When Reflective Feedback Triggers Goal Revision: a Computational Model for Literary Creativity*

Pablo Gervás
Instituto de Tecnología del Conocimiento
Universidad Complutense de Madrid
28040 Madrid, Spain
pgervas@sip.ucm.es

Carlos León
Facultad de Informática
Universidad Complutense de Madrid
28040 Madrid, Spain
cleon@fdi.ucm.es

Abstract

Existing models of the writing task from a cognitive viewpoint agree on the importance of draft revision in the overall process. This is generally assumed to focus on reviewing intermediate drafts in search for feedback on how to modify them to match the driving constraints. However, in literary creativity it is often the case that the feedback leads not to a revision of the current draft but to a redefinition of the constraints that are driving the process. This phenomenon is explicitly described in Sharples' model of writing as a creative task. Yet existing computational models of literary creativity do not contemplate it. The present paper describes a computational model of the creative processes in literary creativity that contemplates the explicit representation of the constraints driving the process, and allows for the feedback from the validation to modify not just the ongoing draft but also the constraints that it is expected to satisfy. This allows the model to represent cases of serendipitous discovery of interesting features.

1 Introduction

Creative processes as carried out by humans are known to involve a significant amount of trial and error. Writers, musicians, painters, poets... rely on a succession of drafts that get polished over many iterations, each one involving feedback from the previous version, and resulting from a process of revision or regeneration of it. Yet computational models developed in AI over the years to emulate these same processes very rarely capture this type of dynamic operation. Sometimes they do, in a limited fashion, when an AI program includes an evaluation function that defines the desired form for outcomes, and this is run over the results of a generative process that produces candidate artefacts of the desired type. However, the dynamics of the creative process in humans are known to be more complex than this, possibly differing significantly across different domains.

*This paper has been partially supported by the project WHIM 611560 funded by the European Commission, Framework Program 7, the ICT theme, and the Future Emerging Technologies FET program.

The present paper focuses on literary creativity, and proposes a computational model for the creative process in this domain based on a number of cognitive model of the task of writing as carried out by humans.

2 Previous Work

The present paper puts forward a proposal that captures in computational terms the operations described in two existing cognitive models of the writing tasks. This section reviews these two models and two competing computational models also based them.

2.1 Cognitive Models of the Writing Task

Flower and Hayes [Flower and Hayes, 1981] define a cognitive model of writing in terms of three basic process: planning, translating these ideas into text, and reviewing the result with a view to improving it. These three processes are said to operate interactively, guided by a monitor that activates one or the other as needed. The planning process involves generating ideas, but also setting goals that can later be taken into account by all the other processes. The translating process involves putting ideas into words, and implies dealing with the restrictions and resources presented by the language to be employed. The reviewing process involves evaluating the text produced so far and revising it in accordance to the result of the evaluation. Flower and Hayes' model is oriented towards models of communicative composition (such as writing essays or functional texts), and it has little to say about literary creativity in particular. Nevertheless, a computational model of literary creativity would be better if it can be understood in terms compatible with this cognitive model. An important feature to be considered is that the complete model is framed by what Flower and Hayes consider "the rhetorical problem", constituted by the rhetorical situation, the audience and the writer's goals.

Sharples [Sharples, 1996] presents a description of writing understood as a problem-solving process where the writer is both a creative thinker and a designer of text. For Sharples, the universe of concepts to be explored in the domain of writing could be established in a generative way by exhaustively applying the rules of grammar that define the set of well-formed sentences. The conceptual space on which a writer operates is a subset of this universe identified by a set of constraints which define what is appropriate to the task at hand.

Sharples explains that the use of a conceptual space "eases the burden of writing by limiting the scope of search through long term memory to those concepts and schemas that are appropriate to the task" [Sharples, 1996, p. 3]. To Sharples, the imposition of these constraints enables creativity in the sense that he identifies creativity in writing (in contrast with simple novelty) with the application of processes that manipulate these constraints, thereby exploring and transforming the conceptual space that they define. Sharples provides a specification of what he envisages these constraints to be. Constraints on the writing task are described as "a combination of the given task, external resources, and the writer's knowledge and experience" [Sharples, 1996, p. 1]. He also mentions they can be external (essay topic, previously written material, a set of publishers guidelines...) or internal (schemas, inter-related concepts, genres, and knowledge of language that form the writer's conceptual spaces).

Sharples also provides a description of how the typical writer alternates between the simple task of exploring the conceptual space defined by a given set of constraints and the more complex task of modifying such constraints to transform the conceptual space. Sharples proposes a cyclic process moving through two different phases: engagement and reflection. During the engagement phase the constraints are taken as given and the conceptual space defined by them is simply explored, progressively generating new material. During the reflection phase, the generated material is revised and constraints may be transformed as a result of this revision. Sharples also provides a model of how the reflection phase may be analysed in terms of specific operations on the various elements. A three step process of reviewing, contemplating and planning the result is suggested as a description of the reflection phase. During reviewing the result is read, minor edits may be carried out, but most important it is interpreted to represent "the procedures enacted during composition as explicit knowledge which can then be integrated with an existing conceptual space". Contemplation involves the process of operating on the results of this interpretation, which are likely to be explicit representations of specific constraints. Planning uses the results of contemplation to create plans or intentions to guide the next phase of engagement.

Sharples also provides an account of how the explicit representation of constraints as elements susceptible of modification is fundamental to achieve this type of cyclic operation. People produce grammatically correct linguistic utterances without being aware of the rules of grammar, but to explore and transform conceptual spaces one needs to call up constraints and schemas as explicit entities, and work on them in a deliberate fashion. For the mind to be able to manipulate the constraints, they have to be subjected to a process of "representational redescription" [Karmiloff-Smith, 1995], re-representing knowledge that was previously embedded in effective procedures as elements susceptible of manipulation.

The problem is that beginners addressing such a cognitive task do not have a vocabulary to describe mental processes to themselves. To learn, they must develop "a coherent mental framework of plans, operators, genres and text types that can guide the process of knowledge integration and transformation" [Sharples, 1996, p. 5]. Experts tend to have such a mental framework that underlies and supports their writing efforts. For beginners, the problem must be addressed with the aid of general knowledge about how to design artefacts, how to transform mental structures and how to solve problems. Because this is difficult to do in the head, some writers resort to capturing the ideas involved in paper, as sketches, lists, plans, notes etc. These external representations stand for mental structures, and they are easier to manipulate. The writer can then explore different ways of structuring the content, apply systematic transformations, establish priorities, and reorder or cluster items. The task of writing addressed in these terms is much closer to recognised design tasks.

The arguments outlined above with respect to how Sharples models the differences between beginners and experts suggests further consideration of the role of the evolution of representation in the progressive acquisition of expertise. In this respect, Karmiloff-Smith [Karmiloff-Smith, 1995] proposes a model of evolving representation called Representational Redescription model.

This model analyses the development of behavioural mastery in a given domain – meaning consistenly successful performance in the domain – in terms of how knowledge about the domain is represented internally by the individual. The model considers three phases of learning. During the first phase the individual focuses on his interaction with the environment, and represents these in the form of raw data received from outside. This may lead to an initial achievement of behavioural mastery. Over the second phase, internal representations are abstracted from the raw data, and processing may start to focus on them. As a result of this introspection, features of the environment may temporarily be disregarded and, as a result, observed behaviour may deteriorate. However, this leads to a recuperation of a more flexible achievement of behavioural mastery, by then based on having achieved reconciliation between internal representation and external data.

This model describes four different levels of cognitive representation: *implicit*, focused on the process itself; *explicit level one* in which basic aggregation of raw data present in the implicit level is performed in terms of data storage but may not yet be accessible to the cognitive system for manipulation operations; *explicit level two*, in which structures from the first explicit level are converted into schemas and thereby become available; and *explicit level three*, a final and "cross-system" representation of concepts that can be verbalized and are fully integrated in a more general cognitive system.

2.2 Existing Implementations of Sharples' Model

MEXICA [Pérez y Pérez, 1999] was a computer model designed to study the creative process in writing in terms of the cycle of engagement and reflection [Sharples, 1999]. It was designed to generate short stories about the MEXICAS (also wrongly known as Aztecs). MEXICA is a flexible tool where the user can set the value of different parameters to constrain the writing process and explore different aspects of story generation. It takes into account emotional links and tensions between the characters as means for driving and evaluating ongoing stories.

MEXICA relies on certain structures to represent its knowledge: a set of *story actions* (defined in terms of pre-

conditions and post-conditions) and a set of *previous stories* (stated in terms of story actions). MEXICA stands out from other systems in that it actually builds its own set of schemas from the set of previous stories. A single type of knowledge structure, known as a *Story-World Context (SWC)*, is used to represent these schemas. Story-World Contexts represent instances of contexts (described in terms of emotional links and tensions between existing characters) in which an action has appeared in a previous story, and they act like rules during the engagement phase: an action is added to the plot if a Story-World Context for that action can be found that matches the plot so far.[1] The reflection phase revises the plot so far, mainly checking it for coherence, novelty and interest. The checks for novelty and interest involve comparing the plot so far with that of previous stories. If the story is too similar to some previous one, or if its measure of interest compares badly to previous stories, the system takes action by setting a guideline to be obeyed during engagement. These guidelines can be considered as a basic implementation of Sharples' constraints, driving which types of action can be chosen from the set of possible candidates.

In MEXICA, the system is actually aware of the emotions of all the characters (and the emotional tensions between them) and uses these to drive and structure the story. But these emotions and tensions are often not mentioned in the final text of the story.

2.3 The ICTIVS Model

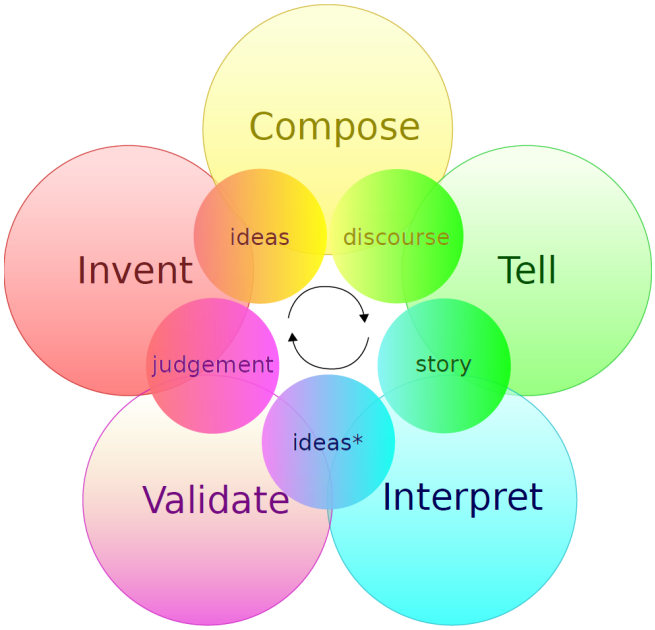

Figure 1: The original ICTIVS model. This version of the model does not take feedback into account.

The ICTIVS model [Gervás and León, 2014] arose as an initial attempt to construct a theoretical model based on an abstract analysis of the task of story construction in the context of a basic communication situation. Figure 1 shows a graphical representation of the ICTIVS model. The communication takes place as an exchange of a linear sequence of text that encodes a complex set of data that correspond to a set of events that take place over a volume of space time, possibly in simultaneous manner at more than one location. To convey this complexity as a linear sequence and recover it again at the other end of the communication process requires a process of condensing it first into a message and then expanding it again into a representation as close as possible to the original. There is a *composer*, in charge of composing a linear discourse from a conceptual source that may also have been produced by himself, and an *interpreter*, faced with the task of reconstructing a selected subset of the material in the conceptual source as an interpretation of the received narrative discourse. In real life, the role of the composer is usually played by a writer and the role of interpreter by a reader, but in the present case a more generic formulation has been preferred for generality.

This overarching act of communication is fundamental because it allows the definition of the purpose of the task in terms of the expected impact of the constructed story on the interpreter. Whatever is produced by the composer will have to be processed by the interpreter, and the impact on the interpreter cannot in truth be considered without taking into account what this process of interpretation involves. With this premise in mind, the ICTIVS model [Gervás and León, 2014] started from a linear description of the complete act of communication from an original purpose in the mind of the composer to a final impression in the mind of the interpreter. This act of communication involves processes of invention of a message and composition of an appropriate form meant to be carried out by the composer. It also involves processes of interpretation and validation carried out by the interpreter. From the point of view of the communicative act, the measure of success of such an act of needs to be established in terms of whether the interpretation by the interpreter matches the message constructed by the composer – success in terms of information transfer – and whether the impression in the mind of the interpreter matches the original purpose of the composer – success in terms of expected impact. As a first approximation, the impression in the mind of the interpreter could be correlated to the results of the validation applied to the message. In order to capture this intuition, the ICTIVS model defines the task of story composition as an iterative cycle of revisions in which the composer progressively generates drafts of his message, and then applies to them an internal process of interpretation and validation intended to match the one that the interpreter will be applying. At each iteration, the results of this estimated interpretation/validation are compared with the original purpose. If mismatches are detected, another cycle is started, and only when a successful match has been found does the resulting version of the message get communicated to the intended audience.

Five specific stages are included in the model: *INVENTION* – coming up with content for the narrative, possibly starting from scratch but often from some specification of

[1] It is important to note that Story-World Contexts (and not the definitions of action in terms of their pre-conditions) are used to find the next action to extend the plot.

purpose; a composer task –, *COMPOSITION* – establishing a form to express the desired content; a composer task –, *INTERPRETATION* – given a story, fill in the gaps, connect the dots, make assumptions on posible background implied, and extend it into a full picture of what the author wants you to "see in your mind"; an interpreter task that the composer needs to model to generate informative feedback for the construction process –, *VALIDATION* – identify the impact that the story, and/or the material interpreted from it, has on the interpreter; as above, an interpreter task but one that the composer needs to model to provide feedback –, and *TRANSMISSION* – passing over the result of the other processes to an audience; this stage establishes the link between the composer and the interpreter. Of these five stages, the first four may take place in an iterative cycle, and the final stage occurs only once after the iterations have lead to a successful draft with potential for achieving the expected impact on the interpreter according to the composer's purpose.

3 Reflective Feedback and Goal Revision in Computational Models of Literary Creativity

The computational models reviewed in section 2.2 and 2.3 capture the essence of the cognitive models described in section 2.1, but they both fail to capture the particular features that concern feedback and goal revision. The model of engagement and reflection in MEXICA has very limited explicit representation of the constraints driving the process, in the form of guidelines set during reflection. The ICTIVS model as originally described was formulated at a more abstract level, but focuses more on the constructive approach to the creative process, with no explicit modelling of the task of revising an already existing draft. It did include the representation of a seed idea or meaning that the composer wants to convey, but no representation of the possibility of this idea being modified as part of the creative process. A refinement on these models is required that can integrate a specification of the purpose for the generation task as an input, that can allow for revision of this specification as part of the process, and that at the same time can take advantage of the existing body of work on narrative generation.

Three relevant insights arise from the consideration of the original ICTIVS model in this enriched context of purpose-driven communication.

First, there will probably be a significant difference in computational terms between the initial iteration, where at each stage new material is generated from the corresponding input, and subsequent iterations, where two different processes may need to be employed: further generation of new material from the specification, and revision of the material generated in previous iterations - where the revision needs to be informed by the initial specification, the earlier drafts, and the identified mismatches. This is important because the computational mechanisms involved in each case may be different, and also because outputs from these two different processes may need to be combined into an integrated output for the corresponding stage.

Second, at the point of deciding whether a given draft is successful in terms of how it matches the original purpose, a truly creative process may consider not only revision of the draft but also revision of the purpose. This may arise whenever the estimated impact of a given draft on the interpreter is considered valuable by the composer beyond his original purpose. By means of this extension, the model can capture the role of serendipity in the creative process [Pease *et al.*, 2013; Corneli *et al.*, 2014].

Third, although computational models of the creative task are traditionally formulated as a cycle, in an ideal creative process cross-fertilization across the type of stages defined would be very positive. This is evident in Flower and Hayes description of the process as a set of transitions between three processes governed by an overall monitoring process that allocates effort to each one of them, and in Sharples' phrasing of his model as a dual cycle between two stages that operate on different data – the text and the constraints. A similar abstraction will need to be considered in our model.

The present section analyses these important concepts in more detail and attempts an initial formulation of such a refinement to take them into account in a manner that better reflects the intuitions arising from the cognitive models.

3.1 Analysing the Tasks Involved in Creative Production

In order to identify the core features that the desired model needs, the tasks involved in generation must be examined explicitly and compared to what the models can currently represent. Following this, we proposed a categorization of generative system according to their capabilities in terms of feedback.

Regarding their internal process, four types of systems can be identified:

- those that take no input and generate outputs determined exclusively by decisions taken during the construction of the system (*mere* generation)
- those that take as input some kind of specification that determines in some way the type of output that is to be obtained (*specification*)
- those that include a module that quantifies in some way the degree to which the outputs obtained satisfy the requirements specified as input (*diagnostic*)
- those that can benefit from the results of a diagnostic module to modify the specification and self correct their output (*reflective*)

Taking into account the kind of input that the systems accept, a parallel axis of classification may be whether the system can generate outputs only by constructing them from scratch (*construction*) or by applying transformations to an initial version of the desired artefacts (*revision*).

When these two axes are combined with the issues described previously, the following set of possible modes of operation arise:

- mere construction: the system generates outputs of a given form as determined by its construction
- construction to a specification: the system generates outputs conforming to a given specification

- construction with diagnostic: the system generates outputs and can provide some quantification of their quality
- reflective construction: the system generates outputs conforming to a given specification and can provide some quantification of their degree of satisfaction, and modify it accordingly.
- mere revision: the system receives an instance of the desired artefact and revises it towards a given goal determined by its construction
- revision to a specification: the system receives an instance of the desired artefact and a given specification and revises the instance of the desired artefact towards the given specification
- revision with diagnostic: the system receives an instance of the desired artefact, revises it towards a given goal determined by its construction, and can provide some quantification of the quality of the revised artefact
- reflective revision: the system receives an instance of the desired artefact and a given specification, revises the instance of the desired artefact towards the given specification, and can provide some quantification of their degree of satisfaction of the specification, and modify it accordingly.

3.2 Summarising the Features of a Creative Process from a Computational Point of View

The cognitive models reviewed in section 2.1 show a number of distinctive features that are relevant for the purpose of the present paper:

1. the creative process is iterative in nature
2. the creative process is driven by a set of constraints that restrict the desired outputs in some way; these constraints may be considered an input to the process
3. a cycle may involve processes of construction and/or processes of revision of prior results
4. at the end of each cycle a diagnostic procedure is applied to the result obtained so far
5. part of the diagnostic may involve quantifying degree of satisfaction of the given constraints
6. subsequent cycles take into account the diagnostic to attempt to improve the results of subsequent cycles
7. consideration of the diagnostic may take the form of planning further operations either on the artefact so far or on the set of constraints
8. the process as a whole includes a stage of meta-level reasoning which decides among the various available operational options applicable to the task at hand, such as, for instance, whether to iterate further or to stop, or, for a given iteration, whether to construct or to revise, whether to act upon the artefact itself or upon the set of constraints, or whether to apply the chosen operation to the whole element or to specific parts of it

3.3 Integrating the Reviewed Tasks and Features into a Computational Model

After having analysed both the tasks and the features involved in the creative process from a computational perspective, we propose the following three extensions for the refined model of the computations involved in literary creativity:

- to consider the explicit representation of constraints as part of the draft itself, so that they can be subject to the same operations as the rest of the draft
- to consider a range of operations that includes both construction and revision
- to consider the possibility of focusing system operation on particular subsets of the draft

The representation on which the creative process operates would therefore need to include at least two different parts:

- the set of constraints to be used to drive the construction process and/or to validate any resulting drafts, known as the *brief*
- the actual *draft* at each point of the creative process

Both the brief and the draft should be represented in such a way that different parts of them may be operated upon in isolation of the rest.

This representation that includes both a brief and a draft will be referred to henceforth as the *work in progress*. Any references to operations upon the work in progress can refer to both operations on the draft or on its specification.

The set of operations to consider would be:

- *reject*: eliminate from the work in progress a particular item for the next cycle
- *generate*: generate anew a particular item of the work in progress during the next cycle
- *revise*: modify a particular item of the work in progress during the next cycle
- *keep*: leave a particular item of the work in progress as it resulted from the previous cycle

Based on this terminology, a computational model for the tasks involved in literary creativity can now be rephrased at a lower level of detail. The same set of general steps can be seen, but each one of them now operates over a representation of the work in progress that includes both a brief and a draft, and at each stage the four types of operation (reject, generate, revise or keep) may be applied to any subset of the work in progress.

The computational model that we propose may now start from a hybrid representation of work in progress. Input may be provided to a creative system either in terms of a brief - a set of constraints that the output should satisfy – or a partial sample of the desired artefact, or as a combination of both modes.

This initial representation of the work in progress would undergo a process of reflection. In this initial reflection process, each of the sections of the work in progress is considered. If only a brief is available, the brief is marked as to be retained for the following construction cycle, and the empty

draft is tagged to be generated. If a brief and a partial draft are available, the partial draft is analysed in the light of the brief. The result of this analysis will be a diagnostic. Based on this diagnostic, the available draft is partitioned into sections, and each of these sections is marked as either to be left as it is (keep), to be regenerated (generate), to be revised (revise) or to be rejected (reject). Additionally, if the brief suggests sections should be added to the partial draft, place holders for them are added to the partial draft tagged as to be generated. If only a draft is available, an interpretation process is run on it to reverse engineer a brief. Based on the resulting brief, the available draft is processed as above. If neither brief nor draft are available, the creative system may follow different procedures, depending on whether a brief or a partial draft is constructed first.

Once the initial reflection phase is over, the system would enter a phase of construction. In spite of the similarities, we do not refer to this stage as engagement, because engagement in the sense used by Sharples applies very specifically to a process of production of new material, and the construction envisaged here may cover other processes such as revision, editing, or omisin. During this phase, each of the sections into which the draft has been partitioned will undergo the operation for which it has been tagged. The draft will therefore be edited by the application of the four basic operations described above. Any sections of the draft that are rejected at this point are stored in a log of fruitless paths.

At the end of the construction phase, the system would enter another phase of reflection. The first aim of this phase would be to ascertain whether the creative process has been concluded satisfactorily. This would arise if the draft matches the brief to perfection.

If the draft does not match the brief, the system would proceed with the rest of the reflection phase as described above, and iterated over another reflection/construction cycle. During reflection cycles other than the first one, the system may also consider modifications to the brief. These may arise from three possible situations. First, if part of the brief has proven impossible to satisfy during the prior cycle, the system may consider abandoning it. This would be plausible behaviour for human creators and should therefore be considered a possibility for artificial models. It would also constitute a very useful addition to allow creative system to steer themselves out of unproductive regions of a conceptual space when the current brief constrains them to restrict the search so. Second, if the reverse engineered brief shows positive features that were not included in the original brief, the system may decide to include them in the brief for the next iteration. This would allow such systems to incorporate the concept of serendipity into their computational models. Third, if the exploration of the conceptual space during a prior construction phase has included an excessive number of choices between possible candidate results, the system may decide to extend the brief to restrict the search to a subset of the conceptual space in question. Extensions to the brief should be compatible with the rest of it, and may take into account information about prior attempts that have failed.

Figure 2 depicts the reflective process in the proposed computational model for the tasks involved in literary creativity, as compared to the classic version that does not address feedback (depicted in Figure 1).

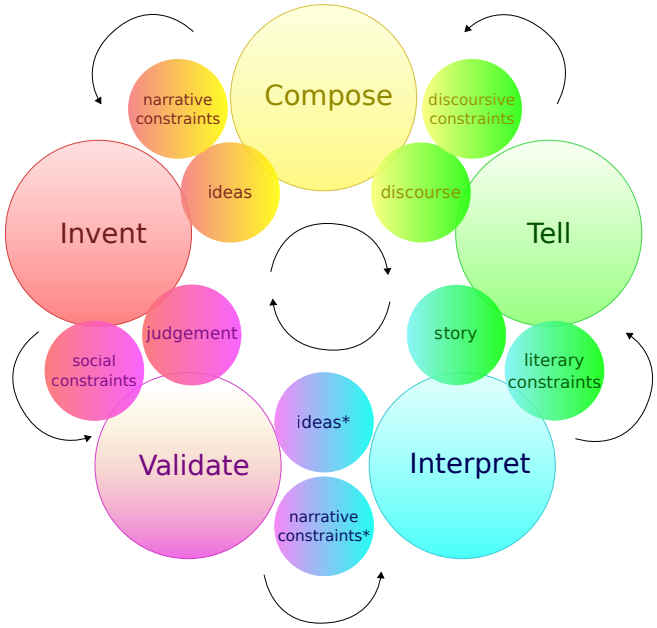

Figure 2: Graphical representation of the ICTIVS model as described in this paper, taking feedback into account. The feedback can trigger modifications of the brief (describing the constraints). The diagram represents draft generation in clockwise direction, and feedback revision in counter-clockwise direction.

4 Discussion

The planning process that Flower and Hayes has a dual purpose of generating ideas – which could correspond to additions to the ongoing draft – and of generating goals for the other processes. While the model that Flower and Hayes propose is not necessarily focused on creative writing, we consider it to be fundamental for describing narrative composition from the point of view of feedback. The ability to generate additions to the ongoing draft would correspond to Sharples process of engagement, with the slight refinement that the ideas generated in that case are restricted to the conceptual space defined by the initial constraints – which would correspond to our definition of the brief. These processes are covered by the generate option of our construction phase. The ability to set goals for other processes as established during the planning process that Flower and Hayes describe in their model corresponds to the establishment of constraints as described by Sharples. These processes are covered in our model by the part of the reflection stage where modifications to the brief are considered – in cycles beyond the initial one.

The translating process described by Flower and Hayes is described as a process of transforming ideas into text. According to our definitions, this would correspond to the task of generating a new instance of a particular section of the draft according to the corresponding brief, as carried out during a construction stage.

Our reflection process combines features of the reviewing process as described by Flower and Hayes – in as much as it involves evaluating the material produced so far and determining which parts of it can stand and which ones need further operations performed upon them – and the reflection stage of Sharples' model in the case of cycles beyond the initial one – where diagnostic leads not to revision of the draft but of the brief, which corresponds closely to Sharples' constraints. These constraints may then affect subsequent processes of revision but also of further construction or reconstruction of material already in the draft. The detailed description of the reflection stage as described by Sharples can be revisited using the proposed new terminology, which allows for finer consideration from a computational point of view of the actual operations involved at each point. The step that Sharples names reviewing involves a process of representational redescription – following Karmiloff-Smith – which clearly goes beyond comparing the results obtained with the original brief. Sharples specifically describes how the results obtained are processed to make available for reasoning within the system knowledge about "the procedures enacted during composition". Following Karmiloff-Smith, the raw data received from the generation processes are interpreted and internal representations are constructed describing valuable properties of these data at a more abstract level. This may take several forms, but a simple solution is to consider attempts to reverse engineer from the resulting draft a hypothetical brief that may have lead to it. This might be a reasonable match for the "explicit representations of specific constraints" that Sharples describes as likely outcomes of the process. The task of comparing such a reverse-engineered brief with the one actually used to drive the construction process would match Sharples step of contemplation, which operates on the results of the reviewing step. Sharples includes in the reflection stage an additional step of planning, where the results of contemplation lead to the creation of plans or intentions to guide the next phase of engagement. To fully capture the subtleties of Sharples analysis we have considered that the validation stage may result in the application of the four operations we have described – reject, generate, revise, keep – to the pair of briefs under consideration: the one used to generate the draft under revision and the one reverse engineered from the actual draft obtained. From this process a new brief will emerge, which can inherit constraints from the original brief, or delete them.

The reflection stage as we have described, in as much as it includes procedures for deciding which operations are to be carried out next on which parts of the work in progress, integrates the task of monitoring the creative process to guide the interaction and alternation between its constituent subprocesses – as described in Flower and Hayes model. The description provided in the present paper, given that it starts from a more fine-grained representation of both the data and the operations under consideration – allows for a more detailed and expressive description. This should allow for easier implementation of instances of creative systems that consider this type of behaviour.

With respect to the ICTIVS model, the model proposed in this paper may be seen as a refinement at a lower level of detail regarding the types of data involved and the types of operation carried out on them, but phrased at a higher level of abstraction with respect to the type of artefact being considered. The ICTIVS model was designed for the specific domain of narrative, and because of this it included separate stages for the ideation of plot or fabula and the composition of such plots or fabulae into sequential discourses. This would correspond to having different narrative levels of representation – fabula and discourse – for the material within the draft, and contemplating a specific process of conversion from one to the other. When we abstract away from these features specific to narrative, we can consider that the stages of *INVENTION* and *COMPOSITION* of the ICTIVS model would correspond to te construction phase that we have described in the current model, the *INTERPRETATION* and *VALIDATION* stage would correspond to the reflection stage, and the *TRANSMISSION* stage would correspond to the actual action of publishing or sending the final draft to an audience, which would correspond to the fulfilling the stopping condition implicit in our current phrasing of the reflection stage. With respect to the low level details described above, the task of reverse engineering a brief from a partial draft would match closely the processing that is considered during the interpretation phase of the ICTIVS model. The task of comparing such a reverse-engineered brief with the one actually used to drive the construction process would match the ICTIVS stage of validation.

The main advantage of this proposal is that two new sources of additional constraints are now included in the model of the creative process. First, the reverse-engineered brief may contain valuable constraints that were not in the original brief. This would correspond to the occurrence of serendipity: the constructive process employed leads to valuable features that were not in the original brief but which may be noticed during contemplation/interpretation of the result, and from then on added explicitly to the brief for subsequent iterations of the process. Second, the model allows for an explicit process for the generation of new constraints. This allows for the design of systems that can autonomously search for their own constraints, which would allow for a broader range of creative process. With respect to Boden's well known taxonomy of creative system [Boden, 2003], as the constraints being considered define the conceptual space that is being explored, a system capable of modifying the constraints that drive it might be capable of achieving transformational creativity. In this way the proposed models allows for a more fine grained representation of data and processes that may lead to the development of more expressive solutions.

The importance of making the system able to reject selected parts of its original brief should not be underestimated. The ability of human creators to depart – sometimes in very radical ways – from their original intentions in search for new aesthetics experiences has long been considered a critical ingredient of creativity of the highest order. It ties in very closely with the concept of transformational creativity, and the implicit ability to shift into new paradigms rather than just explore the old ones. Although all such issues are currently beyond the state of the art of creative systems, it is important to enable our computational models to represent

the types of behaviour that may one lead to implementation of similar behaviours. This would correspond to building explicitly into our computational models the idea of creativity at the metalevel [Wiggins, 2006]. The issue of how such operations might be profitably controlled is beyond the scope of the present paper and will need to be addressed in further work. Overall, a large proportion of the success of a creative system as described in the present proposal will depend on the implementation of suitable strategies for the partitioning of the draft into sections requiring the different operations available, and on the procedures for modifying the brief. These should be the focus of further work along the lines described in this paper.

5 Conclusions

The processes of literary creativity involve a complex web of interacting procedures (generation from a brief, evaluating how a draft matches a given brief, revision of an intermediate draft to fit a given brief, identifying unexpected valuable features from a working draft, editing a brief to optimise the search for creative results,...) and strategies for navigating between them. Existing cognitive models cover this space of solutions, but tend to remain at a high level of abstraction that leaves many of the features relevant for computation underspecified. The existing computational models of the writing task that have tried to take the cognitive models into account have focused on specific features of the process as their engineering mainstays, without trying to address the full complexity of the problem as a whole. The present paper proposes a computational model of the writing task that considers a broader set of ingredients than had been considered before, represented at a lower level of granularity in terms of their computational nature, both in terms of data and in terms of operations. The resulting model shows a strong potential for capturing significant phenomena in the field of creativity not often modelled computationally in the past, such as revision of drafts, working to a given brief, serendipity, and transformational creativity.

A valuable contribution of the proposed model is that it opens for exploration a significant number of lines of research to explore how these various phenomena might be addressed either in terms of working implementations of the proposed computational model or refinements of its basic formulation.

Acknowledgments

This paper has been partially supported by the project WHIM 611560 funded by the European Commission, Framework Program 7, the ICT theme, and the Future Emerging Technologies FET program.

References

[Boden, 2003] Margaret Boden. *Creative Mind: Myths and Mechanisms*. Routledge, New York, NY, 10001, 2003.

[Corneli *et al.*, 2014] Joseph Corneli, Alison Pease, Simon Colton, Anna Jordanous, and Christian Guckelsberger. Modelling serendipity in a computational context. *CoRR*, abs/1411.0440, 2014.

[Flower and Hayes, 1981] L. Flower and J.R. Hayes. A cognitive process theory of writing. *College Composition and Communication*, 32(4):365–387, 1981.

[Gervás and León, 2014] P. Gervás and C. León. Reading and writing as a creative cycle: The need for a computational model. In *5th International Conference on Computational Creativity, ICCC 2014*, Ljubljana, Slovenia, 06/2014 2014.

[Karmiloff-Smith, 1995] A. Karmiloff-Smith. *Beyond Modularity: A Developmental Perspective on Cognitive Science*. A Bradford book. A Bradford Book, 1995.

[Pease *et al.*, 2013] Alison Pease, Simon Colton, Ramin Ramezani, John Charnley, and Kate Reed. A discussion on serendipity in creative systems. In *Proceedings of the Fourth International Conference on Computational Creativity*, page 64–71, Sydney, Australia, jun 2013.

[Pérez y Pérez, 1999] R. Pérez y Pérez. *MEXICA: A Computer Model of Creativity in Writing*. PhD thesis, The University of Sussex, 1999.

[Sharples, 1996] M. Sharples. An account of writing as creative design. In C. M. Levy and S. Ransdell, editors, *The Science of Writing: Theories, Methods, Individual Differences, and Applications*. Lawrence Erlbaum Associates, 1996.

[Sharples, 1999] M Sharples. *How We Write*. Routledge, 1999.

[Wiggins, 2006] G. Wiggins. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7), 2006.