Regent-Dependent Creativity: A Domain Independent Metric for the Assessment of Creative Artifacts

Celso França, Luís Fabrício W. Góes, Álvaro Amorim, Rodrigo Rocha and Alysson Ribeiro da Silva
Programa de Pós Graduação em Engenharia Elétrica
Pontifícia Universidade Católica de Minas Gerais
Minas Gerais, Brasil 30535-901
{celso.franca,lfwgoes,alvaro.amorim,rcor,alysson.ribeiro}@pucminas.br

Abstract

Humans are the ultimate judges on how creative is an artifact. In order to be creative, most researchers agree that an artifact has to be at least new and valuable. However, metrics to evaluate novelty and value are often craft for individual studies. Even within the same domain, these metrics commonly differ. Although this variety of metrics extends the spectrum of alternatives to assess creative artifacts, the lack of domain independent metrics makes hard to compare artifacts produced by different studies, which in turn slows down the research progress in the field. In this paper, we propose an domain independent metric, called Regent-Dependent Creativity (RDC), that assesses the creativity of artifacts. This metric requires that artifacts are described within the Regent-Dependent Model, in which artifacts features are represented as dependency pairs. RDC combines the *Bayesian Surprise* and *Synergy* to measure novelty and value, respectively. We show two case studies from different domains (fashion and games) to demonstrate how to model artifacts and assess creativity through RDC. We also propose and make available a simple API to promptly use RDC.

Introduction

The beauty of computational creativity lies in its diversity, which ranges from music, culinary to science. In recent years, this myriad of possibilities has attracted researchers from across many research fields such as computer science, social sciences and arts into a quest to unveil the processes in which creativity emerges. This interaction of various disciplines led to the proposal of a plethora of creative systems, in which most of them have the ultimate goal of producing creative artifacts. The definition of what makes an artifact creative or not is still an evolving discussion, nevertheless researchers tend to agree that an artifact has to be new and valuable on a particular domain to be considered creative (Boden 2015; Colton et al. 2015; van der Velde et al. 2015). Although the concepts of novelty and value are intuitive to humans, they are not easily translated into a computer program, as they depend on individual knowledge, beliefs, tastes and cultural values (Boden 2015).

In order to tackle such a daunting task, researchers on computational creativity have proposed many solutions to assess how novel and valuable is an artifact. A popular approach is to ask what humans think about it (Lamb, Brown, and Clarke 2015; Karampiperis, Koukourikos, and Koliopoulou 2014), as we are the ultimate judges on creativity. In this way, creative artifacts are indeed evaluated, but the human implicit mechanisms to spot creativity are still left as a black box. A more analytical approach is the use of domain specific metrics to assess novelty, value and any other features that could be related to creativity such as satisfaction, plausibility, faithfulness, generality, etc. This metrics zoo is an expected outcome due to the huge challenges imposed by the complexity that is the assessment of creativity. However, as the sub-fields in computational creativity mature, some have been converging (Góes et al. 2016; Pinel, Varshney, and Bhattacharjya 2015; Maher and Fisher 2012) to standard methods and metrics to evaluate creativity, while others are still proposing new metrics (Tomasic, Znidarsic, and Papa 2014; Schorlemmer et al. 2014; Colton et al. 2014; van der Velde et al. 2015). The latter may still be beneficial on the long run, but it makes hard to compare existing solutions between different works, consequently slowing down the progress in these particular subfields.

As an ambitious goal, a domain independent creativity metric would be ideal to boost scientific research on computational creativity. In fact, previous work (Maher 2010; Maher and Fisher 2012) pursued this objective, but some issues remained: i) the proposed metrics lacked implementation details, making it hard to replicate them; ii) very little or no practical examples were provided, weakening their appeal to experimental researchers; and iii) the lack of quantitative results hindered to show the effectiveness of the metric across different domains.

In this paper, we propose the Regent-Dependent model which describes artifacts as a collection of features, represented by dependency pairs. Once represented in this model, artifacts can be evaluated by our proposed Regent-Dependent Creativity (RDC) metric.

The RDC metric combines the *Bayesian Surprise* and *Synergy* to measure novelty and value, respectively. In order to address some aforementioned issues found in previous work, we present: i) two case studies from different domains; ii) a quantitative evaluation of RDC; and iii) propose and make available a simple API to the research community that implements the RDC metric.

Novelty and Value

Creative artifacts have to be novel and valuable (Boden 2004). In order to evaluate novelty, there are few metrics based on concepts of unexpectedness, expectation and surprise that are commonly used (Grace and Maher 2014). On the other hand, value can be extracted from the associations and rules that bond individual artifacts (Varshney et al. 2013).

A novel artifact may be new only to a particular person or group. Alternatively, it may be entirely new in relation to all human history. The former type of novelty is required to achieve p-creativity (p for psychological), while the latter is concerned to h-creativity (h for historic) (Boden 2015). In practice, psychological creativity is more feasible, since it can be verified for a given dataset of known artifacts. In contrast, historical creativity imposes a dataset to have all existing artifacts, which its completeness is hard to be proved.

However, creativity is not just novelty, a creative artifact must also be valuable. Value is a measure of performance or attractiveness of an artifact which depends on its acceptance by an expert or society (Maher 2010). There are many types of value (e.g. beauty, scientific interest, musical harmony, utility etc.) (Boden 2015), and many of them are difficult to recognize, harder to put into words, and even more difficult to say clearly. For a computational model, however, it must be precisely defined (Boden 2009).

Even in science, values are often transient and sometimes changeable. The meaning of simplicity and elegance, when applied to scientific theories, is something that philosophers of science have long try - and fail - to precisely define (Boden 2004). In addition to it, if a scientific finding or hypothesis is interesting it depends on other current theories of the time and also in the social context. This is where creativity is concerned, the shock of the new can be so great that even for those who are the witnesses, it is difficult to see value in the new idea. This makes the calculation of value specific to a certain context, that is, the value of an artifact depends on the relationships between an artifact and the existing ones, more precisely, the associations of the artifacts features. When evaluating an artifact, its value is determined according to the combination of its features, which in turn are governed by rules that were implicitly created by humans in that context to value certain types of artifacts more than others. Once these rules and associations are expressed in a computer model, the value of an artifact can be determined.

Related Work

The existence of the "islands of creativity" problem, as recently highlighted by Bown (2015), suggests that a significant obstacle for the evaluation of computational creativity resides in the idea that creativity is situated on specific systems without any fluidity between these systems and the rest of the world. In fact, when it is not used a very specific metric, they appeal to random choice.

Some research tackle this problem by employing human computation to assess creativity in computer systems. Joyner et al. (2015) suggest that human computation can be an effective strategy to collect a wide variety of methods for creative tasks. From a set of existing solution methods to the intelligence test Raven's Progressive Matrices (RPM), they developed other new methods using crowd-sourcing, highlighting those that were most different and achieved significant success. On the other hand, Lamb, Brown, and Clarke (2015) point out that the quality of a creativity metric relies on the appropriate choice of human judges, which is addressed by the consensual assessment technique (Amabile 1982) from the field of psychology.

As opposed to the idea of using humans as judges, Cook and Colton (2015) proposed an alternative way to enable a software to make significant decisions. With the use of evolutionary algorithms which evolve short pieces of code called *preference functions*, it makes meaningful and justifiable choices between artifacts. As another approach to measure value, Jordanous, Allington, and Dueck (2015) investigate how to measure subjective and cultural value which have been expressed by members of a community towards other members. Focusing on the activity by electronic musicians on the music social network SoundCloud, they combined qualitative and quantitative research to understand and trace significant 'valuing activities' in Sound-Cloud data.

Maher (2010) proposed a domain independent metric to evaluate creative artifacts. It is based on novelty, value and unexpectedness. Novelty is measured as the distance from clusters of other artifacts in a conceptual space. In addition to it, value is calculated through a set of performance criteria. Finally, unexpectedness looks for variations in attributes by the use of pattern matching algorithms. Despite this research was the first to propose a domain independent creativity metric, it does not verify its applicability in real world examples.

Maher and Fisher (2012) extend the previous model proposed by Maher (2010), where creativity is evaluated based on novelty, value and surprise. In contrast to the previous unexpectedness metric, they use Bayesian inference based on prior probability for measuring the surprise of a given artifact. They suggest an application regarding the design of laptop computers.

Other work also focused on automatically assessing creativity using a creativity model, focused on aesthetics aspects, based on the probabilistic model for Bayesian inference and Shannon's measure of entropy (Burns 2006; 2015). Bayesian inference is applied for evaluating the meaning of a given artifact based on prior evidences and the psychological arousal produced by violating expectations is modeled mathematically by Shannon's measure of surprise (Rigau, Feixas, and Sbert 2007). The aesthetics is finally expressed as the product of Shannon-entropy measure of surprise and the Bayesian-probability measure of meaning. Some previous work have used Bayesian posterior probability or prior probability to model a notion of Bayesian surprise (Itti and Baldi 2009; Baldi and Itti 2010; Maher and Fisher 2012), instead of using the Shannon-entropy.

Similar to those previous work, we also use the Bayesian surprise metric for assessing novelty, which is known to be reasonably effective (Itti and Baldi 2009; Baldi and Itti 2010). However, in contrast to previous work, for modeling and measuring the value of a given artifact, we use concepts of synergy by extracting metrics of a graph-based knowledge representation of the artifact's properties.

Regent-Dependent Model

A single data model to describe each artifact is imperative to allow the creativity evaluation of artifacts produced by different systems. In this paper, we propose a data model to describe an artifact as a set of pairs between its features and their modifiers. This dependency relationship is defined by a pair *P(regent; dependent)* associated with a numeric value v. A *regent* is a feature that contributes to describe an artifact, and may be an action or attribute, while a *dependent* can change the state of an attribute or connect an action to a target. For example, an artifact *car* can be described by a pair $p_i(color; blue)$, where *blue* changes the state of the attribute *color*. The same artifact could also be described by another pair $p_i(drive; home)$, where the dependent *home* connects a target to the action *drive*. In a grammatical example, the famous slogan *"Just do it"* can be described by two pairs: $p_i(do; just)$ and $p_j(do; it)$. The first says that the adverb *just* is a modifier of the verb *do*, while the second pair connects the verb *do* to the direct object *it*.

The value v is important because it can be used to represent the intensity of a specific pair in different contexts. Different cultures have different preferences for culinary recipes, music and art. Even the science progress is weighted by social interests. Thus, the pairs can be modeled to these different situations. For example, a car with the *blue* color may be more common in certain countries than others, so the value v can be set higher than other colors.

With the definition of the presented data model, it is possible to build a *dataset* of existing artifacts and a graph of associations between the artifacts pairs which are required in our proposal to calculate novelty and value, respectively, as explained in the next sections.

Bayesian Surprise as a Novelty Metric

Novelty can only be evaluated compared to a group of existing artifacts. In this paper, we propose that novelty is calculated using a well-known metric called the Bayesian surprise (Baldi and Itti 2010), which enables to evaluate how much new is an artifact compared to existing ones in the *knowledge dataset*. This *knowledge dataset* is the description of existing artifacts, organized in instances following the Regent-Dependent model so that each instance is the representation of an artifact described by its pairs. Physically, a *dataset* is a matrix, where the rows are instances of artifacts and the columns its pairs.

The Bayesian surprise stems from the fact that a new artifact is unusual and surprising for the observer. This surprising effect is an interesting novelty detector that can be calculated by the application of Bayes' theorem, as shown in Equation 1.

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)} \tag{1}$$

According to this theorem, the probabilities represent subjective degrees of belief in hypotheses or models that are updated as new data is acquired. Thus, the degree of conviction of an observer is represented by a subjective probability function $P(h)$ measuring the degree of belief in the observer's hypothesis h. The term $P(h)$ is called *prior distribution*, and reflects the knowledge before new data are considered, whereas the term $P(h|d)$ is called *posterior distribution*, and as the name suggests, reflects the knowledge after consideration of a new fact d has occurred, and be inserted in the hypothesis h. Similarly, $P(d)$ is the probability that d occurs independently of the hypothesis h, and $P(d|h)$ is the probability that event d occurs given that h is true (Kruschke 2015).

Fundamentally, the effect of a new artifact is to transform an observer previous convictions in posterior convictions. Novelty thus can be quantified by considering the difference between the probability distributions that accurately describes how the world view of the observer has changed.

In fact, as shown empirically in recent studies, this approach is able to capture human notions of novelty in different types and levels of abstraction (Itti and Baldi 2009; Baldi and Itti 2010; Varshney et al. 2013). Mathematically, the novelty $n(p_i)$ of a *pair* p_i, regarding a specific artifact a, is calculated by Equation 2, where σ and $\bar{m}$ are respectively the variance and average of an existing *pair* in the *dataset* of artifacts, and μ_i is the value associated with p_i. The total novelty N_a of a given artifact a is defined as $N_a = \sum_{p_i \in a} n(p_i)$. Equation 3 computes the normalized novelty in the range [0,1], by means of an exponential normalization, where λ is a smoothing factor.

$$n(p_i) = \frac{1}{2\sigma^2}\left[\sigma^2 + (\mu_i - \bar{m})^2\right] \tag{2}$$

$$f(N_a, \lambda) = 1 - e^{-\lambda N_a} \tag{3}$$

Synergy as a Value Metric

Strategies for assigning a value to an artifact can be widely distinct from one domain to another. Even experts from the same domain will differ comparing two or more artifacts (Boden 2015). For this reason, there are several metrics to measure value. For instance, pleasantness measures the flavor perception of a recipe (Pinel, Varshney, and Bhattacharjya 2015; Varshney et al. 2013), an aggregation metric defines the quality of a slogan (Tomasic, Znidarsic, and Papa 2014) etc. However, these metrics are designed for specific domains.

On the other hand, artifacts, even in different domains, are composed of a set of elements that have actions and attributes. This set of elements and the interaction among them is what gives value to a certain artifact. For example, a recipe consists of ingredients, each with its own taste, its texture, and its aroma, the final flavor of the recipe, however, is the result of the combined actions of its ingredients (Corning 2012). This feature takes place in other areas, such as in music in which harmony occurs when two or more pitches are combined to produce a chord (Cope 2015), or in some turn-based strategy games, where players perform individual moves with a set of elements that together implement an efficient strategy (Millington 2009).

Moreover, there are plenty of information publicly available that describes artifacts and the elements that constitute them like in *fooDB* [1]. In particular, it is also available how these elements interact and what interactions are most valued in a given context or group of people, which is key to compose a valuable artifact. These facts give evidence that the relationship between the elements of an artifact can be used as a measure of value.

Our *Regent-Dependent Model* allows to represent an artifact by its elements, which in turn are described by regent-dependent pairs. This model also allows to build a graph

[1] Available at http://foodb.ca/

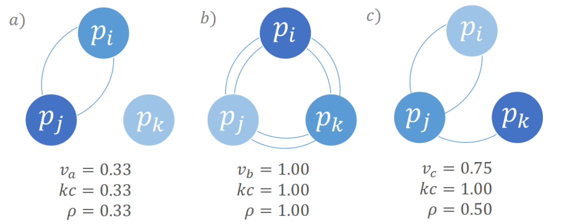

Figure 1: Relationships between pairs used to describe three different artifacts with isolated pairs (a), completely interconnected pairs (b) and reasonably associated pairs (c).

connected by pairs in which their relationship are valued in a particular context. Consequently, the most valuable artifacts in this graph are the ones which have pairs that are more interconnected among themselves.

A metric that captures this cooperative nature is synergy (Corning 2012). It measures the effect produced by various elements (forces, particles, parts or individuals) acting cooperatively in the development of emergent behaviors found in many real-world systems, such as the brain and other neural systems, stock markets, the Internet and social networking systems. The center of gravity of an object, for instance, is actually a synergistic effect, it depends upon how the combined weight of all its parts is distributed. Therefore, the value of an artifact can be measured by the synergy of the elements that exist within it.

To measure synergy, an artifact a is modeled as a graph $G_a = (V, E)$ where the vertex set V consists of pairs belonging to a and there exists an edge between two vertices p_i and p_j if and only if they belong to the same set of synergy. Each set of synergy is defined by two or more pairs that have complementary effects, i.e., they exhibit a better effect together than separately.

The graph is providential because it presents a series of metrics to measure the relationship between vertices. In particular, the *connectedness* and the *density* of the graph define fairly well a measure of synergy and the value v_a of an artifact a can then be defined by Equation 4.

$$v(a) = \frac{1}{2}kc(G_a) + \frac{1}{2}\rho(G_a) \tag{4}$$

where:

G_a: *is the graph which represents the artifact a.*
$kc(G_a)$: *is the Krackhardt's connectedness of* G_a.
$\rho(G_a)$: *is the density of* G_a.

The first term of the Equation 4 measures the associativity among the pairs of an artifact. If all pairs are associated, i.e., all pairs are reachable from every other, then $kc(G_a)$ *Krackhardt's connectedness* of G_a is maximum (Krackhardt 1994). If all pairs are isolated in the artifact, then $kc(G_a)$ is minimum. The second term measures the strength of the connection among the pairs of an artifact. Basically, it measures the average number of connections between two pairs p_i and p_j. Although this is a simple measure of the relationship between vertices, the relationship described in Equation 4 can be more descriptive to contain other graph metrics such as *concentration, diameter and max flow*.

Figure 1 shows graphs of three different artifacts described by pairs p_i, p_j and p_k. Each of these pairs is a vertex and the edges between a pair and another indicates a synergistic relationship. The values v_a, v_b and v_c of respectively artifact a, b and c, calculated by Equation 4, in the range [0,1], is greater when the pairs are fully connected and smaller when less associated the pairs are.

Regent-Dependent Creativity Metric

The proposed *Regent-Dependent Creativity (RDC)* metric, for calculating the creativity of an artifact a, is defined in Equation 5 as the sum of the normalized novelty n_a and value v_a plus an extra penalty term. This penalization is needed to avoid that high novelty artifacts with low value (different but useless artifacts), and high valuable artifacts with low novelty (useful but already known artifacts) are considered creative.

$$rdc(a) = n_a + v_a - p(n_a, v_a) \tag{5}$$

$$p(n_a, v_a) = s_a(1 - e^{-kd_a}) \tag{6}$$

where:

s_a: is the sum of n_a and v_a.
d_a: is the absolute difference of n_a and v_a.

Equation 6 penalizes an artifact depending on the difference among its novelty and its value. The greater the difference between novelty and value of an artifact, its creativity is more penalized. The penalty is more intense as the variable k is higher, however, no artifacts are penalized more than the sum of its novelty and its value. Therefore creativity is in the range [0,2].

Case Studies

In this section, we show two case studies to demonstrate how to model artifacts and assess their creativity using RDC. The first case study is a simplified example from the fashion domain to evaluate apparels. The second one is from the game domain. It is based on a real and large problem to create card combos in the game HearthStone.

Fashion it: Evaluating Creative Apparels

The creation of fashion artifacts is challenging given the diversity of factors such as style, color, patterns, materials, etc (Jagmohan et al. 2014). The challenge lies both in the combination of various elements of clothing with different styles and purposes for creating a complete apparel, and in subsequently ranking them based on certain criteria.

Consider a hypothetical case where in the space of clothing items available to compose an apparel, there is only one type of shoes, one type of pants and one type of shirt, varying only the color as shown in Figure 2(a). Given this space, the process of generating creative apparels reduces to combining the colors of the clothing items available to form a complete apparel *(shoes + shirt + pants)*.

Figure 2(b) shows some existing apparels that are considered interesting combinations for a fashion consultant and provides our prior knowledge.

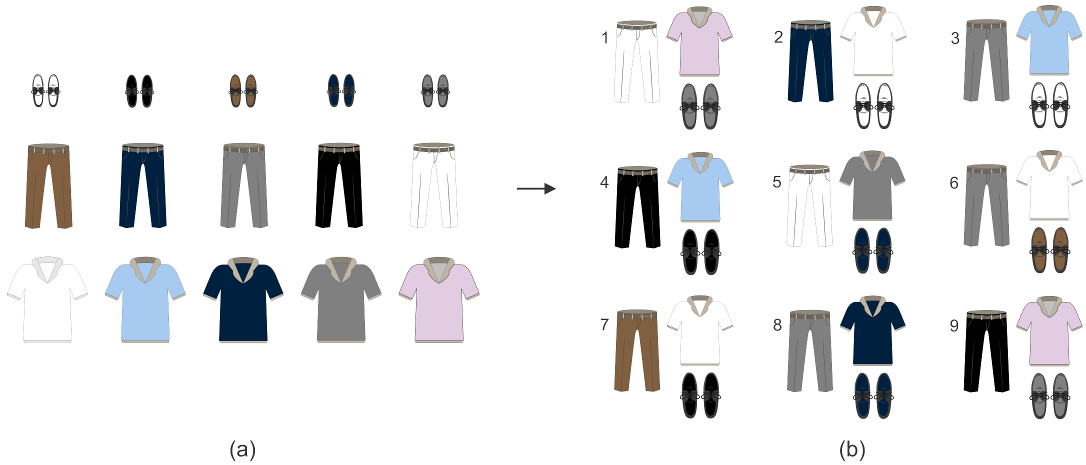

Figure 2: (a) Space of clothing items available to compose an apparel. (b) Existing apparels used to define the knowledge dataset.

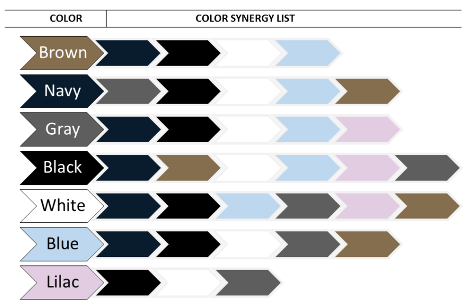

Figure 3: Color Synergy List

As the color is the only feature to be described and there are seven different color options *(white, black, navy, gray, blue, brown, lilac)*, we can use the *regent* to represent the clothing item while the *dependent* is one of the available colors. Thus, the set of *regents* has three elements *(shirt, pants, shoes)* and the set of *dependents* has seven elements. These definitions guide the construction of the *dataset* where each instance is an apparel and the attributes are elements p_i of set P of all pairs used to describe all the clothing item:

P ={*shirt:white, shirt:navy, shirt:gray, shirt:blue, shirt:lilac, pant:white, pant:black, pant:navy, pant:gray, pant:brow, shoes:white, shoes:black, shoes:navy, shoes:gray, shoes:brow*}

For example, the first apparel of Figure 2(b), would be described by the pairs *(shirt, lilac) = 1; (pant, white) = 1, (shoes, gray) = 1*. There are nine known apparels in the *dataset* as described in Table 1. Note that values are set to 1, since all pairs are equally important in this example.

The creativity assessment of an artifact is made according to its novelty and value, as presented in Equation 5. The novelty of an apparel can be calculated by Equation 2, using the apparel dataset.

On the other hand, to calculate value we need to know about the synergy of colors, i.e., what color combinations

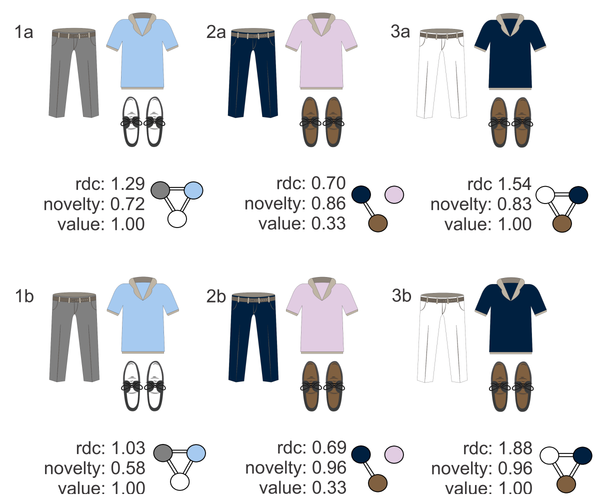

Figure 4: Behavior of the proposal metric for evaluation of different apparel.

are most valued. There are some techniques for combining colors based on color wheels, wherein a set of colors are harmonious when they fit some analogous, triad or pattern. Figure 3 illustrates a synergy list for each color based on the color wheels. The *brown* color for example, has synergy with the colors: *navy, black, white and blue*. Then, according to our proposal for the metric, an apparel would be modeled as a graph where the vertices are a clothing item and there is an edge between one clothing item and another if they have synergistic colors. With the complete graph, the value of the apparel can then be calculated by Equation 4.

Figure 4 shows the behavior of the proposed *RDC metric* in different scenarios. In apparel $1a$, for instance, the colors are all synergistic, so that the graph representing that apparel is completely connected and the application of the Equation 4 returns the maximum value. The novelty, however, is penalized because it is an apparel with an existing pattern in the *dataset*.

Apparel $2a$ has a non existent pattern in the *dataset*, consequently achieving high novelty. On the other hand, the synergy of colors occurs only between pants and shoes,

APPAREL	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}
A_1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0
A_2	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0
A_3	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0
A_4	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
A_5	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0
A_6	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1
A_7	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0
A_8	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0
A_9	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0

Table 1: *Dataset* of existing apparels. Each instance is an apparel and the attributes are the values v_i of the pairs used to describe each apparel. A nonexistent pair has value 0.

making the shirt a isolated vertex in the graph. The effect of this isolation is to reduce the value of this apparel to 0.33 and thereby creativity to 0.7.

Apparel $3a$, in addition to exhibit a non existent pattern in the *dataset*, has synergistic colors for shirt, pants and shoes, so it has a maximum score of creativity.

In order to further explore the RDC metric, the apparels $1b$, $2b$ and $3b$ of Figure 4 show the metric behavior after 10 additional inserts of the apparel $1a$ into the *dataset*. The apparel $1a$ becomes more common, reducing its novelty. However, as the *dataset* increased towards apparel $1a$, the apparels patterns $2b$ and $3b$ become even more novel.

HoningStone: Evaluating Creative Combos for HearthStone

Hearthstone[2], by Blizzard Entertainment, is a DCCG in which human players compete in one-versus-one matches in alternating turns, until a player is defeated. On each turn, a player can play any cards from his hand, use his hero power or minions to attack characters (minions or hero) and particularly combining cards, that is, playing *combos*. Thus, a combo is a group of related cards played in the same turn. In (Góes et al. 2016), a computational creativity system, called HoningStone, was proposed. It automatically generates creative card combos for *Hearthstone* based on the Honing theory of creativity (Gabora 2010). HoningStone used a creativity metric based on surprise and efficiency to generate and evaluate combos. These metrics used a dataset of 31000 distinct combos extracted from real game logs from 10000 decks played in more than 3 million matches obtained from the various public websites.

In this paper, we use the same knowledge dataset to model and evaluate a few card combos generated by HoningStone using RDC. We show how to use synergy as a value metric instead of efficiency.

Each combo is composed of cards, which in turn has effects. Each effect, described in the card's text, is modeled as a pair $P(ability, target)$ which has a value v. In a card which the text is *"destroy 2 minions"*, for instance, it is represented as $P(destroy, minion) = 2$. HearthStone produces 190 distinct pairs when combining all abilities and targets from the existing card set (Góes et al. 2016). Thus, the prior knowledge is composed by 31000 combos, each one represented by those effect pairs extracted from each card. A card c_i is synergistic to another card c_j when they have complementary pairs, i.e., the combined effect of the complementary pairs produces greater advantages than when played separately. Figure 5 shows the pairs' relationship for the cards c_1, c_2 and c_3, while Table 2 lists pairs of each of these cards.

Card	Pairs
C_1	a(plusAttack, alliedMinion) = 2 b(giveCharge, alliedMinion) = 1
C_2	c(enrage, self) = 1 d(plusAttack, self) = 1
C_3	e(dealDamage, character) = 1
C_4	f(dealDamage, minion) = 1 g(plusAttack, minion) = 2
C_5	g(enrage, self) = 1 i(plusAttack, self) = 1
C_6	j(plusAttack, self) = 1 k(plusHealth, self) = 1 l(drawCard, ownHand) = 1

Table 2: A subset of six cards and their respective pairs.

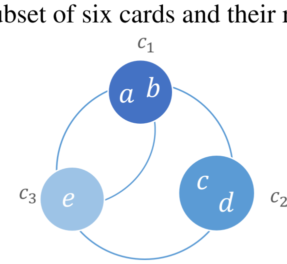

Figure 5: Associations between the effects pairs of cards c_1, c_2 and c_3.

For example, card c_2 and c_3 are synergistic as c_2 enrage effect, represented by pair c, is activated only when this card takes damage. In addition to it, pair e of card c_3 works as trigger that deals damage to card c_2, binding c_2 and c_3. This combination of cards and their complementary effects that makes a combo stronger. The same type of associations can be made to all other cards and effects. The more associations a card has to another, higher is the synergy.

Figure 6 shows the novelty, value and RDC for three combo examples, generated with HoningStone, using cards c_1, c_2, c_3, c_4, c_5 and c_6. Novelty is calculated using the

[2]Available at http://us.battle.net/hearthstone/en/

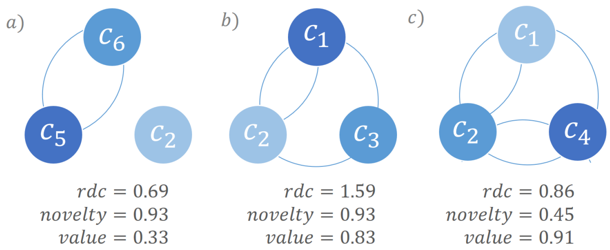

Figure 6: Behavior of the RDC metric for evaluation of different combos.

knowledge dataset of 31000 combos, and synergy uses a simplified set of associations which covers all the effects in cards from c_1 to c_6. Combo b is novel according to the knowledge database and also has a high synergy, leading to a high RDC. On the other hand, combo a presents low value and high novelty. This gap penalizes creativity since it is a new combo but not very effective.

Regent-Dependent Creativity API

In order to complement the presented case studies, an API for evaluating artifacts was developed[3]. Only two inputs are required to evaluate an artifact: a knowledge database that contains existing artifacts of a particular application domain, and a set of relations that represent the synergy among the artifacts' attributes. The knowledge database must contain artifacts encoded in JSON format. In the first example, where clothing items are combined to form an apparel, the knowledge database has the following format:

```
[   {
        "clothingItems": [{
                "type": "SHIRT",
                "color": "LILAC"
            }, {
                "type": "PANTS",
                "color": "WHITE"
            }, {
                "type": "SHOES",
                "color": "GRAY"
            }]
    }, ...
```

A specialized parser is responsible for converting the encoded knowledge database into a collection of instances of artifact objects. The decoded collection of artifact objects is depicted bellow:

```
{
  "1":[0,0,0,0,1,1,0,0,0,0,0,0,0,1,0],
  "2":[1,0,0,0,0,0,0,1,0,0,1,0,0,0,0],
  ...
  "9":[0,0,0,0,1,0,1,0,0,0,0,0,0,1,0]
}
```

Relations representing the synergy of the artifacts are structured as a map between each attribute and its respective synergistic attributes. These relations are illustrated in figure 3, describing the synergy among the colors and their clothing items. The API supports the synergistic relations to be represented as follows:

[3]Source-code for the Regent Dependent Creativity API: `https://github.com/CreaPar/rd-creativity-metric-api`

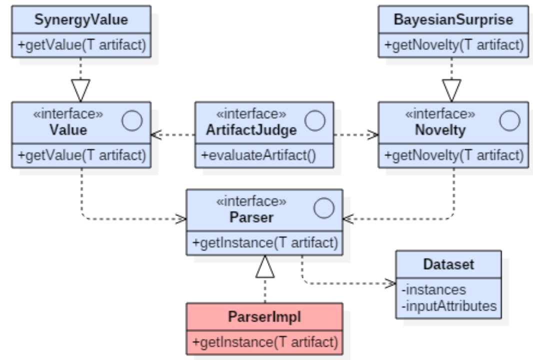

Figure 7: Class diagram

```
{
"WHITE":["NAVY", "BLACK", "BLUE", "GRAY", "LILAC", "BROWN"],
"BLACK":["NAVY", "BROWN", "WHITE", "BLUE", "LILAC", "GRAY"],
"NAVY":["GRAY", "BLACK", "WHITE", "BLUE", "BROWN"],
"BLUE":["NAVY", "BLACK", "WHITE", "GRAY", "BROWN"],
"GRAY":["NAVY", "BLACK", "WHITE", "BLUE", "LILAC"],
"BROWN":["NAVY", "BLACK", "WHITE", "BLUE"],
"LILAC":["BLACK", "WHITE", "GRAY"]
}
```

When the parser loads the knowledge database, it computes the mean and variance of each attribute among all loaded artifacts. These information is useful for the calculation of the RDC metric. The two main classes responsible for the Regent-Dependent creativity metric are: the *SynergyValue* class, responsible for calculating the value metric, in which the method *getValue (T artifact)* will return the synergistic value of the artifact given as parameter; and the *BayesianSurprise* class, responsible for calculating the novelty metric, by using the method *getNovelty (artifact T)*. With a measure of novelty and value, the *evaluateArtifact()* method in ArtifactJudge class, judges how creative is an artifact according to Equation 5 using RDC. Figure 7 shows the implementation details of the API.

Conclusion

Despite the proposal of several metrics to assess the creativity of artifacts, still computational creativity lacks metrics that can be used across different domains. This paper addresses this issue by proposing the Regent-Dependent Creativity (RDC) metric, based on the *Bayesian surprise* and *synergy* to measure novelty and value. The presented results show the use of RDC in two different domains: fashion and games. The fashion case study is simplified but is a throughout example to show each step to model and use RDC. The second one is a real world example to show that the model is applicable to larger problems. This paper also presented an API, which is available online, with a full example so researchers can promptly use RDC to evaluate artifacts.

As future work, RDC can be used into several other domains, such as culinary, arts, music etc. RDC can also be used as a creativity metric to guide the generation of artifacts. The API can be extended to accommodate other metrics and filled up with more examples. In addition to it, we can validate RDC using human experts to assess creativity through techniques such as Consensual Assessment and human computation. We hope that RDC helps the computational creativity community to boost progress in this challenging research field.

Acknowledgment

We would like to thank FIP PUC Minas, FAPEMIG, CNPq and CAPES to support this work.

References

Amabile, T. M. 1982. The Social Psychology of Creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology* 43(5):997–1013.

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.

Boden, M. A. 2004. *The Creative Mind: Myths and Mechanisms*. London: Routledge, 2 edition.

Boden, M. a. 2009. Computer models of creativity. *AI MAGAZINE* 13(2):72–76.

Boden, M. A. 2015. Creativity and ALife. *Artificial Life* 21(3):354–365.

Bown, O. 2015. Attributing creative agency: Are we doing it right? In *Int. Conf. on Computational Creativity*, 17–22.

Burns, K. 2006. Atoms of eve': A bayesian basis for esthetic analysis of style in sketching. *Artificial Intelligence for Engineering Design, Analysis, and Manufacturing* 20(03):185–199.

Burns, K. 2015. Computing the creativeness of amusing advertisements: A bayesian model of burma-shave's muse. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 29:109–128.

Colton, S.; Pease, A.; Corneli, J.; Cook, M.; and Llano, T. 2014. Assessing progress in building autonomously creative systems. In *Int. Conf. on Computational Creativity*, 137–145.

Colton, S.; Pease, A.; Corneli, J.; Cook, M.; Hepworth, R.; and Ventura, D. 2015. Stakeholder groups in computational creativity research and practice. In *Computational Creativity Research: Towards Creative Machines*. Springer. 3–36.

Cook, M., and Colton, S. 2015. Generating code for expressing simple preferences: Moving on from hardcoding and randomness. In *Int. Conf. on Computational Creativity*, 8–16.

Cope, D. 2015. Computational creativity and music. In *Computational Creativity Research: Towards Creative Machines*. Springer. 309–326.

Corning, P. 2012. *Nature's Magic: Synergy in Evolution and the Fate of Humankind*. Cambridge University Press.

Gabora, L. 2010. Revenge of the neurds: characterizing creative thought in terms of the structure and dynamics of memory. *Creativity Research Journal* 22(1):1–13.

Góes, L. F. W.; da Silva, A. R.; Rezende, J.; Amorim, A.; França, C.; Zaidan, T.; Olímpio, B.; Ranieri, L.; Morais, H.; Luana, S.; and Martins, C. A. P. S. 2016. Honingstone: Building creative combos with honing theory for a digital card game. *IEEE Transactions on Computational Intelligence and AI in Games* PP(99):1–1.

Grace, K., and Maher, M. L. 2014. What to expect when you're expecting: the role of unexpectedness in computationally evaluating creativity. In *Int. Conf. on Computational Creativity*, 120–128.

Itti, L., and Baldi, P. 2009. Bayesian surprise attracts human attention. *Vision Research* 49(10):1295–1306.

Jagmohan, A.; Li, Y.; Shao, N.; Sheopuri, A.; Wang, D.; and Varshney, L. R. 2014. Exploring Application Domains for Computational Creativity. In *Int. Conf. on Computational Creativity*, 328–331.

Jordanous, A.; Allington, D.; and Dueck, B. 2015. Measuring cultural value using social network analysis: a case study on valuing electronic musicians. In *Int. Conf. on Computational Creativity*, 110–117.

Joyner, D. A.; Bedwell, D.; Graham, C.; Lemmon, W.; Martinez, O.; and Goel, A. K. 2015. Using human computation to acquire novel methods for addressing visual analogy problems on intelligence tests. In *Int. Conf. on Computational Creativity*, 23–30.

Karampiperis, P.; Koukourikos, A.; and Koliopoulou, E. 2014. Towards machines for measuring creativity: The use of computational tools in storytelling activities. In *Int. Conf. on Advanced Learning Technologies*, 508–512.

Krackhardt, D. 1994. Graph theoretical dimensions of informal organizations. In Carley, K. M., and Prietula, M. J., eds., *Computational Organization Theory*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc. 89–111.

Kruschke, J. 2015. *Doing Bayesian data analysis : a tutorial with R, JAGS, and Stan*. Boston: Academic Press.

Lamb, C.; Brown, D. G.; and Clarke, C. L. 2015. Human competence in creativity evaluation. In *Int. Conf. on Computational Creativity*, 102–109.

Maher, M. L., and Fisher, D. H. 2012. Using AI to evaluate creative designs. In *Int. Conf. on Design Creativity*, 45–54.

Maher, M. L. 2010. Evaluating creativity in humans, computers, and collectively intelligent systems. In *Network Conference on Creativity and Innovation in Design*, 22–28.

Millington, I. 2009. *Artificial Intelligence for Games*. CRC Press.

Pinel, F.; Varshney, L. R.; and Bhattacharjya, D. 2015. A culinary computational creativity system. In *Computational Creativity Research: Towards Creative Machines*. Springer. 327–346.

Rigau, J.; Feixas, M.; and Sbert, M. 2007. Conceptualizing birkhoff's aesthetic measure using shannon entropy and kolmogorov complexity. In *Computational Aesthetics*, 105–112.

Schorlemmer, M.; Smaill, A.; Kai-uwe, K.; Kutz, O.; Colton, S.; Cambouropoulos, E.; and Pease, A. 2014. COINVENT : Towards a Computational Concept Invention Theory. In *Int. Conf. on Computational Creativity*, 288–296.

Tomasic, P.; Znidarsic, M.; and Papa, G. 2014. Implementation of a Slogan Generator. In *Int. Conf. on Computational Creativity*, 340–343.

van der Velde, F.; Wolf, R. A.; Schmettow, M.; and Nazareth, D. S. 2015. A semantic map for evaluating creativity. In *Int. Conf. on Computational Creativity*, 94–101.

Varshney, L. R.; Pinel, F.; Varshney, K. R.; Schorgendorfer, A.; and Chee, Y.-M. 2013. Cognition as a part of computational creativity. In *IEEE Int. Conf. on Cognitive Informatics and Cognitive Computing*, 36–43.