque los núcleos, en el resto del modelo, siempre forman parte de una acción (tal y como se define después). La estructura de las acciones define formalmente la unión entre el núcleo y los átomos.

De acuerdo con los parámetros enlazados al núcleo en forma de átomos o variables, dos tipos de acción se han definido:

- Las *acciones instancias* son acciones cuyos elementos son átomos (*take(john, glass)*).
- Las *acciones de plantilla* son acciones en las cuales algún elemento es una variable (*love(ofelia, x?)* o *attack(x?, y?)*).

12.4. Narraciones

Las *narraciones* son secuencias ordenadas de acciones. Han sido definidas a partir de las partes anteriormente descritas (átomos, núcleos y acciones). La ecuación 12.1 muestra una representación formal de una narración genérica.

$$n = [a_1, a_2, a_3, \dots, a_n] \quad (12.1)$$

donde $[a_1, a_2, a_3, \dots, a_n]$ es una lista de acciones instanciadas y n es una narración.

Capítulo 13

Procesamiento estructural de narraciones

LA PRIMERA PARTE de la investigación tuvo como objetivo la realización de un prototipo funcional en el que se estudiaba la generación de textos narrativos a gran escala mediante un enfoque semántico. El resultado final fue que, si bien tenía sentido la aproximación, el coste de la inserción de conocimiento era muy alto. Por tanto, se decidió centrar la investigación científica en un ámbito puramente estructural, tal como se describe en este capítulo. El trabajo y los resultados de la primera parte de la investigación pueden encontrarse en [León and Gervás, 2010].

Este capítulo muestra cómo una relación estructural de las narraciones, en particular, puede ser la base del proceso automatizado de narraciones. Esta relación ha sido llamada *enlace precondicional*, y las siguientes secciones explican sus principales características. La selección de una relación simple permite analizar computacionalmente las características estructurales de las historias y construir un algoritmo que extrae reglas para los enlaces precondicionales, las llamadas *reglas precondicionales* (capítulo 14).

13.1. De un modelo cognitivo a una definición estructural

Estudiando los resultados recogidos de la evaluación con humanos en el generador semántico de historias, fue detectado que había una correlación muy clara entre dos variables: la *causalidad* y la *cronología*. La causalidad representaba la percepción de que todo pasa por una razón lógica, y la cronología evaluaba la correcta disposición de los eventos en el

tiempo (por ejemplo, que las causas aparecían antes que las consecuencias). Como se trabajó en un dominio muy simple, fácilmente interpretable por todos los evaluadores, todos ellos evaluaron la calidad de estas dos variables con muy pocas diferencias.

Una heurística fue creada para capturar esto: el *enlace precondicional*. Los enlaces condicionales tratan de describir un patrón estructural que está implicada en el reconocimiento humano de la causalidad y la cronología, ignorando el contenido semántico de estos dos conceptos.

13.2. Propiedades estructurales como coherencia estructural

La aproximación propuesta en este trabajo no pretende emular ningún modelo cognitivo. La propuesta consiste en el cambio desde aproximaciones puramente semánticas a otras más estructurales, y consecuentemente la coherencia en las narraciones está definida en estos términos, sin ninguna asunción cognitiva, como se establece a continuación:

Una historia es coherente en el modelo propuesto si es evaluada como tal por evaluadores humanos.

Una historia es *estructuralmente coherente* si tres patrones estructurales están presentes: el *foco*, en *enlace único* y la *conexión completa*.

El *foco*, el *enlace único* y la *conexión completa* se definen más adelante. No han sido explicados antes porque su definición están enlazada con la noción de enlace precondicional. La ausencia de una base cognitiva ha llevado a la definición de estas propiedades en términos del enlace precondicional, que es detallado más adelante.

Una definición formal de coherencia estructural en una historia s se muestra en la ecuación 13.1. La implementación (capítulo 15) se basa en esta ecuación.

$$\begin{aligned} is_structurally_coherent(s) = is_focused(s) & \quad (13.1) \\ & \wedge is_fully_connected(s) \\ & \wedge is_uniquely_linked(s) \end{aligned}$$

La definición de la coherencia estructural ha sido formalizada como una función de evaluación booleana. El modelo establece que una narración es estructuralmente coherente o *no*.

13.2.1. Foco

Una historia tiene *foco* cuando tiene un *sólo* final. Es decir, termina en una *sólo* acción o evento. Esta definición se basa en el trabajo de Trabasso sobre la coherencia en tramas [Trabasso and Sperry, 1985]. Intuitivamente, si una historia se cuenta y no se puede extraer ninguna conclusión de la narración, podría ser considerada que está incompleta. Esto ha sido interpretado como una falta de coherencia. Una definición formal de foco puede verse en la sección 13.4.1.

13.2.2. Conexión completa

Una historia está *completamente conectada* cuando todas las acciones están relacionadas entre sí. Esta definición trata de capturar intuitivamente la restricción de simplicidad de que sólo un hilo narrativo puede existir en las historias simples. La conexión completa captura el hechos de que ninguna acción en la historia debe ser contada sin una razón y que todos los hechos relevantes deben ser contados (de otro modo, la conexión estaría perdida). La conexión completa está formalmente definida en la sección 13.4.2.

13.2.3. Enlace único

Una historia está *únicamente enlazada* si dos eventos están relacionados por una *single* idea o conclusión. Esta es la definición conceptual de otra parte de la simplicidad en las historias coherentes. Si dos eventos están relacionados con otro por varias relaciones, más de un hilo está presente. Por lo tanto, la historia no es *simple* en los términos establecidos. Esta definición conceptual de enlace único está definida en la sección 13.4.3.

13.3. Enlaces precondicionales

La definición de enlace precondicional sólo se *inspira* en la manera en la que, heurísticamente, los humanos llevan a cabo la interpretación de historias, pero no intenta capturar ningún proceso cognitivo. En vez de eso, la propuesta actual lo define como una heurística para máquinas, y no para humanos. Es decir, la definición está estrictamente ligada a la información computacionalmente procesable: la información que los humanos usamos y la manera en la que lo hacemos no es necesariamente compatible.

En una historia formalizada como la secuencia de acciones $\{e_1, \dots, e_n\}$, las acciones $\{e_i, \dots, e_j\}$ (las *precondiciones*) están *precondicionalmente enlazadas* a e_k (la *consecuencia*) si aparecen antes que e_k y el grafo dirigido resultante de crear enlaces precondicionales para todas las acciones tiene coherencia estructural

1. *foco*, teniendo todos los enlaces convergentes a una sola acción.
2. *conexión completa*, si todas sus acciones enraízan directa o indirectamente en una sólo acción, y
3. *enlace único*, según el cual cada par de acciones conectadas está conectada por un sólo enlace *como mucho*.

donde $1 \leq i, j, k \leq n$.

Esta definición es obviamente estructural, es decir, sólo captura propiedades superficiales de historias de acuerdo a una relación sintética. Como ejemplo, si una historia está compuesta por las acciones $\{a, b, c, d, e\}$, enlaces precondicionales válidos aparecerían en la figura 5, pero no en la figura 6.

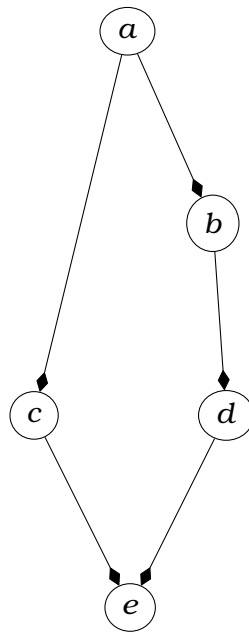
El conjunto de acciones de la precondición de un enlace precondicional representa conjunción. El modelo no considera representar disyunción porque esto llevaría al no-determinismo, que se ignora en este modelo. El proceso de razonamiento sobre la causalidad sería más complejo si una historia no determinara únicamente los enlaces precondicionales de algún hecho.

13.4. Definición formal de patrones estructurales

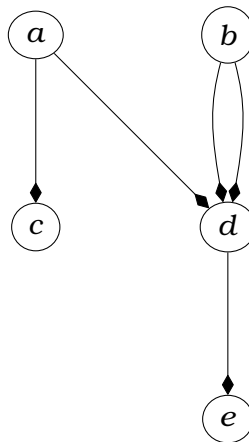
13.4.1. Foco

$$is_focused(s) = |\{a \mid a \in s, \text{ if } is_outcome(s, a)\}| == 1 \quad (13.2)$$

$$is_outcome(s, a) = \forall a' \in s - \{a\}, exist_preconditional_chain(s, a', a) \quad (13.3)$$



Graph 5: Conjunto válido de enlaces precondicionales. Las flechas representan enlaces.



Graph 6: Conjunto no válido de enlaces precondicionales. Las flechas representan enlaces.

13.4.2. Conexión completa

$$is_fully_connected(s) = action_fully_connected(a) \forall a \in s \quad (13.4)$$

$$\begin{aligned}
action_fully_connected(s, a) = preconditional_link(s, a) == root & \quad (13.5) \\
& \vee action_fully_connected(s, a') \\
& \forall a' \in preconditional_links(s, a)
\end{aligned}$$

13.4.3. Enlace único

$$is_uniquely_linked(s) = action_uniquely_linked(a) \forall a \in s \quad (13.6)$$

$$\begin{aligned}
action_uniquely_linked(s, a) = \neg duplicates_in(pl) & \quad (13.7) \\
\text{where } pl = all_preconditional_links_to(s, a)
\end{aligned}$$

13.5. Reglas precondicionales

Las reglas precondicionales son pares de un conjunto de *precondiciones* y una *consecuencia* en la cual las precondiciones son acciones de plantilla (según se definen en la sección 12.3.3), y la consecuencia es una acción de plantilla sencilla. La expresión 13.8 muestra un ejemplo.

$$go(x?, y?) \wedge see(x?, z?) \diamond want(x?, z?) \quad (13.8)$$

13.6. Cálculo de los enlaces precondicionales en una historia

El cálculo de los enlaces precondicionales en una historia consiste en el proceso de asignar los enlaces precondicionales apropiados a los hechos correspondientes en una historia, de modo que se obtenga finalmente una estructura concreta.

El algoritmo 7 muestra una versión en pseudocódigo del algoritmo no determinista para calcular una red precondicional correcta en una historia. Como se muestra en la figura 13.1, la entrada del algoritmo consiste en una historia y un conjunto de reglas precondicionales.

Algorithm 7 Pseudocódigo para asignar una red precondicional a la historia s .

```

1:  $s \leftarrow$  historia actual
2:  $R \leftarrow$  conjunto de reglas precondicionales
3: for  $candidate\_network \in candidate\_preconditional\_networks(R, s)$  do
4:    $l \leftarrow$  apply  $candidate\_network$  on  $s$ 
5:   if  $is\_structurally\_coherent(l) == true$  then
6:     return  $candidate\_network$ 
7:   end if
8: end for
9: return “No se puede encontrar una red precondicional para  $s$ .”

```

Algorithm 8 $candidate_preconditional_networks$: algoritmo en pseudocódigo para encontrar las redes precondicionales candidatas.

```

1:  $R \leftarrow$  conjunto de reglas precondicionales
2:  $s \leftarrow$  historia actual
3: for  $a \in s$  do
4:   for any  $r \in R$  do
5:     if  $a$  puede ser instanciada con  $r$  then
6:       añadir enlaces entre las precondiciones de  $r$  y  $a$ 
7:     end if
8:   end for
9: end for
10: yield  $network$ 

```

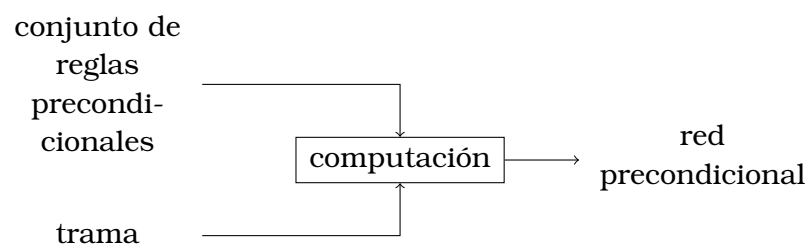


Figura 13.1: Modelo de caja negra el algoritmo de cómputo de enlaces precondicionales.

Capítulo 14

Extracción automática de reglas precondicionales

Este capítulo explica cómo el proceso de adquisición de reglas se lleva a cabo en relación con el modelo.

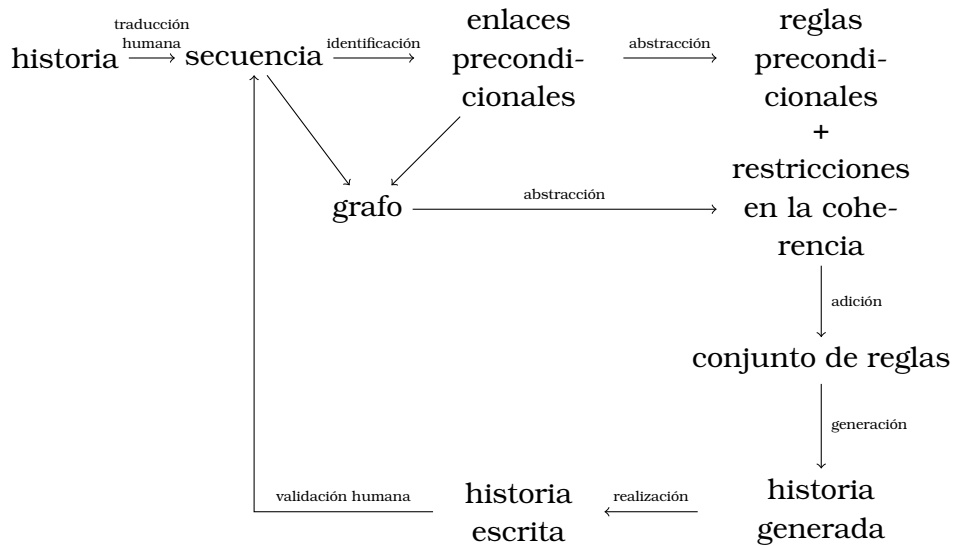


Figura 14.1: Descripción gráfica del proceso de extracción de reglas.

14.1. Algoritmo de extracción de reglas

14.1.1. Corpus de historias de entrada

Se tiene un corpus fijo de entrada, de historias coherentes, al empezar la ejecución. En cada iteración, una nueva historia es generada. Así, se podría decir que un nuevo corpus basado en el anterior se crea: nada evitaría usar las historias coherentes generadas en el proceso como corpus de entrada para otra ejecución, aunque esto no hay sido tratado en este trabajo.

14.1.2. Generación de reglas precondicionales

Para cada historia en el corpus de entrada y para cada historia generada (tal como se muestra en la secciones siguientes), se extraen reglas precondicionales. Esto es llevado a cabo como se detalla en el algoritmo 9. Este algoritmo es análogo al algoritmo para crear enlaces precondicionales mostrado en 7, y abstrae implícitamente las reglas de los enlaces en un proceso de variabilización básico [Charniak and McDermott, 1985].

Algorithm 9 Algoritmo en pseudocódigo para extraer un conjunto de reglas de un corpus de entrada.

```

1:  $s \leftarrow$  historia actual
2: for  $next\_candidate \in$  siguiente conjunto candidato de reglas do
3:    $s_{linked} \leftarrow$  calcular los enlaces precondicionales de  $s$  con  $next\_candidate$ 
4:   if  $\epsilon(s_{linked}) == true$  then
5:     return  $next\_candidate$ 
6:   end if
7: end for
8: return “No se puede encontrar un conjunto de reglas.”

```

Cada conjunto candidato de reglas precondicionales es generado como se describe en el algoritmo 10. La definición formal de la función *candidate_rules* puede ser examinada en el algoritmo 11.

La función *single_rule(story, kernel)* devuelve reglas de acuerdo con el siguiente orden. Se asume que existe una función que crea iterativamente patrones de variables var_a, \dots, var_b .

1. Primero, el átomo especial *raizraiz* \diamond *nucleo*(var_1, \dots, var_n).
2. Después, se genera el kernel de la siguiente acción de la historia *story*: *nucleo*_{accion previa}(var_i, \dots, var_j) \diamond *nucleo*(var_x, \dots, var_y).

Algorithm 10 Orden de generación del los conjuntos de reglas precondicionales candidatas.

```
1:  $s \leftarrow$  historia actual
2:  $K \leftarrow$  conjunto de núcleos de  $s$ 
3:  $R \leftarrow \{\}$ 
4: for  $i \leftarrow 1$  until  $i = |K|$  do
5:   add candidate_rules( $K_i, s$ ) to  $R$ 
6: end for
7: return  $R$ 
```

Algorithm 11 Reglas candidatas para cada núcleo.

```
1:  $s \leftarrow$  historia actual
2:  $k \leftarrow$  núcleo actual
3:  $m \leftarrow$  número máximo de reglas por núcleo
4: for  $i \leftarrow 1$  until  $i = m$  do
5:   yield  $i$  reglas para  $k$  en la historia  $s$ 
6: end for
```

Algorithm 12 Algoritmo de creación de un número concreto de reglas. Auxiliar del Algoritmo 11.

```
1:  $s \leftarrow$  historia actual
2:  $k \leftarrow$  núcleo actual
3:  $j \leftarrow$  número de reglas que van a generarse
4:  $rules \leftarrow \{\}$ 
5: for  $i \leftarrow 1$  until  $i = j$  do
6:   añadir una nueva regla a  $rules$  de acuerdo con single_rule( $s, k$ )
7: end for
8: yield  $rules$ 
```

3. Tras eso, el paso 2 se repite yendo hacia atrás en la historia hasta la primera acción.

Este orden asume un conjunto de variables $\{var_1, \dots, var_j\}$. Las variables son creadas mediante el análisis de los átomos de las acciones en las cuales la creación de reglas precondicionales está basad. El proceso primero examina los diferentes átomos de la historia, devolviendo un conjunto $T = \{t_1, \dots, t_n\}$. Después, un conjunto correspondiente de variables es creado, mapeando cada átomo con una nueva variable $V = \{v_1, \dots, v_n\}$.

Resumiendo, una historia de ejemplo haría que se llevara a cabo la búsqueda de reglas precondicionales para el kernel *go* mediante la exploración de los candidatos siguientes, tras parametrizar, con este resultado:

$$\begin{aligned}
 & root \diamond go(x?, y?) \\
 & root \diamond go(x?, y?); buy(x?, a?) \diamond go(x?, z?) \\
 & \quad buy(x?, a?) \diamond go(x?, z?) \\
 & buy(x?, a?) \diamond go(x?, z?); go(x?, y?) \diamond go(x?, z?) \\
 & \quad go(x?, y?) \diamond go(x?, z?)
 \end{aligned}$$

Finalmente se devolvería este conjunto de reglas:

$$root \diamond go(x?, y?); buy(x?, a?) \diamond go(x?, z?)$$

14.1.3. Generación de historias

Como el modelo teórico no impone ninguna restricción, la definición del sistema computacional para generar historias se detalla como parte de la implementación en la sección 15.2. Es importante dejar claro en este punto que la calidad de este generador de historias está lejos de ser alta, y es un defecto aceptado de la implementación que se propone, por no ser uno de los objetivos de la investigación. En el capítulo 17 se propone un sistema de generación de historias más sofisticado.

14.1.4. Adquisición de criterio humano

De forma análoga a la generación de historias, el modelo teórico de este capítulo no impone ninguna manera particular de aplicar criterio humano en las historias generadas para clasificación. Siguiendo una perspectiva de caja negra 14.2, una historia es insertada en esta parte del proceso, y se espera un valor booleano por parte de un humano que coincida con su opinión sobre la coherencia de la narración.

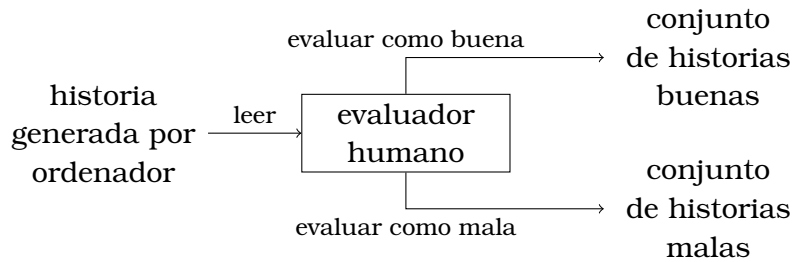


Figura 14.2: Modelo de caja negra de la clasificación booleana de historias por humanos.

14.1.5. Recogiendo y mezclando reglas precondicionales buenas y malas

En una aproximación simple, las reglas precondicionales extraídas sólo desde historias coherentes eran usadas como reglas básicas en las cuales un patrón de unificación simple generaba una historia válida. Tal sistema fue implementado. En las pruebas de la implementación de este sistema, historias claramente no coherentes fueron generadas, lo cual era un error.

Para solucionar esto, el modelo actual considera un conjunto de reglas particionado en *reglas buenas* y *reglas malas* tras la intervención humana. Para abstraer esta clasificación de reglas, tras la clasificación de *la historia*, sus reglas son extraídas y añadidas a uno u otro conjunto.

- El conjunto de reglas buenas contiene los enlaces precondicionales de las historias buenas.
- El conjunto de las reglas malas conjunto de reglas de historias consideradas no coherentes siempre que estas reglas no estén en el conjunto anterior.

Es decir, si una historia considerada correcta, s_i , tras el análisis, genera las reglas precondicionales $\{r_1, r_2, r_3\}$ y una historia s_j , no correcta, crea las reglas $\{r_1, r_4, r_5\}$, el conjunto *bueno* contiene las reglas $\{r_1, r_2, r_3\}$ y el conjunto *malo* contiene las reglas $\{r_4, r_5\}$.

Capítulo 15

Implementación, evaluación y resultados

15.1. Corpus de historias

15.1.1. Historias de asesinatos

El primer prototipo funcional de la función de evaluación fue construido para tramas cortas de historias de asesinato. La intención era estudiar la plausibilidad de una función de evaluación para historias, no enlazada únicamente a la corrección [León and Gervás, 2010]. Esta implementación consideraba un conjunto de variables mayor, las cuales fueron elegidas sin ningún modelo narratológico o psicológico. La selección estaba basada en la intuición del autor sobre lo que constituye una buena historia.

Los resultados empíricos era prometedores. La evaluación humana mostró que crear una función de evaluación que mide varias variables, basada en el análisis secuencial de versiones formales de los textos puede llegar a asemejarse al criterio humano. Sin embargo, se hizo patente que simplemente cambiando el foco no era suficiente: el cuello de botella de adquisición de conocimiento aún estaba presente.

15.1.2. Aesop's Fables

Una vez que los límites de la aproximación basada en conocimiento fuera claramente experimentados, el proceso estructural con fábulas fue llevado a cabo según se explica en los capítulos 13 y 14. Se encontró un problema relacionado con el dominio en particular: Aunque las fábulas tienen una estructura similar y muy simple, muchos eventos diferentes aparecen en ellas (muchos verbos). Como esto ocurre, el número de reglas

precondicionales adquiridas por cada núcleo es muy bajo. Generalmente, en este caso, sólo hay una o dos reglas por cada caso en todo el corpus, lo cual es muy bajo. Este caso hace que sea muy difícil experimentar dado que muy pocas reglas por núcleo siempre dan lugar a las mismas historias. Es decir, aunque estas nuevas historias serían coherentes, serían muy parecidas unas a otras.

15.1.3. Óperas

Después de descubrir el problema de las fábulas, se decidió trabajar con óperas de siglo XIX. Estas óperas clásicas siguen una patrón muy clásico basados en las pasiones humanas y finales trágicos. Esto sugirió que era posible adaptar las óperas de tal modo que una traba básica fuese abstraída como historia simple. Esta adaptación fue requerida porque las óperas no cumplen el requisito de ser cortas y con un sólo hilo. Sin embargo, fue posible identificar hilos simples principales en las óperas.

15.2. Generación de historias simple

La generación de historias propuesta es muy simple y directa. Tiene que ser tenido en cuenta en la generación de historia como tal no es el objetivo de esta investigación, por lo tanto, esta aproximación a la generación de historias puede ser fácilmente mejorada. Sin embargo, esto ha sido dejado como trabajo futuro (capítulo 17). El sistema de generación de historias propuesto informa la operación de exploración mediante la aplicación de las reglas precondicionales extraídas hasta el momento de la generación. Es decir, en vez de generar candidatos sin ninguna restricción, sólo aquellos candidatos que pueden posteriormente satisfacer las reglas precondicionales son creados. Como estas reglas han sido cogidas de la aplicación del proceso de adquisición, las reglas deberían capturar, hasta cierto punto, esquemas narrativos coherentes. Las pruebas para esta hipótesis se detallan en las secciones siguientes.

15.3. Ejecución de un usuario

La secuencia de evaluaciones se presenta en la figura 15.1.

Para representar de una manera más gráfica la evolución del criterio sobre las historias generadas, es posible representar la proporción entre las historias coherentes y no coherentes para las últimas 6 generaciones.

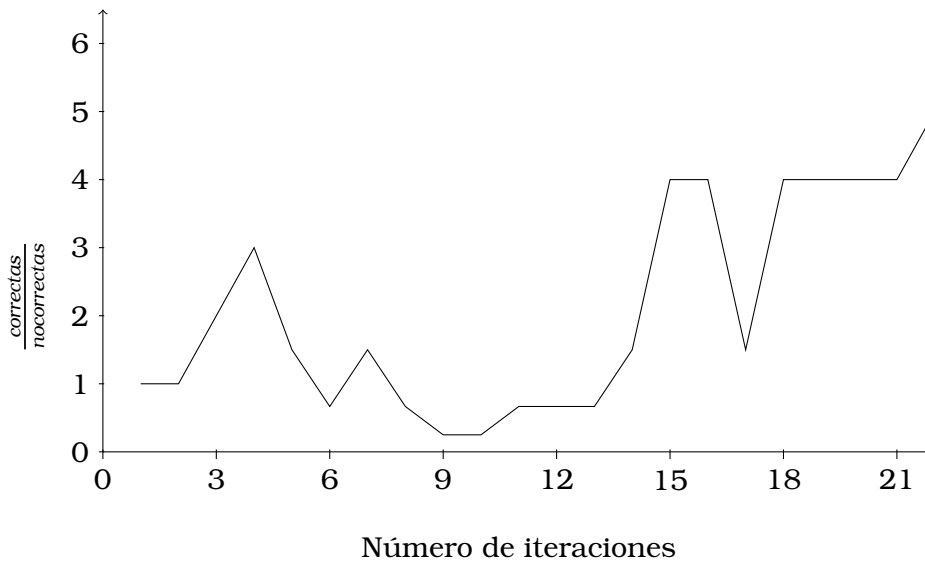


Figura 15.1: Curva de aprendizaje para un usuario.

El experimento concreto tardó en ejecutarse aproximadamente 8 minutos y medio. Esto significa un tiempo medio de 25 segundos por historia, aproximadamente. Aunque la creación de reglas a mano para el mismo dominio no ha sido llevada a cabo, se asume que este sistema pseudoautomático es mucho más rápido, basándonos en la experiencia. Por tanto, la solución es prometedora de acuerdo con sus objetivos. Al menos, para dominios simples como el que se muestra en estos resultados. Se planea estudiar dominios más complejos como parte del trabajo futuro.

15.3.1. Saturación

Un límite se impuso en el experimento: una vez que el usuario llega a un punto en el que la mayoría de las historias son evaluadas como coherentes, el experimento se detiene. Puede comprobarse en la sección anterior que este límite ha sido establecido en 5. Esto es así porque durante los primeros ensayos se encontró un punto de *saturación*.

Tras algunas evaluaciones, el procedimiento actual fue incapaz de extraer más reglas de una manera robusta. Mientras que es teóricamente posible encontrar más reglas, esto no ocurre tan rápido como en la primera fase de la ejecución (*fase de no saturación*). Fue empíricamente comprobado que una vez que 5 ó 6 historias seguidas son clasificadas como coherentes, la probabilidad de estas en la zona de saturación es muy alta,

así que se decidió establecer este punto como la condición de parada.

La saturación no tiene lugar por el sistema de experimentación en sí mismo. La manera en la que se hacen modificaciones a las historias generadas para crear nuevas tiene un límite. Para mantener la semejanza con las historias que pueden ser generadas mediante la pura aplicación de reglas, algunas restricciones están podando demasiadas posibilidades. La cantidad de patrones diferentes que esto puede crear es pequeña, por lo tanto, se agota, en un momento dado, este conjunto de patrones. Las propuestas para el trabajo futuro, en el capítulo 17, examinan otras aproximaciones posibles a la generación de historias para superar esta limitación.

15.4. Resultados globales

Se ha medido el tiempo medio en alcanzar la saturación. Si la adquisición pseudoautomática de reglas para un dominio es más lenta que hacerlo a mano, la utilidad de esta solución es discutible. Como media, se tardó 8.78 minutos, con una desviación estándar de 2.752. La evaluación más rápida tardó 4.224 minutos, y la más lenta 13.407. El tiempo medio por cada historia fue de 26.20 segundos, con una desviación estándar de 7.02.

En relación a la comparación con la creación manual de historias, tiene que ser tenido en cuenta que el sistema propuesto es capaz de generar un conjunto de reglas a partir de un corpus de entrada sin intervención humana, si se desea. Mientras que ha sido demostrado que esta aproximación es incompleta, es claro que el beneficio en términos de tiempo es notable.

La figura 15.2 muestra la coherencia estructural media del proceso de adquisición de reglas de todos los evaluadores con los que se ha experimentado. Puede ser comprobado cómo la proporción entre historias coherentes y no coherentes aumenta casi linealmente durante la ejecución de las pruebas.

Se asume que el contenido de las historias afecta a la evaluación, es decir, no todas las tramas y no cualquier corpus de entrada daría estos resultados. Las razones se explican en este capítulo, principalmente en la sección 15.1. Mientras que estos resultados son prometedores y la evaluación empírica muestra que adquirir reglas de forma pseudoautomática es posible, la aplicación de esta aproximación a otros dominios (principalmente a aquellos más complejos) se planea como parte del trabajo futuro.

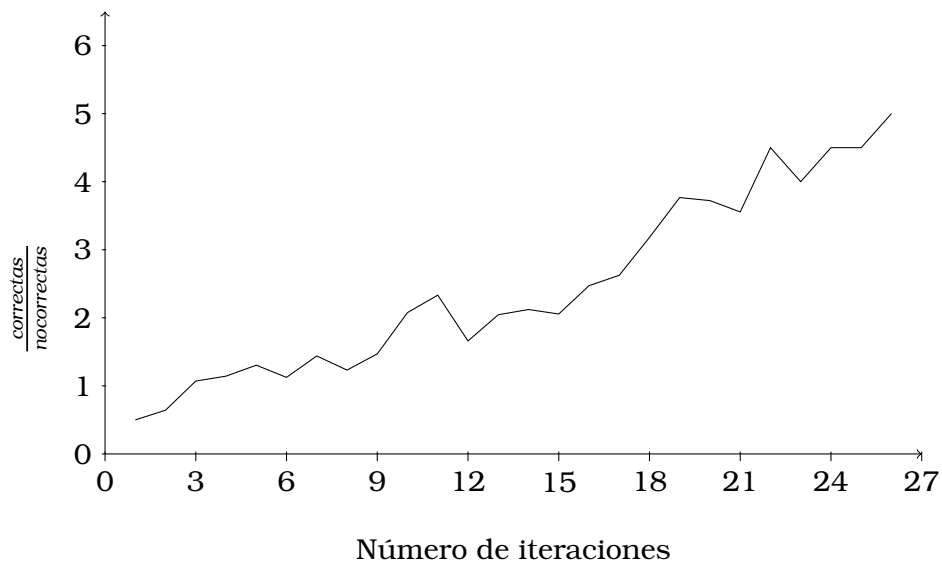


Figura 15.2: Proporción media de historias coherentes y no coherentes durante el proceso de adquisición de reglas. Se muestra la proporción de las últimas 5 historias.

Capítulo 16

Discusión

16.1. Aspectos conceptuales de esta solución

La principal asunción en esta investigación es que los resultados empíricos son válidos, y que el hecho de que el método modelado no imite los procesos psicológicos humanos no es influyente. Como se presentó en el capítulo 10, la reducida cantidad de conocimiento disponible sobre los procesos mentales humanos en narrativa hace muy difícil implementar sistemas que emulen *realmente* a los humanos.

16.1.1. Número de historias en el proceso de extracción de reglas

El algoritmo 9 sólo aprende de *una* historia. Mientras que estos resultados ofrecen resultados útiles (capítulo 15), el sistema podría aprender contenido de más historias al mismo tiempo haciendo una variación básica en el algoritmo. Esto sería beneficioso en principio porque hay contenido implícito en el *conjunto* de historias que podría ser extraído, no sólo de las historias individualmente. Por ejemplo, el sistema podría extraer una regla r_a de una historia s_a una una regla r_b de s_b . Aunque por separado podrían ser útiles, juntas no lo serían.

Como el algoritmo de extracción de reglas debe devolver un solo conjunto de reglas, deben ser coherentes como conjunto, y no sólo de forma independiente. Hasta ahora, el algoritmo sólo asegura la coherencia para historias simples.

16.1.2. Impacto del corpus de entrada en el resultado final

El conjunto de historias elegido como corpus de entrada afecta a los resultados generales de los experimentos mostrados en el capítulo 15 porque son la base de las historias generadas. Corpus diferentes crearía diferentes historias con distintas características. Mientras que se asume que las reglas resultantes dependerán en el contenido de entrada, es importante discutir la manera en la que la *calidad* de los resultados es afectada por el proceso de creación del corpus.

En la implementación en particular, las óperas han sido resumidas y traducidas a una representación formal a mano, lo que introduce el criterio del autor en la ejecución. Tiene sentido, entonces, tener en cuenta que esto es una fuente de error. Actualmente, sin embargo, no hay manera de hacer esto automáticamente.

16.1.3. Influencia de otros aspectos diferentes a la coherencia

Tal y como se muestra en el capítulo 15 se les pedía a los evaluadores que se centraran en la coherencia narrativa cuando evaluaban historias. Esta restricción explícita fue claramente expuesta en el cuestionario para reducir la influencia de otros aspectos de las historias en el resultado. La experiencia adquirida presentada en la primera parte de la investigación [León and Gervás, 2010] muestra que la opinión sobre la calidad global que los lectores perciben cuando leen una historia influye en la valoración de otras variables más concretas. Es decir, si un lector encuentra que la historia es graciosa o interesante, considerará que la historia es buena, por lo cual es posible encontrar una buena valoración para una historia que no es realmente coherente.

Para afrontar este hecho parcialmente, el modelo asume que la evaluación con criterio humano es la manera global de evaluación. Es decir, si los evaluadores humanos consideran una historia coherente, las reglas son *correctas*. Esta decisión es la consecuencia de la definición estructural de la propuesta, que intenta evitar modelos cognitivos. Si las historias que no son coherentes son evaluadas como tal por humanos, es válido para el sistema y el alcance de esta investigación.

16.1.4. Influencia de la *opinión* humana en el proceso de extracción de reglas

Establecer una definición general de coherencia en historias válida por todas las historias posibles no tiene sentido porque la opinión de individuos particulares está implicada en el proceso. No hay manera, actualmente, de formalizar la opinión humana, y no hay razón, a priori, para dar más relevancia a una opinión que a otra en todos los casos. Por lo tanto, el conjunto particular de evaluadores humanos en el proceso de adquisición de reglas definido en la sección 14.1 afecta al conjunto final de reglas adquiridas. Es decir, el conjunto final puede ser inválido para otros humanos.

Los resultados recogidos en la sección 15.4 sugiere que la evaluación de la coherencia es razonablemente uniforme para un dominio simple bajo ciertas condiciones más o menos restringidas, pero esto no tiene que ser el caso para todos los escenarios. Se hace la hipótesis de que dominios más complejos en los cuales la evaluación de coherencia en historias no será uniforme pueden ser encontrados. El estudio de esta hipótesis y sus consecuencias se planea como trabajo futuro.

16.2. Comparación con otras aproximaciones

Para que se puedan recoger reglas, el algoritmo necesita, como entrada, un conjunto de historias para usarlo como “corpus de entrenamiento”. Este corpus, sin embargo, debe satisfacer ciertas restricciones. Cualquier conjunto de textos reunidos no es suficiente, en general, para hacer que el sistema funcione. Esta aproximación está basada en conocimiento estructural en narrativa que narraciones simples y cortas incluyen, no es datos estadísticos.

La mayor parte de algoritmos de aprendizaje máquina pueden aprender la mayor parte de las veces sin importar la calidad del corpus. La exactitud del aprendizaje depende de la calidad del corpus. Sin embargo, el algoritmo propuesto carece de esta habilidad: si las historias de entrada no son procesables por la función de evaluación, nada podrá ser aprendido.

Por otro lado, los resultados de este proceso de aprendizaje son útiles totalmente tras su adquisición: el conjunto de reglas es perfectamente modificable y legible por humanos.

La otra aproximación considerada que merece la pena comprar es la adquisición de reglas causales desde una base de conocimiento externa. El beneficio general es claro: no hay necesidad de crear una base de conocimiento, y esto ahorra tiempo y esfuerzo. Sin embargo las opciones posibles

como OPENCYC [CyCorp, 2010] o CONCEPTNET [Liu and Singh, 2004] tienen ciertas características que hacen difícil aplicar la aproximación actual a la evaluación de contenido narrativo.

Se concluye que la opción de aprender reglas es la más beneficiosa. Su coste es bastante bajo en comparación con el coste de crear las reglas a mano, y provee una cobertura robusta dentro de dominios concretos. Tiene el coste de desarrollar el algoritmo, pero una vez éste ha sido creado, la adaptación a nuevos dominios es, en principio, directa.

Capítulo 17

Conclusiones y trabajo futuro

17.1. Conclusiones de la investigación

Tras el análisis de los resultados mostrados en el capítulo 15 puede ser concluido que las hipótesis han sido *parcialmente* validadas. A través de la implementación y las pruebas ha sido demostrado cómo *es posible* construir un sistema que lleva a cabo un proceso estructural que crea historias que son consideradas *coherentes* por humanos.

Un énfasis especial es puesto en el hecho de que las hipótesis han sido sólo demostradas *parcialmente*. Esto es considerado así porque:

- Los experimentos implican criterio humano. Esto significa que la evaluación de coherencia es sólo parcial, y que muchos más evaluadores humanos podrían haber sido usados para llevar a cabo un estudio más específico. Sin embargo, debido a que el sistema es un prototipo, ha sido considerado innecesario.
- El foco inherente de la tesis está relacionado con narrativa, que es un término muy complejo que no tiene una definición única aceptada. Por lo tanto, sentenciar que las historias son *coherentes*, mientras que plausible, no puede ser formalmente demostrado.

17.2. Beneficios e inconvenientes

Se ha concluido que la contribución actual reduce, hasta cierto punto, el esfuerzo humano que se requiere para mejorar el dominio de generación de narrativas para una aplicación. Mientras que esto se considera un beneficio claro, tiene que quedar claro que el sistema propuesto, hasta aquí, sólo es aplicable a dominios simples y a narraciones cortas y sencillas.

Mientras que un trabajo más detallado podría llevar a un modelo mejorado que provea cobertura a formatos narrativos más sofisticados, esto aún no se ha hecho.

Teóricamente, el modelo propuesto cubre el espacio completo de soluciones para todos los algoritmos propuestos. Esto permite una cobertura teórica de las soluciones, lo cual es bueno. Sin embargo, este proceso no ha sido realizable en la práctica porque los espacios en los que se trabaja son tan grandes que una exploración completa es intratable en términos de computación. Por lo tanto, la ventaja teórica está bloqueada por las limitaciones prácticas.

Finalmente, se considera que una parte importante de la contribución es la propuesta de un paradigma totalmente estructural para el proceso computacional de contenido narrativo. Ha sido mostrado cómo esto puede ser positivo con un modelo original. Pero se hace la hipótesis de que, mientras que esto podría abrir nuevos caminos de investigación, es bastante probable que el proceso estructural por sí mismo no sea suficiente, y que aproximaciones semánticas sean necesarias para alcanzar un grado decente de calidad, al menos en comparación con las historias de otros sistemas de generación automática y de humanos.

17.3. Trabajo futuro

17.3.1. Mejora del modelo

Se ha demostrado empíricamente que el proceso estructural de historias tiene sentido, al menos para algunos dominios. Sin embargo, la aplicación de la aproximación propuesta no está limitada al sistema presentado, y se hace la hipótesis de que soluciones más sofisticadas y potentes pueden ser desarrolladas.

Por ejemplo, la clasificación de historias podría ser mejorada de forma que la función de evaluación y el criterio humano no sean booleanos, sino que estén en un intervalo real. De este modo, una separación simple entre historias *coherentes* y *no coherentes* podría ser mejorada y un *ranking* de historias podría ser creado. Esto llevaría a un concepto de historia en el que unas son “más coherentes” que otras.

La figura 14.1 muestra un diagrama del sistema de adquisición de reglas precondicionales expuesto. Mientras que ha sido demostrado que tiene sentido construir conjuntos de reglas de este modo, algunas mejoras pueden ser aplicadas al algoritmo de modo que la aproximación a la adquisición de reglas sea más potente. Por ejemplo, el primer candidato aceptado se toma

como solución. Esto no es óptimo necesariamente, y más opciones pueden ser tenidas en cuenta. Por ejemplo, también, sería posible elegir el *mejor* candidato en base a una clasificación no booleana según la coherencia.

El modelo puede ser mejorado también mediante la adición de relaciones de tiempo más complejas a la definición de *historias simples*. Las asunciones que se han hecho consideran un modelo extremadamente restrictivo. Esto se hizo para mantener el modelo simple y enfocado, pero se hace la hipótesis de que el modelo en su estado actual podría ser aplicado a historias en las que la duración de cada acción es mayor que una sola unidad de tiempo, por ejemplo. Relaciones temporales más complejas ampliarían el campo de aplicación.

17.3.2. Generación de historias mejorada

El proceso de adquisición de reglas se basa en crear historias candidatas que están clasificadas bajo supervisión. La calidad de el sistema de generación de historias afecta al tiempo requerido para llegar al punto de saturación en los experimentos y en la capacidad de encontrar reglas nuevas.

Tal y como ha sido diseñado, el proceso de adquisición puede sólo recoger las reglas precondicionales que están implícitamente presentes en las historias de entrada porque no se incluyen otras posibilidades. El punto de saturación se alcanza cuando no hay nada nuevo que “aprender”, es decir, un máximo local en relación a las reglas precondicionales se ha alcanzado. Para evitar esto, se puede introducir *ruido* en el proceso.

17.3.3. Mejoras en la implementación

La implementación del modelo teórico ha sido llevada a cabo con motivos de prueba, y no ha sido adaptada para mejor uso tras esto. Para hacer posible que se lleve a cabo una investigación más avanzada, y dado que la demostración de la validez de tal modelo se hace empíricamente, tiene sentido crear una versión de la implementación que puede ser usada de una manera más robusta.

Por una parte, la implementación podría ser ejecutada más rápido. Durante las fases medias del desarrollo del prototipo, la paralelización del código fue tenida en cuenta. Mientras que fue finalmente descartada para mantener un proceso de desarrollo simple, debido al tipo de algoritmos que se usan, la paralelización puede hacer posible expirar subespacios mayores en la búsqueda de reglas precondicionales y enlaces en la generación de historias.

Por otra parte, el desarrollo de tal modelo intenta ser útil para la comunidad científica, de tal modo que modificar el prototipo siguiendo un patrón de diseño software que haga posible implementar y publicar el sistema como una librería -o cualquier otra forma de sistema distribuible- tiene sentido y utilidad.

Bibliography

- [Aesop, 1992] Aesop (1992). *Aesop's Fables*. Houston: Advantage International, The PaperLess Readers Club.
- [Allen, 1991] Allen, J. F. (1991). Time and time again: The many ways to represent time. *International Journal of Intelligent Systems*, 6:341-355.
- [Applebee, 1978] Applebee, A. (1978). *The Child's Concept of Story: Ages Two to Seventeen*. Chicago: University of Chicago Press.
- [Aristotle, 1974] Aristotle (1974). *Poética*, volume 8. Coleccion Biblioteca Románica Hispánica IV.
- [Bal, 1998] Bal, M. (1998). *Narratology: Introduction to the Theory of Narrative*. University of Toronto Press, second edition.
- [Barthes and Duisit, 1975] Barthes, R. and Duisit, L. (1975). An introduction to the structural analysis of narrative. *New Literary History*, 6(2):237-272.
- [Bernstein, 2009] Bernstein, M. (2009). On hypertext narrative. In *HT '09: Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 5-14, New York, NY, USA. ACM.
- [Bramsen et al., 2006] Bramsen, P., Deshpande, P., Lee, Y. K., and Barzilay, R. (2006). Inducing temporal graphs. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, EMNLP '06*, pages 189-198, Morristown, NJ, USA. Association for Computational Linguistics.
- [Bringsjord and Ferrucci, 1999] Bringsjord, S. and Ferrucci, D. (1999). *Artificial Intelligence and Literary Creativity: Inside the mind of Brutus, a StoryTelling Machine*. Lawrence Erlbaum Associates, Hillsdale, NJ.

- [Chambers and Jurafsky, 2008] Chambers, N. and Jurafsky, D. (2008). Unsupervised learning of narrative event chains. In *Proceedings of ACL-08: HLT*, pages 789–797, Columbus, Ohio. Association for Computational Linguistics.
- [Chambers and Jurafsky, 2009] Chambers, N. and Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 602–610, Suntec, Singapore. Association for Computational Linguistics.
- [Charniak and McDermott, 1985] Charniak, E. and McDermott, D. (1985). *Introduction to artificial intelligence*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- [Chatman, S., 1986] Chatman, S. (1986). *Story and Discourse: Narrative Structure in Fiction and Film*, volume Second Edition. Cornell University Press, USA.
- [Crawford, 2005] Crawford, C. (2005). *On Interactive Storytelling - New Riders: Game Design and Development*. Peachpit Press, Berkeley, California, first edition.
- [Cullingford, 1981] Cullingford, R. E. (1981). *Inside Computer Understanding: Five Programs Plus Miniatures*, chapter SAM. Lawrence Erlbaum Associates.
- [CyCorp, 2010] CyCorp (2010). Opencyc. <http://www.opencyc.org/>.
- [Dehn, 1981] Dehn, N. (1981). Story Generation After Tale-Spin. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 16–18.
- [Elhadad and Robin, 1996] Elhadad, M. and Robin, J. (1996). An overview of surge: a reusable comprehensive syntactic realization component. Technical report, Ben-Gurion University of the Negev.
- [Field, 1998] Field, S. (1998). *Screenplay: The Foundations of Screenwriting*. Amazon.com, third edition.
- [Finlayson, 2009] Finlayson, M. (2009). Deriving narrative morphologies via analogical story merging. *New Frontiers in Analogy Research*, pages 127–136.

- [Foster, 1941] Foster, E. M. (1941). *Aspects of the Novel*. Edward Arnold, London.
- [Gatt and Reiter, 2009] Gatt, A. and Reiter, E. (2009). SimpleNLG: A realisation engine for practical applications. In *ENLG-2009*.
- [Genette, 1966] Genette, G. (1966). *Figures I*. Seuil.
- [Genette, 1969] Genette, G. (1969). *Figures II*. Seuil.
- [Genette, 1972] Genette, G. (1972). *Figures III*. Seuil.
- [Genette, 1979] Genette, G. (1979). *Narrative Discourse: An Essay in Method*. Cornell University Press.
- [Gervás, 2007] Gervás, P. (June, 2007). Tap: a text arranging pipeline. Technical report, Natural Interaction based on Language Group, Universidad Complutense de Madrid, Spain.
- [Gervás and León, 2010] Gervás, P. and León, C. (2010). Story generation driven by system-modified evaluation validated by human judges. In *First International Conference on Computational Creativity*, Lisboa, Portugal.
- [Graesser et al., 1994] Graesser, A. C., Singer, M., and Trabasso, T. (1994). Constructing Inferences During Narrative Text comprehension. *Psychological Review*, (101):371–395.
- [Hassan et al., 2007a] Hassan, S., León, C., Gervás, P., and Hervás, R. (2007a). A computer model that generates biography-like narratives. In *International Joint Workshop on Computational Creativity*, London.
- [Hassan et al., 2007b] Hassan, S., Pavón, J., Arroyo, M., and León, C. (2007b). Agent based simulation framework for quantitative and qualitative social research: Statistics and natural language generation. In *The Fourth European Social Simulation Association Conference*, Toulouse, France. F. Amblard, F. Amblard.
- [Herman, 2000] Herman, D. (2000). Narratology as a cognitive science. *Image and Narrative*.
- [Jones and Galliers, 1996] Jones, K. S. and Galliers, J. R. (1996). *Evaluating Natural Language Processing Systems: An Analysis and Review*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

- [Joyce et al., 1989] Joyce, M., Kaplan, N., McDaid, J., and Moulthrop, S. (1989). Hypertext, narrative, and consciousness. In *Hypertext'89 Proceedings, November 5-8, 1989, Pittsburgh, Pennsylvania, USA*, pages 383–384. ACM.
- [Kelly, 1955] Kelly, G. (1955). *The Psychology of Personal Constructs*, volume I,II. Norton, New York.
- [Klein et al., 1973] Klein, S., Aeschliman, J. F., Balsiger, D., Converse, S. L., Court, C., Foster, M., Lao, R., Oakley, J. D., and Smith, J. (1973). Automatic novel writing: A status report. Technical Report 186, Computer Science Department, The University of Wisconsin, Madison, Wisconsin.
- [Kolodner, 1980] Kolodner, J. L. (1980). *Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model*. PhD thesis, University of Yale, New Haven, CT, USA.
- [Ladkin, 1987] Ladkin, P. B. (1987). The logic of time representation.
- [Lang, 1999] Lang, R. R. (1999). A declarative model for simple narratives. In *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*, pages 134–141. AAAI Press.
- [Lebowitz, 1983] Lebowitz, M. (1983). Storytelling as planning and learning. In *International Joint Conference on Artificial Intelligence*.
- [Lebowitz, 1985] Lebowitz, M. (1985). Storytelling as Planning and Learning. *Poetics*, 14:483–502.
- [Lee, M., 1994] Lee, M. (1994). A Model for Story Generation. Master's thesis, University of Manchester.
- [León and Gervás, 2010] León, C. and Gervás, P. (2010). The Role of Evaluation-Driven rejection in the Successful Exploration of a Conceptual Space of Stories. *Minds and Machines*, 20(4):615–634.
- [León et al., 2008] León, C., Peinado, F., and Navarro, A. (2008). An intelligent plot-centric interface for mastering computer role-playing games. In *ICIDS08, 1st International Conference on Interactive Digital Storytelling*.
- [Lerusalimschy et al., 2006] Lerusalimschy, R., Figueiredo, L. H. d., and Celes, W. (2006). *Lua 5.1 Reference Manual*. Lua.Org.

- [Leslie Kaelbling, 1994] Leslie Kaelbling (1994). Associative Reinforcement Learning: A Generate and Test Algorithm. In *Machine Learning*.
- [León and Gervás, 2008] León, C. and Gervás, P. (2008). Creative storytelling based on transformation of generation rules. In *5th International Joint Workshop on Computational Creativity*.
- [León et al., 2007a] León, C., Hassan, S., and Gervás, P. (2007a). From the event log of a social simulation to narrative discourse: Content planning in story generation. In Olivier, P. and Kray, C., editors, *Conference of the Artificial and Ambient Intelligence*, page 402–409, Culture Lab, Newcastle University, Newcastle upon Tyne, UK.
- [León et al., 2007b] León, C., Hassan, S., Gervás, P., and Pavón, J. (2007b). Mixed narrative and dialog content planning based on bdi agents. In *XII Conferencia de la Asociación Española para Inteligencia Artificial*, Salamanca, Spain.
- [Liu and Singh, 2004] Liu, H. and Singh, P. (2004). ConceptNet: A Practical Commonsense Reasoning Toolkit. *BT Technology Journal*, 22.
- [Lytinen, 1992] Lytinen, S. L. (1992). Conceptual Dependency and its Descendants. *Computers and Mathematics with Applications*, 23(2–5):51–73.
- [Mancini, 2000] Mancini, C. (2000). From cinematographic to hypertext narrative. In *HYPERTEXT '00: Proceedings of the eleventh ACM on Hypertext and hypermedia*, pages 236–237, New York, NY, USA. ACM.
- [Mani et al., 2006] Mani, I., Verhagen, M., Wellner, B., Lee, C. M., and Pustejovsky, J. (2006). Machine learning of temporal relations. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, ACL-44, pages 753–760, Morristown, NJ, USA. Association for Computational Linguistics.
- [Mateas and Sengers, 1999] Mateas, M. and Sengers, P. (1999). Narrative Intelligence: An Introduction to the NI Symposium. In *Working Notes of the Narrative Intelligence Symposium*, pages 1–10.
- [Mateas and Stern, 2005] Mateas, M. and Stern, A. (2005). Structuring content in the Façade interactive drama architecture. In *Proceedings of AIIDE*, pages 93–98.

- [McIntyre and Lapata, 2009] McIntyre, N. and Lapata, M. (2009). Learning to tell tales: a data-driven approach to story generation. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1*, ACL-IJCNLP '09, pages 217–225, Morristown, NJ, USA. Association for Computational Linguistics.
- [McKee, 1997] McKee, R. (1997). *Story. Substance, Structure, Style and the Principles of Screenwriting*. Regan Books, New York.
- [Meehan, 1976] Meehan, J. (1976). *The Metanovel: Writing Stories by Computer*. PhD thesis, Yale University.
- [Meehan, 1981] Meehan, J. R. (1981). *Inside Computer Understanding: Five Programs Plus Miniatures*, chapter Tale-Spin and Micro Tale-Spin. Lawrence Erlbaum Associates.
- [Michie, Donald, 1968] Michie, Donald (1968). Memo Functions and Machine Learning. *Nature*, 218:19–22.
- [Mueller, 1987] Mueller, E. T. (1987). *Daydreaming and Computation: A Computer Model of Everyday Creativity, Learning and Emotions in the Human Stream of Thought*. PhD thesis, UCLA, Computer Science Department.
- [Murray, 1997] Murray, J. H. (1997). *Hamlet on the Holodeck. The Future of Narrative in Cyberspace*. MIT Press, Cambridge, MA.
- [Neil and Simborowski, 1993] Neil, P. and Simborowski, N. (1993). *The Complete Fairy Tales of Charles Perrault*. Houghton Mifflin Harcourt.
- [Niehaus and Young, 2009] Niehaus, J. and Young, R. M. (2009). A Computational Model of Inferencing in Narrative. In *Intelligent Narrative Technologies II*. AAAI Press.
- [Peinado et al., 2008] Peinado, F., Navarro, A., and Gervás, P. (2008). A Testbed Environment for Interactive Storytellers. In *International Conference on Intelligent Technologies for Interactive Entertainment*. ACM Digital Library.
- [Pérez y Pérez, 1999] Pérez y Pérez, R. (1999). *MEXICA: A Computer Model of Creativity in Writing*. PhD thesis, The University of Sussex.
- [Propp, 1928] Propp, V. (1928). *Morphology of the Folk Tale*. TX: University of Texas.

- [Pérez y Pérez et al., 2007] Pérez y Pérez, P., Sosa, R., and Lemâitre, C. (2007). A Computer Model for Visual-Daydreaming. In *AAAI 2007 Fall Symposium on Intelligent Narrative Technologies*. AAAI.
- [R Development Core Team, 2010] R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- [Real Academia Española, 2010] Real Academia Española (2010). *Diccionario de la Real Academia Española*.
- [Reiter and Dale, 2000] Reiter, E. and Dale, R. (2000). *Building Natural Language Generation Systems*. Cambridge University Press.
- [Riedl, 2004] Riedl, M. (2004). *Narrative Planning: Balancing Plot and Character*. PhD thesis, Department of Computer Science, North Carolina State University.
- [Riedl and León, 2008] Riedl, M. and León, C. (2008). Toward vignette-based story generation for drama management systems. In *INTETAIN, Workshop on Integrating Technologies for Interactive Stories*.
- [Riedl and León, 2009] Riedl, M. and León, C. (2009). Generating story analogues. In *Proceedings of the 5th Artificial Intelligence and Interactive Digital Entertainment*.
- [Riedl and Young, 2006] Riedl, M. and Young, R. (2006). Story Planning as Exploratory Creativity: Techniques for Expanding the Narrative Search Space. *New Generation Computing, Computational Paradigms and Computational Intelligence. Special Issue: Computational Creativity*, 24(3):303–323.
- [Riedl and Sugandh, 2008] Riedl, M. O. and Sugandh, N. (2008). Story planning with vignettes: Toward overcoming the content production bottleneck. In *Proceedings of the 1st Joint International Conference on Interactive Digital Storytelling (Formerly TIDSE and ICVS)*.
- [Ritchie, 2008] Ritchie, G. (2008). Uninformed Resource Creation for Humour Simulation. In *Proceedings of the 5th International Joint Workshop on Computational Creativity*, pages 147–151, Madrid.
- [Rowe et al., 2009] Rowe, J. P., McQuiggan, S. W., Robison, J. L., Marcey, D. R., and Lester, J. C. (2009). Storyeval: An empirical evaluation framework for narrative generation. In *AAAI Spring Symposium*.

- [Rumelhart, 1975] Rumelhart, D. E. (1975). Notes on a schema for stories. *Representation and Understanding: Studies in Cognitive Science*, pages 211–236.
- [Schank and Abelson, 1977] Schank, R. and Abelson, R. (1977). *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. L. Erlbaum, Hillsdale, NJ.
- [Schank, 1982] Schank, R. C. (1982). *Dynamic Memory : A Theory of Reminding and Learning in Computers and People*. Cambridge University Press.
- [Sharples, 1996] Sharples, M. (1996). An account of writing as creative design. *The Science of Writing*.
- [Sharples, 1999] Sharples, M. (1999). *How We Write*. Routledge.
- [Swartjes and Theune, 2006] Swartjes, I. and Theune, M. (2006). A Fabula Model for Emergent Narrative. In S. Gbel, R. M. and Iurgel, I., editors, *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment, Third International Conference, TIDSE 2006*, number 4326 in Lecture Notes in Computer Science, pages 49–60, Heidelberg.
- [The LaTeX Project, 2010] The LaTeX Project (2010). \LaTeX - a document preparation system.
- [The Project Gutenberg Team, 2010] The Project Gutenberg Team (2010). Project Gutenberg.
- [Theune et al., 2003] Theune, M., Faas, E., Nijholt, A., and Heylen, D. (2003). The virtual storyteller: Story creation by intelligent agents. In *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment*, pages 204–215.
- [Thomas and Chad Fowler, 2005] Thomas, D. and Chad Fowler, A. H. (2005). *Programming Ruby: The Pragmatic Programmers' Guide*. Pragmatic Bookshelf, Raleigh, NC, 2. edition.
- [Todorov, 1977] Todorov, T. (1977). *The Grammar of Narrative*, pages 108–119. The Poetics of Prose. Ithaca: Cornell UP.
- [Trabasso and Sperry, 1985] Trabasso, T. and Sperry, L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, 24:595–611.

- [Turner, 1992] Turner, S. (1992). *MINSTREL: A Computer Model of Creativity and Storytelling*. PhD thesis, University of California at Los Angeles, Los Angeles, CA, USA.
- [van den Broek, 1988] van den Broek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language*, 27:1-22.
- [Viktor Shklovsky, 1917] Viktor Shklovsky (1917). *Art as Technique*. University of Nebraska Press.
- [Vilain et al., 1986] Vilain, M., Kautz, H., and Beek, P. (1986). Constraint propagation algorithms for temporal reasoning. In *Readings in Qualitative Reasoning about Physical Systems*, pages 377-382. Morgan Kaufmann.
- [Vinet and Griffin, 2010] Vinet, J. and Griffin, A. (May, 2010). Arch Linux. <http://www.archlinux.org/>.
- [Weyhrauch, 1997] Weyhrauch, P. (1997). *Guiding interactive drama*. PhD thesis, Computer Science, Carnegie Mellon Univ., Pittsburgh, PA.
- [Wielemaker, 2010] Wielemaker, J. (2010). SWI-Prolog. <http://www.swi-prolog.org/>.
- [Wiggins, 2006] Wiggins, G. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7).
- [Wilensky, 1981] Wilensky, R. (1981). *Inside Comptuer Understanding: Five Programs Plus Miniatures*, chapter PAM, pages 136-179. Lawrence Erlabaum Associates.
- [Wu, 1992] Wu, X. (1992). Inductive learning: Algorithms and frontiers. *Artificial Intelligence Review*, 6.

