

Meta-level Evaluation and Transformational Creativity; An analysis of MEXICA

Juan Alvarado¹ and Geraint A. Wiggins^{2,1}

¹ School of Electronic Engineering and Computer Science, Queen Mary University of London, United Kingdom

² Department of Computer Science, Vrije Universiteit Brussel, Belgium
{j.alvaradolopez, geraint.wiggins}@qmul.ac.uk, Geraint.Wiggins@vub.be

Abstract

We present a proposal for meta-level evaluation in creative systems. In this paper, the Mexica creative system is the subject of our analysis. To achieve meta-level evaluation, we have constructed a program that interacts with Mexica's inputs and its outputs. The meta-level system evaluates the results and makes decisions based on them. For decision-making it includes heuristics that are defined in terms of Mexica's results; these are rules particular to the object-level system. We also include other general metrics that apply in any other creative system. For example, Ritchie's criteria for evaluating creative systems are included in this proposal.

Introduction

Creative systems have been developed for various purposes and fields. We can find them for the generation of stories, melodies, paintings, poems, etc. There are also creative systems for non-artistic domains because creative behaviours are important here too.

There are many things to analyse in creative systems, for example, the outputs they produce, the rules, the input examples and the evaluation methods they use, etc. An important aspect of the analysis is the evaluation of the work of these systems. One way to approach the evaluation of a system is in terms of the output it produces (Ritchie 2007), we can call this an *Output Evaluation*. Another way to evaluate it is to analyse the process that the system follows (Colton, Pease, and Charnley 2011; Jordanous 2012), we can call this a *Process Evaluation*. We could also consider a combination of both types of evaluation where we analyse both the output and the process, we can call it a *Process-Output Evaluation*. Humans can evaluate the creativity of systems, and creative systems can evaluate their performance themselves (Human Evaluation, System Evaluation). There also could be a meta-evaluation; an evaluation of their evaluation (Jordanous 2011).

There is another way to evaluate creative systems, which is to go to the meta-level. In the meta-level there are rules, but they do not need to be the same as at the object level. As an example, take this paper. It was generated by an actor (one/many authors, a system, etc.). The paper is the output. The actor has rules to generate a conceptual space, find novel concepts and evaluate them. This is the object level.

In the meta-level, a reader looking at this paper might have a distinct set of rules and different knowledge. Therefore, the evaluation of the paper might be different as this meta-level evaluation might not share the evaluation rules in the object level. Following the terms suggested before, assuming the reader is a human being, this would be a *Human Meta-Level Output Evaluation*.

In this paper, we take the creative system Mexica as an example and explore its meta-level output evaluation. To achieve this, we have built a new system that interacts with Mexica. The meta-level system interacts with the graphical interface of Mexica and the outputs it generates and changes the parameters Mexica uses to produce an output. Therefore, we are talking about a *System Meta-Level Output Evaluation*.

To achieve meta-level evaluation, we have to consider the distinct parts of Mexica. For this we have used the ideas on creative systems expressed by Boden (1990) and the Creative Systems Framework (CSF) proposed by Wiggins (2006). With these tools, we can formalise the elements of Mexica in a model that identifies its components and how they interact.

We also use the criteria proposed by Ritchie (2007) for evaluating creative systems. These criteria can give us information about the typicality and quality of stories generated by Mexica. Using this evaluation along with the information from the files Mexica generates, we can know, for example, the distribution of stories in the groups proposed by these criteria, the characteristics these stories have, the result of the self-evaluation for each story and the set of parameters used to generate them. This information allows us to build a meta-level system that admits transformational creativity.

This paper is about how to do reflection in a creative system as suggested by Buchanan (2001), structured as transformational creativity in Boden (1990) terms. One of our reviewers summed up the intent of this paper, extremely well; "The result is a meditation on computational creativity, transformational creativity and conceptual spaces that is relevant to the topic."

Background

We have brought together the following ideas in this work. These are important because we use them to distinguish the distinct parts of the creative systems and how they interact.

This way, we are in a better position to perform meta-level evaluation and later modification in Mexica and thus achieve transformational creativity.

Conceptual Spaces and Transformational Creativity

Boden (1990) has expressed various ideas about creative systems. One of the key components of her ideas is conceptual spaces. They contain the creative ideas that systems can find. Conceptual spaces, as expressed by Boden, are spaces delimited by accepted rules in a social group. They define the space of solutions of a creative system.

Boden (1990) also points out that there are distinct types of creativity in systems. One of the important types of creativity that she points out is transformational creativity. For Boden, this type of creativity is characterised by changing the rules that define the conceptual space (Boden 1990). This means that the accepted rules in a social group have changed and therefore the concepts that can be found in a conceptual space have also changed. This could mean the existence of more, fewer, better concepts. In any case, it means a change in the conceptual space and a change in the concepts or ideas available to a particular system.

Wiggins (2006) adds to the proposal of Boden (1990) that we can have transformational creativity by changing not only the conceptual space but also the rules a system uses to traverse it. This way we are also making different concepts available. They could be more, fewer, better, etc. but the important thing is that there has been a transformation.

Mexica

Mexica (Pérez y Pérez 1999) is a creative system that generates short stories. Mexica does not have pre-defined goals. It uses an Engagement and Reflection cycle (Pérez y Pérez and Sharples 2001; Sharples 1996). Content generation is guided by a set of constraints and previous stories available.

Mexica includes different rules for Engagement and Reflection. Alvarado and Wiggins (2018a) point out that in the original Sharples (1996) account each phase produces different results because conceptual spaces are not the same. Similarly, Alvarado and Wiggins (2018b) point out that Mexica produces a story through the interaction of both phases in which conceptual spaces are not the same either. In one phase, the system generates a type of content which, passing to another phase, is restricted and changed.

In Reflection, the system evaluates the story in progress. The evaluation takes into account the parameters provided by the user. This can lead to the system changing parameters to improve the story. For example, if the story is getting boring or if the evaluation determines that the story in progress is very similar to one/some previous stories in the database.

The system delivers as a final output a story and several files that show part of the process that Mexica followed to generate the story. In these files, the system shows the parameters it has used, the actions it has added to the story and the results of the self-evaluations.

Creative Systems Framework

Wiggins (2006) has proposed a framework for creative systems. In this framework, he proposes a formalisation of creative systems in such a way that we can identify different aspects in a creative system and the rules that operate on them. This model takes into account ideas of Boden (1990) for conceptual spaces and formalises them. Wiggins (2006) highlights the rules systems use and the concepts that can be found using them in conceptual spaces.

In this framework, Wiggins (2006) points out that there are rules that a system can use to define its conceptual space. He takes very much into account the process that the system follows for the generation of artefacts. He considers a set of rules that the system uses to explore the conceptual space. He also considers the use of a set of rules by which the system can determine the quality of the objects produced. Different systems apply different rules, and their number is variable.

Wiggins (2006) suggests an interpretation function that uses the mentioned rule sets to find creative objects. We can change rule sets to change the behaviour of a system. This results in many interesting behaviours. For example, rules that define the conceptual space can change and they can generate new conceptual spaces. The rules that serve to traverse the conceptual space can change, giving rise to having access to new concepts which were previously inaccessible. In principle, it would be possible to change the evaluation rules, allowing previously rejected ideas to be accepted.

Ritchie's Criteria

Ritchie (2007) has proposed some criteria for evaluating creative systems based on the quality of the output that they produce. These criteria are useful for finding the proportion of examples that fall into different important categories in creative systems analysis. For example, the percentage of new artefacts to the examples known to the system, or the percentage of artefacts that are high-quality and untypical.

In this paper, we also suggest the use of this set of criteria to feed the meta-level system with information on the evaluation of the products generated by Mexica. This way, the meta-level system can make decisions about the parameters that can be varied to change the behaviour (and outputs) of Mexica.

Alvarado and Wiggins (2019) propose a model that integrates the Engagement and Reflection account, with the Creative Systems Framework and Ritchie's criteria. Here the similarities between the CSF and Ritchie's criteria are highlighted. A particular example of the coincidences between the two proposals is the importance given to the generation of untypical but good quality examples, which could be seen as an achievement for creative systems. This is so because untypical examples generally correspond to solutions that are not found in the conceptual space and evaluation processes, taking into account only this, could show they are poor quality. Finding untypical and good quality artefacts implies that the evaluation rules have been able to find good quality artefacts despite not being in the conceptual space. Wiggins (2006) calls this aberration to highlight the negative

connotation with which artistic examples of good quality but not following pre-established rules have been received in the past.

Alvarado and Wiggins (2020) show the application of Ritchie's criteria to the output generated by Mexica. They highlight here the outputs' trends for Mexica. This can be useful when evaluating an instance of Mexica or deciding how to find a new instance.

Implementation

For the implementation of this project, a program that can interact with the Mexica's graphical interface and its output has been built. The purpose of this program is to change the input of Mexica based on a meta-level evaluation.

Object level

At the object level, we have Mexica and its components. These include inputs, outputs, and rules that it uses. As inputs, we have the initial action that the system uses as a starting point, the previous stories and a set of parameters that are defined in its user's interface. The outputs include the log files, the generated story and the set of all stories generated in previous runs of Mexica. The log files contain information on the process that the system has followed to generate the stories.

Alvarado and Wiggins (2018b) have analysed Mexica and identified distinct groups of rules for Engagement and Reflection. They argue that these different rules produce different results in each stage.

Before Mexica generates a story, it analyses all previous stories. It takes each action in each story and computes the context. Mexica creates structures with the generated contexts. It groups all actions in previous stories that share the same context into the same context structure.

When Mexica generates a story, it receives an initial action. Mexica computes the context for the initial action. Then, using this context, it looks for a similar one among the context structures. Mexica retrieves all the associated actions from those similar contexts. Those actions will be candidates to continue the story. It generates something similar to a tree with a variable number of branches to continue the story. It analyses each branch with other conditions that filter the candidates. After the filtering, Mexica selects one of the remaining candidates to continue the story in progress. Mexica computes the context of the story in progress with each new action added and repeats the process in each execution of Engagement.

Context similarity checking requires some parameters the user can set in the graphical user interface. The number of matches between the context structures (and potential candidates) depends on the value of these parameters. Candidates filtering also requires parameters the user can set in the graphical interface.

The lack of options to continue the story (i.e. there are no matches with context structures) can cause an impasse. The reflection stage has mechanisms to solve this problem, but if it cannot do so, the system gives up and the story is abandoned.

Interface

To interact with the object level, the meta-level system requires an interface we have built. This interface is a program capable of interacting with Mexica's graphical interface, changing parameters and starting its execution. This interface has access to Mexica's output files. It reads from the object level the previous stories, the log files, the generated story, the set of all generated stories and the set of parameters defined by the user. This interface communicates directly with the meta-level system. The output of this interface (which comes from the meta-level system), includes parameters and the initial action to adjust the future behaviour of Mexica.

Meta level

Wiggins (2006) argues that we can view transformational creativity as exploratory creativity at the meta-level. This way we can use the same components he uses in the analysis of creative systems in the meta-level.

In the meta-level system, some rules define the conceptual space. Here, the conceptual space corresponds to the space of all possible instances of Mexica. Mexica's user inputs constraint this space. The user in this case is the meta-level system. It only changes the parameters, not the previous knowledge.

A set of rules allows the meta-level system to traverse its conceptual space. They do so by varying Mexica's parameters. With this, they generate a new instance of Mexica. These rules allow decision-making with the information provided by the meta-level evaluation.

There are meta-level evaluation rules that can be used with instances of Mexica found in the meta-level conceptual space. The evaluation rules in the meta-level include heuristics that take into account the information in Mexica's log files (e.g. the result of the self-evaluations, or the number of actions retrieved in the first execution of Engagement). For example, they take into account the probability of success of a story in progress when there is an impasse in the first execution of Engagement, which is low.

Following this example, when the meta-level system detects an impasse in the first execution of Engagement, it raises a flag because of the meta-level evaluation. Then using the traversing rules, this evaluation and the current set of parameters, the meta-level system makes a decision. As a result, the meta-level system establishes a new value for the similarity of contexts parameter, so Mexica can get more options. With this, using Mexica's graphical interface, the interface sets the new set of parameters and a new initial action and runs Mexica. This way, in this example, Mexica can create a story, avoiding the impasse because now it has more actions (or different) to choose from. This modification considers the self-evaluation of Mexica and the meta-level evaluation to produce a story that gets a better overall evaluation. The purpose of the meta-level system is to make the evaluation at the object level be satisfied more and more often.

Alvarado and Wiggins (2020) show the result of applying criteria of Ritchie (2007). They provide performance-related

information on past Mexica runs. The results show that there are not untypical and good quality examples produced by Mexica. We can include this result in the meta-level evaluation and traversing strategy to make a more informed decision on how to change parameters to get better stories, but more than that, explore the possibility of finding untypical and good quality concepts.

Discussion

We have shown one way that we could meta-level evaluate creative systems. Here, we have used Mexica, a creative system that produces stories. We can think of many examples of creative systems. They follow different processes and approaches to be developed and use different rules. What we present here is an idea and a way of doing meta-level evaluation in a particular example.

Pérez y Pérez (1999) points out that changing Mexica's parameters offers great flexibility to experiment with the creative process. What the meta-level system does is precisely this; it interacts with Mexica, examines the output, evaluates at the meta-level, decides and establishes new parameters for Mexica, and it runs Mexica again. It interacts as a user so it cannot change the inner workings of Mexica.

In this paper, we propose to use the criteria of Ritchie (2007) to determine Mexica's performance. The criteria form a good tool that gives information regarding the outputs that Mexica generates. While their first intention is to be used by humans to evaluate the outputs of the systems, Ritchie (2012) points out that it is possible to use them as internal components of the system and not external judgements. This means that we can apply these criteria in distinct ways: Internal: How does the system work, on its own terms? External: How does the system work, in terms of independent measures such as human judgements? Considering the internal way, we can do what we report in this paper. It should be possible to use these criteria as part of a system that evaluates another system. There is still the external part in which we should include human intervention to evaluate the meta-level system and its meta-level evaluation.

Ritchie's criteria do not have a particular field of application, which makes them a generic part of this meta-level system. Perhaps it is necessary to include other general approaches/methodologies for the evaluation (e.g. Colton, Pease, and Charnley 2011; Jordanous 2012).

There are other elements of this meta-level evaluation that are not general. For example, they largely depend on the results that Mexica delivers and on the results of its self-evaluations. The same happens with other creative systems. Further refinement is necessary to incorporate general ideas in creative (meta-level) (meta) evaluation to improve this proposal.

Acknowledgements

We would like to thank the reviewers for their comments which helped improve this paper. Thanks to the National Council of Science and Technology (CONACYT) in México for its sponsorship. The research was also supported by

funding from the Flemish Government under the "Onderzoeksprogramma Artificiële (AI) Vlaanderen" programme.

References

- Alvarado, J., and Wiggins, G. A. 2018a. Exploring the Engagement and Reflection Model with the Creative Systems Framework. In François Pachet, A. J., and León, C., eds., *Proceedings of the Ninth International Conference on Computational Creativity (ICCC'18)*, 200–207. Salamanca, Spain: Association for Computational Creativity.
- Alvarado, J., and Wiggins, G. A. 2018b. Understanding MEXICA: an analysis of an Engagement-Reflection system. In *5th AISB Symposium on Computational Creativity*.
- Alvarado, J., and Wiggins, G. A. 2019. Creative Systems Framework, Empirical Criteria and Engagement and Reflection: A unified framework. In *6th AISB Symposium on Computational Creativity*.
- Alvarado, J., and Wiggins, G. A. 2020. Assessing Creativity of MEXICA: An Application of Ritchie's Criteria. In *7th AISB Symposium on Computational Creativity*.
- Boden, M. A. 1990. *The Creative Mind: Myths and Mechanisms*. Abacus.
- Buchanan, B. G. 2001. Creativity at the metalevel AAAI-2000 presidential address. *AI Magazine* 22(3):13–28.
- Colton, S.; Pease, A.; and Charnley, J. 2011. Computational creativity theory: The FACE and IDEA descriptive models. In *Second International Conference on Computational Creativity*, 90–95.
- Jordanous, A. 2011. Evaluating Evaluation: Assessing Progress in Computational Creativity. *International Conference on Computational Creativity* 102–107.
- Jordanous, A. 2012. A Standardised Procedure for Evaluating Creative Systems: Computational Creativity Evaluation Based on What it is to be Creative. *Cognitive Computation*.
- Pérez y Pérez, R., and Sharples, M. 2001. Mexica: A computer model of a cognitive account of creative writing. *Journal of Experimental and Theoretical Artificial Intelligence* 13(2):119–139.
- Pérez y Pérez, R. 1999. *MEXICA : A Computer Model of Creativity in Writing*. Ph.D. Dissertation, The University of Sussex.
- Ritchie, G. 2007. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines* 17(1):67–99.
- Ritchie, G. 2012. A closer look at creativity as search. In *International Conference on Computational Creativity*, 41–48.
- Sharples, M. 1996. An account of writing as creative design. *The science of writing: Theories, methods, individual differences, and applications*. (January):127–148.
- Wiggins, G. A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.