Encouraging p-creative behaviour with computational curiosity

Kazjon Grace[1],[2], Mary Lou Maher[2] Maryam Mohseni[2], and Rafael Pérez y Pérez[3]
[1]The University of Sydney, [2]UNC Charlotte, and [3] UAM Cuajimalpa
[1]Australia, [2]USA, and [3] Mexico
kazjon.grace@sydney.edu.au, {m.maher, mmohseni}@uncc.edu, rperez@correo.cua.uam.mx

Abstract

A concept, design or other artefact is p-creative when it is simultaneously novel and valuable for a specific individual. This is defined by contrast to h-creative artefacts, which are novel and valuable for a society as a whole. When we talk about p-creativity in computational systems we usually mean that something is creative to the system itself: the system has its own experiences and goals, and with them judges novelty and value. We propose an alternative approach aimed at *simulating* what a specific human user will find p-creative in order to *stimulate* that user towards p-creative behaviour. We define a framework for simulating curiosity, explore several domains in which it could be applied, and describe some preliminary results from a system designed to suggest papers for students to read that they would find surprising. We end the paper with a discussion of how this model can be extended to generate *framing narratives* that combine content from different artefacts that encourages p-creative behaviour.

Introduction

An artefact observed by a creative agent (human or artificial) is p-creative ("psychologically" creative) when that agent considers it novel and valuable (Boden 2004), regardless of whether other agents or the society as a whole would agree. Increasing the number of such observations is clearly beneficial for creativity, but we argue it may also be of benefit outside creative domains. The more different kinds of food we eat, the healthier we tend to be (Vadiveloo et al. 2015). The more broadly we read fiction, the greater our ability to understand others' emotions (Kidd and Castano 2013). Employees with both breadth and depth – "t-shaped" people – are sought out for their collaboration skills (Berger 2010). In each of these situations it is *seeking p-creative experiences*, rather than any specific goal, that is desirable. This paper describes a framework for interactive systems that encourage p-creative behaviour in their users, with applications both within and outside creativity.

Encouraging p-creative behaviour in a user is distinct from encouraging creative behaviour in general. The majority of computationally co-creative systems drive the user towards their best estimate of h-creativity (or "historical" creativity – artefacts judged to be creative by society as whole). Encouraging a user to pursue p-creativity requires knowing what *they* will find novel and valuable. Our reasoning for trying to directly encourage p-creativity is based on its impact on the user's motivations. Unexpected discoveries play an important role in driving the user towards more creative outcomes (Suwa, Gero, and Purcell 2000), a partial explanation for which may be the impact of curiosity on learning and performance (Reio and Wiswell 2000). Our goal is to develop creative systems capable of simulating the user's novelty and value functions with sufficient fidelity to suggest p-creative actions. We hypothesise that these systems may act as a kind of *curiosity engine*, repeatedly stimulating curiosity and encouraging behavioural diversification.

This paper is structured in three parts. We first describe a framework for this kind of system which we call the *Personalised Curiosity Engine*, or PQE (pronounced "pique"). We then describe an initial prototype of one component of the PQE system in the domain of text: a model of what makes a document unexpected. We conclude with a discussion of how the task of suggesting content in PQE systems could be framed as a form of narrative generation.

PQE systems persuade their users to take p-creative actions by *simulating* their novelty and value functions in order to *stimulate* their curiosity. They could be applied to creative tasks as a way to overcome fixation (Purcell and Gero 1996) and encourage diversity. They could also be applied to any other situation in which diverse action is of benefit, such as food or reading. PQE systems applied outside of traditionally "creative" domains are a kind of *persuasive computing* (Fogg 2009). These systems draw on computational creativity techniques to inspire curiosity and persuade users to consider more diverse actions.

p-Creative experiences are motivating

Behaviour change systems are a class of persuasive interactive system concerned with encouraging users to make sustainable changes to their behaviour in domains like education and health (Fogg 2002). Despite some early successes, behaviour change has remained largely an unsolved problem, for the simple reason that old habits die hard (Klasnja, Consolvo, and Pratt 2011). Using extrinsic rewards or social reinforcement to motivate significant sustained behaviour change is extraordinarily challenging, regardless of whether technology is involved. Our approach leverages a different motivator: curiosity.

An individual's motivation to perform an activity can belong to two broad classes: the extrinsic motivation to perform an activity for a reward (e.g. money, status or grades), and the intrinsic motivation to perform an activity for its own sake. Intrinsic motivation leads to greater interest, confidence, and performance (Ryan and Deci 2000).

Intrinsic motivation is tightly connected to the drive for self determination and self growth, and represents a principal source of enjoyment throughout life (Csikszentmihalyi and Rathunde 1992). Curiosity is one model for intrinsic motivation: it is the drive to seek novel stimuli for the sake of learning alone (Merrick and Maher 2009; Barto, Singh, and Chentanez 2004). The state of curiosity (as distinct from the trait of curiosity, see (Berlyne 1966)) has been modeled as seeking stimuli that are new enough to arouse interest but not so new as to cause disgust (Saunders and Gero 2001), and as seeking stimuli that most improve one's model of the world (Schmidhuber 2010).

Our approach leverages these cognitive models to simulate the curiosity of an individual and suggest recipes that will best stimulate their intrinsic motivational drive. Past computational models of curiosity have been used to generate novel content (Saunders and Gero 2001; Merrick and Maher 2009). Our approach differs in that we will simulate what an individual user will find curious, rather than imbue curiosity within an intelligent interactive system.

PQE: The Personalised Curiosity Engine

The PQE framework is a high-level approach to building systems that encourage p-creative behaviour, both in creative domains and everyday life. The underlying assumption is that a human user is engaged in taking actions within a particular domain with the goal of having p-creative experiences. The actions a user takes will always include engaging with artefacts within the domain, but may also include creating new ones. For example, in a music domain the user can listen to tracks composed by others, and may or may not be engaged in composition herself. In a cooking domain the user will eat and/or cook recipes written by others, and may or may not create her own. Describing interactions with our framework this abstractly lets us also apply it in cases where the user is not creating new artefacts, such as helping a graduate student discover new research papers and topics. This parallels the notion of a "serendipitous" recommendation (Herlocker et al. 2004), a notion which has also been studied in creative contexts (Corneli et al. 2014), although we avoid that term as it evokes the notion of such discoveries occurring by chance. Our long-term hypothesis is that repeatedly engaging with p-creative stimuli suggested by a PQE system would, over time, diversify a user's preferences.

Simulating p-creative evaluation in PQE

To make p-creative suggestions to our user (i.e. *stimulate* them towards p-creative behaviour) we must first estimate what they will find p-creative (i.e. *simulate* their evaluation of p-creativity). Standard models of novelty are built from the system's knowledge, with the assumption that it reflects the domain as a whole. In our case we must first develop a novelty model based on the knowledge of the user.

Techniques for approximating the knowledge (for novelty modelling) and preferences (for value modelling) of an individual fall under the area of *personalisation and user modelling* (Kay 1994). Modelling the value function of an individual (their "p-value", by analogy to p-creativity) requires knowing that individual's preferences within the domain.

Perhaps the most common paradigm for modelling preferences is the recommender system (Ricci, Rokach, and Shapira 2011). Recommenders use a model of the user and/or the items to select a set of results likely to be chosen by the user. The goal of these approaches is to use the preference ratings of a small set of known artefacts to estimate preferences across the whole domain. Very efficient algorithms exist for both user-based ("collaborative") and item-based ("content-based") preference modelling, and systems implementing those models are near-universal in online commerce. These preference models present a natural starting point for simulating users' value functions.

Modelling individual novelty (i.e. "p-novelty") is more complex. Our previous work has developed models of novelty based on expectations (Grace and Maher 2014; Grace et al. 2015). We define a novel artefact as one that violates an observer's expectations, where those expectations are based on the observer's past experiences (Baldi and Itti 2010). This makes novelty fundamentally subjective: to simulate the novelty of an individual one must estimate their knolwedge of the domain. We propose an approach to simulating a user's novelty that does not require a full record of what domain artefacts they have observed.

Stimulating p-creative behaviour in PQE

We propose systems that encourage p-creative behaviour by making suggestions for what actions to take. The simplest approach to delivering suggestions to the users is to wait for the user to query the PQE system and then provide a list of suggestions as a kind of search result. This is the approach adopted by recommender systems: the user is known to be looking for something, so tailor the information that is retrieved based on the available user model. While simple to implement, this approach may not effectively stimulate p-creative behaviour. Firstly, users may not be in a mindset compatible with the "search" metaphor, as they may be actively creating or acting independently.

Suggestions, like creative artefacts, might reasonably be expected to be more appreciated when explained (Moran and Carroll 1996). This draws to mind the concept of creative systems that provide *framing* for their generative acts (Charnley, Pease, and Colton 2012). One possibility for going beyond simple recommendation is to provide compelling framing alongside the suggestions made by our systems, an approach we discuss later in this paper.

Structure of the PQE framework

The structure of the framework can be seen in Figure 1. PQE Users provide feedback on their actions in the domain in the form of preferences and what surprises them. This feedback is used to estimate their value and novelty functions respectively. This feedback may be prompted in response to observing a specific new artefact, or it may come in the form

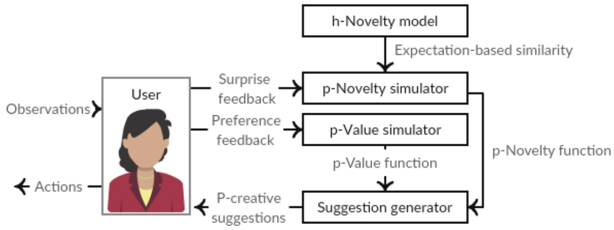

Figure 1: The three processes of the domain-general PQE framework and how they interact with a user to encourage p-creative behaviour.

of more general questions about likes, dislikes, familiarities and surprises. This feedback is then used by the *novelty simulator* and *value simulator* processes.

The **value simulator's** task is to estimate the user's value function for the whole domain, given information about some subset of that domain for which feedback has been provided. This feedback comes in the form of preferences or ratings: value judgements of artefacts, attributes of artefacts, or classes to which artefacts belong. The value simulator process must infer the user's values across the whole domain (their personal "value function") from this feedback. It may draw on knowledge about the relationships between artefacts and/or users to do so.

The **novelty simulator's** task parallels that of the value simulator: it must estimate the user's novelty function for the whole domain feedback. Like the user's value feedback, this comes in the form of judgements about novelty. The user reports that they find some artefacts or features to be low-novelty ("unsurprising") and others high-novelty ("surprising"). Unlike the value simulator this sparse feedback is insufficient to infer a novelty function for the domain. The h-novelty model provides the required additional knowledge.

The **h-novelty model** is based on our prior work in expectation-based models of novelty. These models take the entire database and construct a set of expectations in the form of conditional probabilities. PQE uses this model to provide a similarity metric that the novelty simulator can use to compare objects that the user has rated to objects that they have not. This similarity metric treats objects that are based on similar expectations as being similar. This results in estimates of how surprising a user will find a new artefact that are based on how surprising they have found similar combinations of surprising features elsewhere. For example, a user who found a mix of sweet and sour flavours totally unsurprising in a stir fry recipe will likely find the same combination unsurprising in a salad. This then allows the novelty simulator to operate in the same way that the value simulator does: using known ratings to infer unknown ones.

The **suggestion generator** uses the inferred value and novelty functions to provide recommendations to the user. These are intended to influence what areas of the domain the user explores (i.e. what artefacts they observe, create, consume or otherwise engage with). We do not specify the exact form these suggestions may take, but examples include search results or a single recommended action.

The PQE framework is applicable to computational co-creativity contexts (in which a human's creative acts are being supported or enhanced), but also to other domains that are not traditionally considered "creative". Discovering new things that are simultaneously novel and valuable to you can happen in any domain, this is a core component of the notion of *everyday creativity* (Runco and Richards 1997). In the following section we describe an pilot implementation of one component of the framework, the h-novelty model, in the domain of research papers.

Applying PQE to Research Papers

We are prototyping PQE in the domain of research papers, as part of an eventual future system that encourages students to read more broadly. Research papers are an example of a creative domain where unexpected discoveries are valuable and may lead to future creative research acts that generate their own papers. Our current goal is to understand what makes a research paper unexpected and to build an h-novelty model for this domain. The implementation presented here does not yet include personalisation.

In this paper we describe one promising model of research paper novelty, and demonstrate some preliminary results in the form of highly novel and highly not novel papers. Our next step will be to validate these results against the novelty judgements of humans who are experts in the research field of the papers in our corpus. If the experts are in agreement, then they can be expected to correlate with h-creativity in this domain (Amabile 1996).

For the purposes of these experiments we use a database of 298,000 research paper abstracts from the ACM Digital Library[1]. These documents span many of the most prominent journals, magazines and conferences in computing research. Abstracts were used as a summary of each paper.

Modelling expectations in text

For this prototype we defined novelty as exhibiting unexpected conceptual combinations, a kind of *relational expectation* (Grace et al. 2015). In this section we explain the details of the process by which we identify concepts within text documents, compute their relationships, and then assess the unexpectedness of the combinations that appear in abstracts of the published papers.

We adopt a topic modelling approach to inducing concepts from our corpus (Blei, Ng, and Jordan 2003). A topic model is a probabilistic graphical model (a type of statistical model) for learning the themes that occur in a collection of documents. First each document is represented as a "bag of words", an approach that ignores word order and context in order to provide a unified vector representation for each document. These models then produce a set of "topics", each consisting of a distribution over all the words in the corpus. This is based on the modelling assumption that topics are a probabilistic mixture of all the words in the corpus. Words which feature strongly in a topic are assigned a relatively high probability. For example, our model produced

[1] http://dl.acm.org

a topic which assigned the greatest probability mass to the words "network", "node", "protocol", "communicat+"[2], and "rout+". This topic clearly relates to computer networking.

Documents are then assigned different proportions of each topic, as in a mixture model. For example a paper on the Internet of Things may be drawn from 60% the above networking topic and 40% from a combination of topics about sensing, the Internet and experimentation. Topics are not labelled by the system and they are not guaranteed to be comprised of a single theme that is easily human-comprehensible, but they are usually at least moderately interpretable as in the example above.

We base our model of expectations on an extension of the basic topic modelling algorithm called a Correlated Topic Model, or CTM (Blei and Lafferty 2007). The advantage of this specific algorithm is that it allows for topics to be more or less correlated, i.e. the networking topic described above occurs more frequently with a topic about cybersecurity than it does with a topic about image recognition, as those themes are more conceptually related. This forms the basis of our expectation model: topics are concepts inferred from the dataset, and the correlations between topics give us a basis for what combinations of concepts are unexpected. Our prototype uses the R package "STM" (Roberts, Stewart, and Tingley 2014) to construct these models (STM is another topic model extension that is equivalent to a fast CTM implementation in some configurations). We use the default number of topics, 40, in our investigation.

In our previous models of relational expectation we have argued that the overall novelty of an artefact should be equal to the most novel concept or combination of concepts within that artefact (Grace et al. 2015). As a thought experiment on why we prefer this approach over averaging or combining multiple surprising artefact components, consider two fish. The first fish is slightly longer than expected, a slightly unexpected shade of blue, and has slightly bigger eyes than you would normally see. The second fish is physically unremarkable but can sing like a classically-trained soprano.

Our model of expectation bases the novelty of a text as the lowest (i.e. highest negative) correlation coefficient among all pairs of topics significantly present in that text, and the proportion of the document which contains that pair. We determine whether a topic is "significantly present" in a document using a topic proportion threshold of 0.1 (i.e. the document is at least 10% comprised of that topic). The formula can be seen in Equation 1, given a document $d = [t_i, t_j, \ldots, t_n]$ consisting of the set of topics significantly present. t_i and t_j are pair of topics in d which have the lowest (i.e. highest negative) correlation coefficient.

$$\left(CovMat_{t_i,t_j} \Big/ \min_{k=1\ldots K, l=1\ldots K}(CovMat_{k,l})\right) * 2(\min(prop(d,t_i), prop(d,t_j))) \quad (1)$$

[2]We use "+" to denote a stem formed from the combination of multiple words with the same root, e.g. "communicate", "communication' and "communicator". Singular word forms are always combined with their plurals and are not marked.

Where $CovMat$ is the covariance matrix for the topic model and $prop(d,t)$ is a function that returns the proportion of a document d that is comprised of a topic t. The first term is the novelty of the document's most novel topic combination, expressed as a proportion of the most novel topic combination in the model. The second term is twice the smaller of the smaller of the two proportions of the document that come from the novel topics. The product of the two gives the normalised unexpectedness of the most unexpected topic combination weighted by how much of the document is made up of that combination.

The second term of the novelty equation was originally a sum of the two topic proportions, but we found this to favour documents that just passed the significance threshold with one topic, and were thus not particularly surprising. We adopted the minimum of the two topic proportions to weight our novelty measure towards documents that contained substantial amounts of both unexpected topics.

A document composed of 50% of each the two most novel topics in the model would be given a novelty of 1. A document containing at-most independent topics would have a novelty of 0. Documents containing only positively correlated topics have negative novelty. Documents containing a lot of a moderately novel topic combination will be rated more novel than documents containing only a little of the most novel pair of topics.

Results

The top five words of each of the 20 topics in our CTM trained on the ACM Digital Library's abstracts were:

1. *image+, method, object, use, surfac+*
2. *application, service, mobile, provide+, resourc+*
3. *model, simulat+, use, process, operat+*
4. *comput+, will, student, learn, course*
5. *program, languag, code, use, implement+*
6. *system, design, develop, softwar, tool*
7. *perform, memor+, parallel, execut+, processor*
8. *search, propos+, method, feature, result*
9. *network, node, protocol, sensor, rout+*
10. *user, inform+, web, content, use*
11. *data, quer+, database, efficien+, large*
12. *algorithm, problem, graph, comput+, time*
13. *framework, structur+, specif+, approach, relat+*
14. *method, test, measur+, predict, use*
15. *result, analys+, evaluat+, stud+, effect*
16. *problem, strateg+, agent, decision, mechan+*
17. *power, design, energ+, circuit, propos+*
18. *interact, user, use, interface, game*
19. *secur+, attack, policy, privacy, can*
20. *research, social, stud+, group, communit+*

The majority of these topics have clear meanings within the domain of computer science research. For example, topic 2 is clearly about mobile services and their associated infrastructure, while topic 5 is clearly about programming languages and their features. Topics 6, 8, 13, 14, and 15 relate to the language used to describe research itself, rather than specific sub-fields of content. Topics 4, 7, 17, 18 and 19 clearly refer to specific research topics, respectfully computing education, parallel computing, chip engineering, human-computer interaction, and cybersecurity. We have not performed any formal verification of the validity of this model, but they pass casual inspection as reasonably reflective of the abstracts from the ACM digital library.

Using this model we present three of the most unexpected and three of the most expected papers from our dataset of abstracts. In each case we provide a link that can be used to view the abstracts at the ACM Digital Library. The three highly surprising documents can be seen in Table 1

The first paper in Table 1 combines Topic 2 (mobile service infrastructure) with a Topic 12 (algorithms research). While to a casual observer these seem highly related topics, they are actually fairly contra-indicated within computing research. Mobile services are a highly applied topic, while algorithms are basic research. The abstract for this paper is also highly unusual, containing mathematical formalisms to describe the problem the paper solves. The topic model picked up on this abstract's unusual wording.

The second paper in Table 1 has an extremely short abstract, which appears to have resulted in the topic model assigning very concentrated topic proportions to it. This abstract belongs to a magazine article, a format which does not traditionally have article abstracts – closer inspection reveals the abstract to be a pull quote from the first page. This, other than highlighting the challenge of obtaining quality data even from one of the most reputable archives of computing research, shows the disadvantage of working with abstracts. Neither of these papers is particularly novel-seeming, but their abstracts are certainly highly atypical.

The third paper is the most traditionally "novel" according to our topic model approach. The paper combines topic 20, which appears to be associated with social science in computing, with topic 12, the algorithms and graph research topic. This is because the paper describes algorithms for finding particular kinds of groups of people in social graphs. We also note that topic 12 was present in all of the top three papers, and actually the top four most surprising topic combinations in the dataset. The algorithms research theme appears to the be most polarising of all twenty topics: it is highly correlated and highly uncorrelated with many other topics, and independent from few.

By contrast all three of the papers in Table 2 are highly concentrated combinations of two most related topics in our model: 13 and 5. Topic 5 is about programming languages, while topic 13 is about the more ephemeral artefacts of computing research: frameworks, approaches, models, and other structures. In all three papers the abstract describes a new structure (two frameworks and an approach) that contributes to programming languages research in some way. This – for better or worse – appears to be by far the least surprising kind of paper in the ACM DL.

Our current prototype only consists of the h-novelty model process in the PQE framework (see Figure 1). We are developing a recommender systems approach to answering user searches with a list of results that is simultaneously fit to the query and novel to the user. In the following section we discuss some of the shortcomings of this recommender systems approach, and describe an alternative interaction model for PQE that draws more closely on computational creativity techniques.

Beyond Recommendation: Narrative framing for behavioural suggestions

The novelty model we presented above is one component of the PQE architecture in Figure 1. We are in the process of developing a complete implementation containing personalisation of both novelty and value as well as a suggestion generator. In this section we discuss an approach to that suggestion generation component that goes beyond the "recommender systems" paradigm of providing a list of results. One problem with the recommendation approach is that it ignores the iterative nature of the PQE task. It is not possible to build *towards* any artefact or concept that might not be appreciable by the user given their current knowledge.

Result lists also do not typically provide a compelling framing for *why* each suggestion was made. Explanations have been incorporated into recommender systems (Tintarev and Masthoff 2011), but to our knowledge they have not explicitly been designed to be compelling or persuasive. We intend to compare the recommender systems model (providing an unordered, unframed set of suggestions) with a list of suggestions designed to be consumed sequentially, with each entry having accompanying framing. This framing will explain what the user should look for in each artefact, and what that will contribute to the overall goal of the sequence.

Our proposed system draws an analogy between this task of ordering and explaining suggestions and plot generation. Our system will be based on the Engagement-Reflection (ER) model, which has previously been employed to characterise plot generation (Pérez y Pérez and Sharples 2001), interior design (Pérez y Pérez, Aguilar, and Negrete 2010) and visual compositions (Pérez y Pérez, De Cossío, and Guerrero 2013), among other domains. In general terms, this model is composed of elements, relations between those elements, and actions that progress those relations. In the case of storytelling the elements are the characters in the narrative, between whom there exist emotional links and conflicts. Story-actions progress the tale by evolving those emotional links and tensions between characters. The advantage of this narrative analogy over viewing this as a planning problem is the ability to model *tension and interestingness*. A planning approach may produce a sequence of resources to satisfy the goal, but we hypothesise that a narrative generation approach could additionally *keep the user interested*.

In our application of the ER model to framing for behavioural suggestion elements will be the "concepts" within the domain, and their relationships will be determined by how those concepts interact within the user's past knowledge. The story actions are a sequence of artefacts, each of which introduces, relates, or elaborates on the user's knowledge about the concepts in the "plot". The start of our "story" is the the user's current level of understanding of the domain. The culmination is the user appreciating a desired goal artefact that would not be comprehensible given their current level of knowledge. Framing will be generated

Title	URL	Unexpected topic combination
1: "Facility location with Service Installation Costs"	`http://dl.acm.org/citation.cfm?id=982953`	[*algorithm, problem, graph, comput+, time*] (43%) & [*application, service, mobile, provide+, resourc+*] (35%)
2: "Volunteer computing: the ultimate cloud"	`http://dl.acm.org/citation.cfm?id=1734164`	[*algorithm, problem, graph, comput+, time*] (39%) & [*comput+, will, student, learn, course*] (36%)
3: "Star Search: Effective Subgroups in Collaborative Social Networks"	`http://dl.acm.org/citation.cfm?id=2810062`	[*research, social, stud+, group, communit+*] (31%) & [*algorithm, problem, graph, comput+, time*] (26%)

Table 1: The three most unexpected paper abstracts in our database.

Title	URL	Expected topic combination
1: "A simple rewrite notion for call-time choice semantics"	`http://dl.acm.org/citation.cfm?id=1273947`	[*framework, structur+, specif+, approach, relat+*] (43%) & [*program, language, code, use, implement+*] (43%)
2: "Constrained kinds"	`http://dl.acm.org/citation.cfm?id=2384675`	[*program, language, code, use, implement+*] (43%) & [*framework, structur+, specif+, approach, relat+*] (42%)
3: "Slicing as a program transformation"	`http://dl.acm.org/citation.cfm?id=1216375`	[*program, language, code, use, implement+*] (50%) & [*framework, structur+, specif+, approach, relat+*] (42%)

Table 2: The three least unexpected paper abstracts in our database.

to explain each story action's contribution towards the goal.

Engagement-Reflection for PQE suggestions

Figure 2 shows a schema of the ER model, with the engagement state on the top and the reflection state underneath. The process starts with a special instance of the Add Action process in which the initial world state – characters and their contexts – is added to the new action sequence. The model then iteratively probes its memory to recall parts of previous sequences that match, selects one action from a matching sequence to add, and then adds it. After a certain number of engagement cycles (3 in Mexica) reflection occurs and the emerging sequence is evaluated for interestingness, cohesion and whether it fits provided constraints. When no sufficiently similar stories exist in memory it enters Reflection to break the impasse, modifies the sequence, and then returns to Engagement.

In PQE we propose that this cycle could take place within the *suggestion generator* process, with the initial world state being provided by the novelty and value models of the user. ER requires a set of existing narratives that can be recalled during the process, which would need to be learnt or provided. A separate process would need to select a (currently unreachable due to being too novel) goal artefact for the ER model to strive towards. This process of inferring a goal from a particularly novel and/or valuable state parallels with the model of specific curiosity in (Grace et al. 2017).

Inducing concepts for plot generation

Any implementation of the narrative suggestion process outlined above must grapple with defining what a "concept" is, how they will be learned, and how they will relate. In order to be useful for both PQE and ER, concepts are required to have certain properties:

1. Concepts must capture the themes that characterise artefacts within the domain (i.e. be useful for describing artefact meanings).
2. It must be possible to model the likelihood of concepts co-occurring so that they can be used to form the h-novelty model described in the PQE framework above.
3. Concepts must be comprehensible, so that the framing process can communicate them to the user.
4. Concepts must be algorithmically learnable from the data. Manual labelling or pruning may be appropriate, but the process should be mostly automated.
5. Concepts must have a notion of *accessibility* – what other concepts are sufficiently related that they should be appreciable by a user who comprehends this concept. This is required for the ER model to construct sequences of artefacts that lead to an eventual goal.

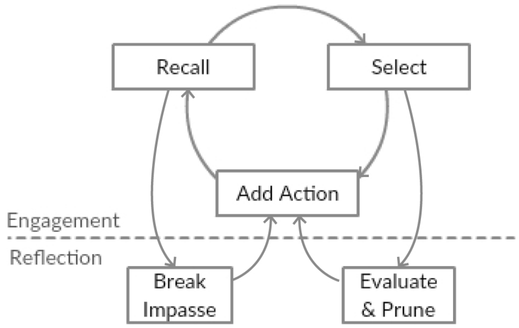

Figure 2: The processes of the ER model as it will be integrated with PQE. This cycle takes place inside the Suggestion Generator process.

From these properties we can see that the topic model representation described in this paper is likely insufficient. It satisfies the first, second and fourth properties, and perhaps the third as some methods for visually conveying topic models have been developed (Chuang, Manning, and Heer 2012). Topic models offer no simple way to describe what topics are adjacent or accessible to each other in terms of learnability. We are pursuing ways to develop a model of the concepts within research papers that also satisfies this fifth property. One possibility is to use the citation networks that exist between papers to characterise how new work builds on old. Another possibility is to identify what is mentioned when a paper is cited, and use that to infer what concepts a paper explains.

What determines "tension" for a framing narrative?

One role of the Reflection process in the ER model is to evaluate the interestingness of the emerging narrative by comparing it to a desired narrative structure. Narratives typically follow a degradation-improvement structure in which the situation gets worse and worse for the characters before a climactic moment after which things start to get better. This may happen cyclically, often with degradations of increasing magnitude. In ER this is called the tensional representation of a narrative.

In the PQE narrative framing task tension is not derived directly from the characters (the concepts in the artefacts being suggested), but from the user's knowledge of those concepts. New concepts being introduced to the narrative raise the tension, as they can be reasonably be expected to increase the user's confusion. Equally, concepts that are not currently connected to the other concepts within the story increase the tension, as the goal of these narratives is to render a single specific goal artefact comprehensible. Story actions (i.e. suggestions) that connect previously disparate concepts will decrease tension. We will also investigate modelling the tension between concepts that are (or seem) contradictory, although this will require a more detailed representation than the current topic models approach provides.

Conclusion

The PQE framework describes how computational creativity techniques might be used to encourage users to diversify their behaviour. The core principle of this framework is suggesting p-creative actions, based on the hypothesis that they will stimulate the user's curiosity (their intrinsic motivation to explore the domain). The PQE framework is also an opportunity to apply computational creativity techniques outside of domains that are traditionally thought of as "creative". The approach applies to domains, such as education and nutrition, where divergent behaviour is beneficial.

We have prototyped an h-novelty model, one component of the PQE system, in the domain of research paper recommendation. We model novelty in unstructured text documents using topic models, a machine learning approach that induces the dominant "themes" of a database and describes how they relate. We base our novelty model on papers that exhibit unexpected combinations of these topics. Our prototype, which operates only on research paper abstracts rather than the full text, can successfully identify abstracts that are highly novel as well as those that are highly conventional. We are working on developing this novelty model into a recommender-based system that can suggest papers that are simultaneously fit to the user's search queries, and surprising according to their knowledge of the domain.

We also describe our plans to incorporate this novelty model into a system that provides a specific sequence of artefact suggestions along with explanations of why the user should engage with each. This is based on treating the p-creativity stimulating suggestion task as one of narrative generation. We propose a approach based on the ER model that describes how such a system could construct a compelling journey through the domain, building at each step the user's capacity to appreciate a goal artefact.

In summary, our key contribution in this paper is an approach to suggesting unexpected research papers by identifying novel combinations of topics. We show how we can select documents that include topic combinations that have a very low probability of occurring together. We use a bag of words representation for each paper and the Correlated Topic Model algorithm to generate the distribution and correlation of topics across the corpus of research papers. We propose that this core element of our co-creative system has the potential to deliver surprising research papers to a student at a minimum, and beyond that to provide a plan for generating a surprising research paper by suggesting unusual combinations of topics for new research papers.

References

Amabile, T. 1996. *Creativity in context*. Westview press.

Baldi, P., and Itti, L. 2010. Of bits and wows: a bayesian theory of surprise with applications to attention. *Neural Networks* 23(5):649–666.

Barto, A. G.; Singh, S.; and Chentanez, N. 2004. Intrinsically motivated learning of hierarchical collections of skills. In *Proceedings of the 3rd International Conference on Development and Learning*, 112–19. Citeseer.

Berger, W. 2010. *CAD Monkeys, Dinosaur Babies, and T-Shaped People: Inside the World of Design Thinking and How It Can Spark Creativity and Innovation*. Penguin.

Berlyne, D. E. 1966. Curiosity and exploration. *Science* 153(3731):25–33.

Blei, D. M., and Lafferty, J. D. 2007. A correlated topic model of science. *The Annals of Applied Statistics* 17–35.

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3(Jan):993–1022.

Boden, M. A. 2004. *The creative mind: Myths and mechanisms*. Psychology Press.

Charnley, J.; Pease, A.; and Colton, S. 2012. On the notion of framing in computational creativity. In *Proceedings of the Third International Conference on Computational Creativity*, 77–81.

Chuang, J.; Manning, C. D.; and Heer, J. 2012. Termite: Visualization techniques for assessing textual topic models. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, 74–77. ACM.

Corneli, J.; Jordanous, A.; Guckelsberger, C.; Pease, A.; and Colton, S. 2014. Modelling serendipity in a computational context. *arXiv preprint arXiv:1411.0440*.

Csikszentmihalyi, M., and Rathunde, K. 1992. The measurement of flow in everyday life: toward a theory of emergent motivation. In *Nebraska Symposium on Motivation*, volume 40, 57–97.

Fogg, B. J. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002(December):5.

Fogg, B. J. 2009. A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology*, 40. ACM.

Grace, K., and Maher, M. L. 2014. What to expect when youre expecting: the role of unexpectedness in computationally evaluating creativity. In *Proceedings of the 4th International Conference on Computational Creativity, to appear*.

Grace, K.; Maher, M. L.; Fisher, D.; and Brady, K. 2015. Modeling expectation for evaluating surprise in design creativity. In *Design Computing and Cognition'14*. Springer International Publishing. 189–206.

Grace, K.; Maher, M. L.; Wilson, D.; and Najjar, N. 2017. Personalised specific curiosity for computational design systems. In *Design Computing and Cognition'16*. Springer. 593–610.

Herlocker, J. L.; Konstan, J. A.; Terveen, L. G.; and Riedl, J. T. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)* 22(1):5–53.

Kay, J. 1994. The um toolkit for cooperative user modelling. *User Modeling and User-Adapted Interaction* 4(3):149–196.

Kidd, D. C., and Castano, E. 2013. Reading literary fiction improves theory of mind. *Science* 342(6156):377–380.

Klasnja, P.; Consolvo, S.; and Pratt, W. 2011. How to evaluate technologies for health behavior change in hci research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3063–3072. ACM.

Merrick, K. E., and Maher, M. L. 2009. *Motivated reinforcement learning: curious characters for multiuser games*. Springer Science & Business Media.

Moran, T. P., and Carroll, J. M. 1996. *Design rationale: concepts, techniques, and use*. L. Erlbaum Associates Inc.

Pérez y Pérez, R.; Aguilar, A.; and Negrete, S. 2010. The eri-designer: A computer model for the arrangement of furniture. *Minds and Machines* 20(4):533–564.

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

Pérez y Pérez, R.; De Cossío, M. G.; and Guerrero, I. 2013. A computer model for the generation of visual compositions. In *Proceedings of the Fourth International Conference on Computational Creativity*, 105. Citeseer.

Purcell, A. T., and Gero, J. S. 1996. Design and other types of fixation. *Design studies* 17(4):363–383.

Reio, T. G., and Wiswell, A. 2000. Field investigation of the relationship among adult curiosity, workplace learning, and job performance. *Human Resource Development Quarterly* 11(1):5–30.

Ricci, F.; Rokach, L.; and Shapira, B. 2011. *Introduction to recommender systems handbook*. Springer.

Roberts, M. E.; Stewart, B. M.; and Tingley, D. 2014. stm: R package for structural topic models. *R package version 0.6* 1.

Runco, M. A., and Richards, R. 1997. *Eminent creativity, everyday creativity, and health*. Greenwood Publishing Group.

Ryan, R. M., and Deci, E. L. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25(1):54–67.

Saunders, R., and Gero, J. S. 2001. Artificial creativity: A synthetic approach to the study of creative behaviour. *Computational and Cognitive Models of Creative Design V, Key Centre of Design Computing and Cognition, University of Sydney, Sydney* 113–139.

Schmidhuber, J. 2010. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development* 2(3):230–247.

Suwa, M.; Gero, J.; and Purcell, T. 2000. Unexpected discoveries and s-invention of design requirements: important vehicles for a design process. *Design studies* 21(6):539–567.

Tintarev, N., and Masthoff, J. 2011. Designing and evaluating explanations for recommender systems. In *Recommender Systems Handbook*. Springer. 479–510.

Vadiveloo, M.; Dixon, L. B.; Mijanovich, T.; Elbel, B.; and Parekh, N. 2015. Dietary variety is inversely associated with body adiposity among us adults using a novel food diversity index. *The Journal of nutrition* 145(3):555–563.