

Measuring analogical creativity - An experiment

Stergos Afantenos, Leonardo Cortez Bernardes, Henri Prade

Institut de Recherche en Informatique de Toulouse (IRIT), Univ. Toulouse, Toulouse INP, CNRS
stergos.afantenos@irit.fr, bernardesleonardo0@gmail.com, henri.prade@irit.fr

Abstract

Analogical reasoning is said to be creative. After discussing views of computational creativity and its evaluation, we consider a particular form of analogical inference based on analogical proportions, which are quaternary relations comparing the differences inside two pairs of items. This inference mechanism is capable of generating a new item from three items described on the same set of features. This mechanism is used, in an experiment, to generate new prompts that are addressed to an LLM-based chatbot to generate constrained text. The prompts specify the emotions at work in a short text describing a mundane situation. Hellinger distance and Kullback-Leibler divergence are used for provisional assessments of creativity at the prompt level or the text level. The relative merits of these two measures of creativity assessment are discussed. The paper proposes an original creative mechanism coupling analogical inference and LLM-based text generation.

Introduction

Creativity refers to the ability to develop novel, and preferably valuable, ideas or productions. This clearly applies to a wide range of works. Human creativity ranges from arts, literature and humour to problem solving which itself covers a variety of situations encountered in mathematics and sciences as well as in problems of everyday life. Creativity as a major asset of the human mind, has been a challenging issue in artificial intelligence research for a long time (Boden 1994; McCormack and d’Inverno 2012; Colton and Wiggins 2012), with illustrations in particular in music (Wiggins and Forth 2018), in the visual arts (Cook and Colton 2011), or in literature (Van de Cruys 2020).

Drawing analogies has a reputation for being creative, especially in problem solving. This has been advocated by many authors, in cognitive psychology, e.g., (Dreistadt 1969; Halpern 1986), in artificial intelligence (Veloso 1993), or in mathematics (Polya 1954). Creativity seems all the more evident and striking as the analogy brings together two distant situations, or two quite different frames of reference.

Analogical reasoning can take various forms according to different modeling frameworks.¹ In structure mapping theory and related approaches (Gentner 1983; Falkenhainer,

Forbus, and Gentner 1989; Hummel and Holyoak 1997), analogical reasoning is considered to be the ability to draw parallels between two universes (or, more simply, two situations) by matching the relationships between elements of the first universe (and their properties) to those applicable to elements of the second universe, and to draw potential conclusions in the second universe based on the similarity of relationships in the two universes (the second universe is supposed here to be less known than the first). This kind of analogical reasoning can be encoded in a classical logical framework (requiring second-order anti-unification) (Gust, Kühnberger, and Schmid 2006) where one can, for example, capture Rutherford’s famous creative analogy, which compares the solar system (playing the role of the first universe) to Rutherford’s atomic model (the second universe) (Schmidt et al. 2014).

Case-based reasoning (Kolodner 1993; Aamodt and Plaza 1994; Fuchs et al. 2020) is often considered as a simple form of analogical reasoning which uses a repertory of cases, made of pairs (problem, solution), and proposes a solution for a target problem by adapting the solution of a similar problem found in the repertory of cases. However, the creativity of case-based reasoning remains relatively modest, to the extent that the problems being compared are similar.² However, the creativity of case-based reasoning remains rather modest inasmuch the problems brought together are similar. A parallel is thus established between the pair (problem, solution) used and the target problem with its proposed solution. One can say that the proposed solution is to the target problem as the solution of the similar case is to the problem of this case.

The above statement that links four items is an analogical proportion and is of the form “ a is to b as c is to d ”. In fact when one parallels a pair (a, b) in a first universe with the corresponding pair (c, d) in the second universe, one obtains

can also mention the iconic approach of analogical reasoning (Kunda, McGreggor, and Goel 2013) which has been applied to the solving Raven’s Progressive Matrix tests, a problem which can also be dealt with an analogical proportion-based approach (Afantenos et al. 2024; Correa Beltran, Prade, and Richard 2016)

²Several works rely on case-based reasoning that have improved upon ideas on creativity that appear in (Schank and Kass 1990); let us mention, for example, (Ramalho and Ganascia 1994), (de Mántaras 1998) in the field of music creation.

¹Apart from the approaches reviewed in this introduction, we

such proportions, as pointed out in (Hesse 1959). For instance, “the electron is to the nucleus as the planet is to the sun” in Rutherford’s analogy. In such an example, the analogical proportion involves two (quite different) universes; in case-based reasoning, one has another type of split between problems and solutions.

There is a third type of analogical reasoning that is directly based on analogical proportions, where knowing that “ a is to b as c is to d ”, one infers d (or an unknown part of the description of d) from the triplet (a, b, c) . It requires that the four items be of the same kind in the sense that they can be described on the same set of attributes, which is not necessarily the case in the two previous approaches. It is worth pointing out (Barbot, Miclet, and Prade 2019) that analogical proportions may either involve four items describable by the same set of attributes, as in “a puppy is to a dog what a kitten is to a cat”, or associate two universes (that of a and c , and that of b and d) which refer to two distinct sets of attributes, as in “wine is to the French as beer is to the English”. In this paper, we only focus on the first kind of analogical proportions.

Analogical proportions appear to be creative in the sense that a, b, c and d can be all different, and where d — which can be obtained from a, b and c — is described by a recombination of the features of a, b and c . However the question of evaluating such a creativity is raised. In this paper, we propose an experiment that illustrates this creative process in the area of text generation, and try to measure the novelty of what is produced.

This paper investigates creativity using analogical proportion-based inference. It is organized into three main parts. The first part reviews different aspects of creativity and ways to assess it. It also considers means of measuring creativity that have been proposed, such as Hellinger distance or Kullback-Leibler divergence, and discusses their relative merits. The second main part provides a background on analogical proportions and analogical inference, and points out their potential for creativity, which is theoretically evaluated. The third main part describes the setting of an experiment: the analogical proportion-based inference is used to generate combinations of features pertaining to a short text that reports a sequence of ordinary facts. These combinations of features are then used as prompts that are addressed to an LLM-based chatbot to generate constrained texts. Thus the creativity can be both measured at the prompt level and at the text level. Then the results of the experiment are reported, which is followed by a final discussion.

Creativity and its evaluation

In this section, we first review the work on computational creativity aimed at formalizing or assessing it. We then recall two statistical divergence measures proposed to assess novelty, a major aspect of creativity.

Review of works on computational creativity

Computational creativity refers to AI systems capable of imitating human creative processes. This brings us back to human creativity. It is why it may be useful to start with

the definition of (human) creativity given in a highly-cited paper (Rhodes 1961): “The word creativity is a noun naming the phenomenon in which a person communicates a new concept (which is the product). Mental activity (or mental process) is implicit in the definition, and of course no one could conceive of a person living or operating in a vacuum, so the term *press* is implicit. The definition begs the questions as how new the concept must be and to whom it must be new.” This first proposed view is known for introducing four “p”: person, product, process and press, involved in any creative processes.

There has been an active research trend in computational creativity for more than thirty years. Broadly speaking, in addition to work on specific computer programs capable of producing new works in the fields of art, literature or mathematics, a series of works have attempted to define, categorize and formalize computer creativity in general, or to propose general methods for assessing creativity.

This started with the influential work of Boden (Boden 1994; 2003), who distinguished between three kinds of creativity: combinational, exploratory, and transformational. In combinational creativity, the result is obtained by a mere combination of standard ideas, while the two other forms of creativity take place in a conceptual space, the first one explores this space, while the second transforms it (e.g., by relaxing some of the constraints defining the conceptual space). Transformational creativity is regarded as the highest form of creativity which goes beyond mere novelty.

Bundy (1994) points out that according to Boden’s view (Boden 2003), an automated theorem prover solving an open conjecture in mathematics would not be considered truly creative, the new result being obtained by an exploratory procedure. In contrast, when the open conjecture is solved by a human, it is likely that the human work will be considered creative. Bundy emphasized the need to take into account the possible great complexity of exploring the conceptual space. Let’s add that in mathematics, apart from the difficulty of the result to be established, an innovative proof based on a new way of looking at the problem can be considered an example of transformational creativity.

Besides, Boden (2003) distinguishes between p-creativity (“p” for psychological) and h-creativity (“h” for historical) where the former is judged in relation to the agent’s productions, while the result of the latter creativity is regarded as new by society.

Boden’s inspiring work has led some researchers to attempt to make more formal and refine her definitions of a conceptual space or transformational creativity, see, e.g., (Ritchie 2006) for a discussion. One noticeable line of research is the development of a ‘Formal Creativity Theory’ by Wiggins (Wiggins 2006a; 2006b; 2019), where the conceptual space is understood as a set of possible artefacts associated with a set of rules that select those that are acceptable (just playing the role of a quality checking), and where another set of rules attribute a value to a created artefact (Wiggins 2006b).

Wiggins views transformational creativity as exploratory creativity at the meta-level, taking up an idea defended in (Buchanan 2001). Creativity is viewed as a generalized

search process i) without closed-world assumption, ii) allowing for the combination of existing ideas, iii) distinguishing between heuristics that evaluate the quality of solutions and those that are used for finding a solution efficiently, and iv) preferring the notion of comparison to that of value. This view of creativity as search has been further discussed and restated in greater details in (Ritchie 2012). Recently, the setting of Formal Creativity Theory has been used for arguing that “while novelty is neither necessary nor sufficient for transformational creativity in general, when using an inspiring set, rather than a sequence of experiences, an agent actually requires novelty for transformational creativity to occur” (Santo, Wiggins, and Cardoso 2024).

Another line of research inspired by Boden’s work, is the approach, more pragmatic, proposed by Ritchie (2001; 2007), which proposes a series of criteria for attributing creativity to a computer program, based on three indicators. Namely, in addition to the two recognized indicators for judging creativity that are *novelty* (the produced item should be different of what is known) and *quality* (the produced item must have value in its genre), Ritchie adds *typicality* to assess creativity (the produced item should be representative of its class). Moreover, the creative program may use an *inspiring set* that it is designed to recreate differently (as far as possible). The high-valued output artefacts of a creative program, which are not already in the inspiring set, constitute the *creative set* (Colton, Pease, and Ritchie 2001).

Maher (2010) advocates that evaluating creativity is a combination of the evaluation of three essential criteria for creativity: novelty, value, and *unexpectedness* (or surprise), where “unexpectedness is measured by how far the artifact is from the expected next artifact” (taking into account the sequential nature of expectation).

Schmidhuber (2010) takes a (reinforcement) learning perspective for modeling creative processes, where “the creativity reward of the action selector is the current surprise or novelty measured by the improvements of the world model”. The idea is to use reinforcement learning for “creating / discovering / learning novel patterns in the growing history of actions and sensory inputs, where the theory formally specifies what exactly is a pattern, what exactly is novel or surprising, and what it means to incrementally learn novel skills leading to more novel patterns”.

This is not to be confused with the use of modern machine learning tools for computational creativity, where, for instance, surprising images can be generated by a text-to-image model from prompts. In such a case, the creativity part relies in the (usually human-originated) prompt that specifies a requirement for an unusual *combination* of elements which are not encountered together in the real world.

Lastly let us mention some general ideas commonly associated with creativity. Creativity is often the result of a conceptual blending (Turner and Fauconnier 2003), which corresponds to the association of two different frames of reference that can be joined, connected together. Analogical reasoning, especially when making a parallel between two distinct universes through a structure mapping is often considered as a form of conceptual blending.

Lateral thinking (de Bono 1971), in problem solving,

amounts to move from a familiar model to a new, unexpected one. This is close to the idea of “thinking out of the box”, which itself sounds as a transformational operation for enlarging, modifying the conceptual space.

Potential measures of novelty

The previous survey has shown that everybody agrees with the idea that novelty is a central criterion in the evaluation of creativity. The quality, the value of the items produced seem more easy to maintain through constraints by the computer program. In this subsection we review two measures that have been used in practice for assessing novelty: the Hellinger distance and the Kullback-Leibler divergence.

These measures compare two probability distributions. Applied to creativity, one may consider that the first distribution describes the population of the items in the inspiring set and the second one the items produced by the system. These measures evaluate to what extent are the produced items dissimilar to the existing examples we start with.

Hellinger distance This distance is used in the assessment of creativity, e.g., in (Fele et al. 2022). It was first introduced in 1909 by Hellinger, a German mathematician. Given two discrete probability distributions $P = (p_1, \dots, p_n)$ and $Q = (q_1, \dots, q_n)$ over the same space $X = \{x_1, \dots, x_n\}$ (where $p_i = P(\{x_i\})$, $\sum_{i=1,n} p_i = 1 = \sum_{i=1,n} q_i$), their Hellinger distance is defined as

$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1,n} (\sqrt{p_i} - \sqrt{q_i})^2}$$

By an easy calculation, it follows

$$H^2(P, Q) = 1 - \sum_{i=1,n} \sqrt{p_i q_i}$$

As can be seen the Hellinger distance is related to the Euclidean distance. We have $0 \leq H(P, Q) \leq 1$. $H(P, Q) = 0$ if and only if $p_i = q_i$ for all i . $H(P, Q) = 1$ if and only if P and Q are mutually singular, i.e., $\forall i, p_i > 0 \Rightarrow q_i = 0$ and conversely $\forall i, q_i > 0 \Rightarrow p_i = 0$.

The quantity $\sum_{i=1,n} \sqrt{p_i q_i}$ is known as the Bhattacharyya coefficient, a well-known measure of similarity between two probability distributions P and Q . So $H(P, Q)$ is a measure of dissimilarity between these distributions.

Kullback-Leibler divergence This divergence (Kullback and Leibler 1951) is a measure of relative entropy that evaluates how much the probability distribution Q is different from the probability distribution P . It is defined by

$$KL(P||Q) = \sum_{i=1,n} p_i \ln \frac{p_i}{q_i} = \sum_{i=1,n} (p_i \ln p_i - p_i \ln q_i)$$

As a natural generalization of the Shannon entropy this measure has been characterized in different ways and applied in different fields. It is positive : $KL(P||Q) \geq 0$, the divergence is non-symmetric $KL(P||Q) \neq KL(Q||P)$, so it is not a distance. The classical use of this divergence is to judge to what extent Q is an acceptable approximation of P , in that perspective, the smaller the divergence, the better, the

more similar P and Q since $KL(Q||P) = 0$ if and only if $p_i = q_i$ for all i .

Besides, $KL(P||Q)$ is finite only if the support of P is contained within the support of Q ($\forall i, p_i > 0 \Rightarrow q_i > 0$; but if we have $\exists i, p_i > 0$ and $q_i = 0$, then $KL(P||Q)$ will be infinite. Thus, when $KL(P||Q) > 0$ (and not close to 0), it indicates that Q is dissimilar with respect to P .

The Kullback-Leibler divergence has been advocated as a measure of ‘‘Bayesian surprise’’ (Itti and Baldi 2009). It has been used to detect surprise in short narratives, showing that a change in the model’s prediction corresponds with a significant change in the participants’ rating (Chieppe, Sweetser, and Newman 2022). Moreover, the use of Kullback-Leibler divergence has enabled the retrieval of key authors and works that shape the broad outlines of literary change in a vast corpus of French novels from the XIXth and early XXth centuries (Barré and Poibeau 2023). The use of this divergence seems appropriate for assessing novelty.

Lastly, some authors have proposed to use algorithmic complexity as a basis for attributing creativity (McGregor 2007; Mondol and Brown 2021; Brown and Ventura 2022). Conditional Kolmogorov complexity is proposed for evaluating novelty. For value, the idea is that an artefact is of high quality if it represents the output of a significant amount of computational effort. Although theoretically appealing, the approach is not so easy to use in practice, even if approximate evaluations could be obtained in principle. We have therefore not retained this option for our experiment, in spite of the fact that analogical proportions can be also thought in terms of Kolmogorov complexity (Bayouhd, Prade, and Richard 2012; Prade and Richard 2018).

Analogical inference and creativity

We first recall what are analogical proportions and what is their Boolean modeling, before explaining how they can be used for inference and for creative purposes. We end this section with a theoretical study of their creative power.

Analogical proportions

Analogical proportions are statements linking four entities of the same kind. They are of the form ‘‘ a is to b as c is to d ’’. In principle, these entities can be anything: numbers, word, images, sentences, ... Analogical proportions appeared for the first time (at least in Occident) with Aristotle, who takes lessons from works by mathematicians of the time just before him on arithmetic proportions ($a - b = c - d$) and geometric proportions ($\frac{a}{b} = \frac{c}{d}$). These numerical proportions equalize differences and ratios respectively. Aristotle extended such a pattern to words (or concepts).

As can be seen, such proportions operate both a comparison inside pairs, and a comparison between pairs. This is precisely what state ‘‘ a is to b as c is to d ’’. Such a proportion is denoted by $a : b :: c : d$. As suggested by the parallel with numerical proportions, analogical proportions are supposed to obey three postulates:

- $a : b : a : b$ (*reflexivity*)
- $a : b :: c : d \Rightarrow c : d :: a : b$ (*symmetry*)

a	b	c	d
0	0	0	0
1	1	1	1
0	0	1	1
1	1	0	0
0	1	0	1
1	0	1	0

Table 1: Truth table of $a : b :: c : d$

- $a : b :: c : d \Rightarrow a : c :: b : d$ (*central permutation*)

While the two first postulates are not surprising, the stability under central permutation seems unexpected. However, if we consider the following example ‘‘puppies are to dogs as kittens are to cats’’, we should admit that one may also state that ‘‘puppies are to kittens as dogs are to cats’’.

The application of the symmetry and central permutation (in sequence), shows that an analogical proportion satisfies

- $a : a :: b : b$ (sameness);
- $a : b :: c : d \Rightarrow d : b :: c : a$ (external permutation);
- $a : b :: c : d \Rightarrow b : a :: d : c$ (internal reversal);
- $a : b :: c : d \Rightarrow d : c :: b : a$ (complete reversal).

One can easily check that the above properties make sense on the ‘‘cats / dogs’’ example.

The entities we consider are vectors of Boolean values. Thus, any entity is represented by a vector $\vec{a} = (a_1, \dots, a_n)$ where a_i is the Boolean value of a feature / attribute i . The analogical proportions between vectors (of the same size) is defined componentwise, namely

$$\vec{a} : \vec{b} :: \vec{c} : \vec{d} \text{ if and only if } \forall i \in \{1, \dots, n\}, a_i : b_i :: c_i : d_i.$$

We first deal with one component and four Boolean variables representing the value of some feature of interest for four distinct items.

The reflexivity postulate $a : b :: a : b$ forces a Boolean analogical proportion to be true for all a and all b , which entails that the four valuations $(0, 0, 0, 0)$, $(0, 1, 0, 1)$, $(1, 0, 1, 0)$, and $(1, 1, 1, 1)$ should correspond to true analogical proportions. Due to the central permutation postulate, the minimal Boolean model that satisfies the three postulates is true for the 6 following patterns for the Boolean variables a, b, c, d (and false for the $2^4 - 6 = 10$ other possible patterns) given in Table 1 (Prade and Richard 2018).

The following quaternary logical formula is true only for the 6 above patterns of values (and thus false for the 10 others) (Miclet and Prade 2009):

$$a : b :: c : d = ((a \wedge \neg b) \equiv (c \wedge \neg d)) \wedge ((\neg a \wedge b) \equiv (\neg c \wedge d)) \quad (1)$$

It precisely expresses that ‘‘ a differs from b as c differs from d , and b differs from a as d differs from c ’’ (and ‘‘when a and b do not differ, c and d do not differ’’).

Let us take the cat / dog example, ‘‘a puppy is to a dog as a kitten is to a cat’’ for illustrating the analogical proportion between vectors that represent the four animals on a set of 6 attributes, here mammal, bird, canid, feline, young, adult, as in Table 2.

As can be seen, the attributes in the example are such that each of them, vertically, illustrates one of the 6 patterns

	<i>mam.</i>	<i>bird</i>	<i>canid</i>	<i>felid</i>	<i>young</i>	<i>adult</i>
puppy	1	0	1	0	1	0
dog	1	0	1	0	0	1
kitten	1	0	0	1	1	0
cat	1	0	0	1	0	1

Table 2: A Boolean validation of *puppy:dog::kitten:cat*

in Table 1. It validates the fact the analogical proportion *puppy:dog::kitten:cat* holds.

Analogical inference

Given a triplet $a, b, c \in \{0, 1\}^3$, one may look for $x \in \{0, 1\}$ such that $a : b :: c : x$ holds true. From truth Table 1, it can be checked that if a solution x exists, it is unique, but there are two cases where the equation has no solution. Indeed, there is no $x \in \{0, 1\}$ such that $1 : 0 :: 0 : x$ and $0 : 1 :: 1 : x$ are true analogical proportions. It can be checked that the Boolean analogical equation $a : b :: c : x$ is solvable if and only if $(a \equiv b) \vee (a \equiv c)$ is true, where \equiv is the equivalence connective. In that case, the *unique* solution is given by $x = a \equiv (b \equiv c)$ (remember \equiv is associative); thus $x = b$ if $a = c$ and $x = c$ if $a = b$.

Analogical proportion-based inference, when applied to classification, relies on a simple principle: if four (Boolean) vectors \vec{a} , \vec{b} , \vec{c} and \vec{d} make a valid analogical proportion component-wise for each 4-tuple of values pertaining to the same attribute, then it is expected that their class labels also make a valid proportion³ (e.g., (Bounhas, Prade, and Richard 2017)).

$$\frac{\vec{a} : \vec{b} :: \vec{c} : \vec{d}}{cl(\vec{a}) : cl(\vec{b}) :: cl(\vec{c}) : cl(\vec{d})} \quad (2)$$

If the class labels for vectors \vec{a} , \vec{b} and \vec{c} are known (i.e., they belong to the sample set), the classification of a new Boolean vector \vec{d} is possible using (2) only if i) the equation $cl(\vec{a}) : cl(\vec{b}) :: cl(\vec{c}) : x$ is solvable, and ii) the analogical proportion $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ holds true on all components. If these two conditions are met, $cl(\vec{d})$ is the unique solution of $cl(\vec{a}) : cl(\vec{b}) :: cl(\vec{c}) : x$. In practice there are usually several triplets $(\vec{a}, \vec{b}, \vec{c})$ leading to several solutions for $cl(\vec{d})$, which may be different. Then a majority vote decides the class assigned to \vec{d} . Such a procedure yields good results on benchmarks (Bounhas, Prade, and Richard 2017).

In (Afantenos et al. 2024), a slightly different inference pattern has been proposed for creativity purposes. From three items \vec{a} , \vec{b} , \vec{c} , one can build another item \vec{d} , provided that for each attribute i the Boolean analogical equation $a_i : b_i :: c_i : x_i$ is solvable, which requires that

$$(\vec{a} \equiv \vec{b}) \vee (\vec{a} \equiv \vec{c})$$

holds true component-wise. This can be summarized by the following inference pattern

³Similarly to Boolean attribute values, the only valid analogical proportions between two class labels, possibly identical, s and t , are $s : t :: s : t$ and $s : s :: t : t$, while $s : t :: t : x$ has no solution (except for for $s = t = x$).

$$\frac{\vec{a}, \vec{b}, \vec{c}}{\vec{d} = (\vec{a} \equiv \vec{b} \equiv \vec{c})} \quad (3)$$

It can be shown that \vec{d} is different from $\vec{a}, \vec{b}, \vec{c}$, provided that $\vec{a} \neq \vec{b}$ and $\vec{a} \neq \vec{c}$. The next section studies how much different \vec{d} can be from the three other vectors.

A theoretical evaluation of analogical proportion-based creativity

An analogical proportion $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ establishes a parallel between the pairs (\vec{a}, \vec{b}) and (\vec{c}, \vec{d}) . The 6 patterns of Table 1 making an analogical proportion true enables us to partition the n attributes on which the items are described ($\vec{a} = (a_1, \dots, a_n)$, $\vec{b} = (b_1, \dots, b_n)$, $\vec{c} = (c_1, \dots, c_n)$, $\vec{d} = (d_1, \dots, d_n)$), into 3 subsets:

- i) the subset of attributes i where $a_i = b_i = c_i = d_i = 1$ or $a_i = b_i = c_i = d_i = 0$, i.e., where the four items $\vec{a}, \vec{b}, \vec{c}, \vec{d}$ are equal;
- ii) the subset of attributes j where $a_j = b_j$ and $c_j = d_j$ but $a_j \neq c_j$, i.e., where the items are equal within the pairs (\vec{a}, \vec{b}) and (\vec{c}, \vec{d}) , but not in the same way;
- iii) the subset of attributes k where $a_k = c_k$ and $b_k = d_k$, i.e., the subset of attribute(s) whose value(s) change(s), in the same direction, from \vec{a} to \vec{b} and from \vec{c} to \vec{d} .

This is pictured in Table 3.

items	All equal	Equal by pairs	Same Differences
\vec{a}	1 0	1 0	1 0
\vec{b}	1 0	1 0	0 1
\vec{c}	1 0	0 1	1 0
\vec{d}	1 0	0 1	0 1

Table 3: The 3 types of patterns of an analogical proportion

Note that the central permutation of \vec{b} and \vec{c} exchanges the last two columns in Table 3. It allows to state the result:

$$\vec{a} : \vec{b} :: \vec{c} : \vec{d} \Leftrightarrow \begin{cases} Equ(\vec{a}, \vec{b}) = Equ(\vec{c}, \vec{d}) \\ Dif^{10}(\vec{a}, \vec{b}) = Dif^{10}(\vec{c}, \vec{d}) \\ Dif^{01}(\vec{a}, \vec{b}) = Dif^{01}(\vec{c}, \vec{d}) \end{cases} \Leftrightarrow \begin{cases} Equ(\vec{a}, \vec{c}) = Equ(\vec{b}, \vec{d}) \\ Dif^{10}(\vec{a}, \vec{c}) = Dif^{10}(\vec{b}, \vec{d}) \\ Dif^{01}(\vec{a}, \vec{c}) = Dif^{01}(\vec{b}, \vec{d}) \end{cases}$$

where, given two vectors $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$,

$$\begin{aligned} Equ(\vec{x}, \vec{y}) &= \{i \mid x_i = y_i\} = Equ^0(\vec{x}, \vec{y}) \cup Equ^1(\vec{x}, \vec{y}), \\ Equ^0(\vec{x}, \vec{y}) &= \{i \mid x_i = y_i = 0\}, \\ Equ^1(\vec{x}, \vec{y}) &= \{i \mid x_i = y_i = 1\}, \\ Dif^{10}(\vec{x}, \vec{y}) &= \{i \mid x_i = 1, y_i = 0\}, \\ Dif^{01}(\vec{x}, \vec{y}) &= \{i \mid x_i = 0, y_i = 1\}. \end{aligned}$$

The Hamming distance between \vec{a} and \vec{b} , $Ham(\vec{a}, \vec{b})$, is equal to the cardinality of $Dif^{10}(\vec{a}, \vec{b}) \cup Dif^{01}(\vec{a}, \vec{b}) = Dif(\vec{a}, \vec{b})$. We have $Ham(\vec{a}, \vec{b}) = |Dif(\vec{a}, \vec{b})| \triangleq e$. Note that $Ham(\vec{a}, \vec{b}) = Ham(\vec{c}, \vec{d})$. Let $r = |\{i \text{ s.t. } a_i = b_i = c_i = d_i = 1 \text{ or } a_i = b_i = c_i = d_i = 0\}|$. Thus $Ham(\vec{c}, \vec{d}) = e$; $Ham(\vec{b}, \vec{d}) = n - e - r$; $Ham(\vec{a}, \vec{d}) = n - r$.

If $e = 0$, $\vec{a} = \vec{b}$ and $\vec{c} = \vec{d}$; if $r + e = n$, $\vec{a} = \vec{c}$, $\vec{b} = \vec{d}$. In both cases, the proportion $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ is trivial: we get $\vec{a} : \vec{a} :: \vec{b} : \vec{b}$ or $\vec{a} : \vec{b} :: \vec{a} : \vec{b}$. If $r = n$, the 4 items are equal.

Thus \vec{d} is all the more distant from \vec{a} and \vec{c} as r is small and e is large. But if $r = 0$, $e = n$, we get $\text{Ham}(\vec{b}, \vec{d}) = 0$. So we have to equally (or almost equally, if n is odd) balance e and $n - e$, with $r = 0$, if we want to maximize the distance of \vec{d} to $\{\vec{a}, \vec{b}, \vec{c}\}$ equal to $\min(e, n - e - r, n - r)$.

If n is even this maximal distance is just $n/2$. So the vector \vec{d} , entirely obtained from $\vec{a}, \vec{b}, \vec{c}$ can be very different from these three items, thus quite novel with respect to them.

Thus the analogical inference (3) appears to be a controlled manner to explore a universe of potential possibilities from existing options. This is better than a random selection (as early proposed in (Kaufmann 1969)).

Framework of the experiment

In the following we report an experiment illustrating the effectiveness of the creative power of analogical proportion-based inference. This is a two-step process: first prompts are generated, and then texts are produced. The novelty is checked both at the prompt level and at the text level. We first present our source of inspiration for the experiment, before explaining how analogical proportion-based inference is used for creating prompts, in turn used for generating texts.

The source of inspiration The source of inspiration is a well-known book by Queneau, *The Exercises in Style* (Queneau 1947), where the author rewrites the same story 99 times, a short text where the narrator gets on a bus, witnesses a dispute between a man and another passenger, and then sees this person again two hours later getting advice on adding a button to his overcoat. Each time the story is written in a new style. This can be viewed as an example of (repeated) exploratory creativity in Boden’s sense.

Queneau is also famous for another book (Queneau 1961) published later, that exemplifies combinational creativity. This book is based on ten sonnets (a poem of 14 verses) where each line is printed on a separate strip of thick paper. The ten sonnets have the same rhyme scheme and the same rhyme sounds. Thus any line from a sonnet can be combined with a line from the nine others, allowing for 10^{14} (= 100,000,000,000,000) different poems. This then led Queneau, together with a mathematician, Le Lionnais, to start the “Ouvroir de littérature potentielle” (Oulipo) gathering a group of people interested in experimental literature.

Creating prompts Queneau, in his book *The Exercises in Style* (Queneau 1947) varies the style each time he rewrites the story, making use of a wide variety of rhetorical devices. For our experiment, we transpose the idea, by choosing a limited number of features that can apply to a text. We use 10 Boolean features corresponding to the presence or absence of the following 10 emotions: joy, fear, anger, tranquility, nostalgia, shame, confusion, curiosity, love, and res-

ignation.⁴ This has the advantage of providing vectors that: i) can be processed by analogical inference; and ii) can be associated with prompts sent to an LLM-based chatbot able to generate a constrained text.

For building an inspiring set of 100 vectors (with 10 components) in our experiment, we start with an initial set of 11 vectors: these vectors are the vector (0, 0, 0, 0, 0, 0, 0, 0, 0, 0) and the 10 vectors with a single 1 (and 9 ‘0’) corresponding to each feature. There are $2^{10} = 1024$ distinct Boolean vectors with 10 components. In order to obtain the $100 - 11 = 89$ missing vectors of the inspiring set, we proceeded as follows. To a set of vectors we can associate a probability distribution describing the probability of obtaining any combination of feature values using this set as a sample set. For completing the inspiring set we add vectors one by one to the current set. The principle is to use the probability distribution of the current set and to add a vector (not yet in the set) such that the Shannon entropy of the resulting set is minimized. This procedure stops when the current set includes 100 vectors. This is our inspiring set. Let us call it \mathcal{I} . Note that \mathcal{I} is rather “homogeneous” in the sense that \mathcal{I} gathers vectors with a small number of ‘1’.

From this set \mathcal{I} we generate an output set \mathcal{O} by means of analogical inference. In order to do so, we use triplets of vectors $(\vec{a}, \vec{b}, \vec{c})$ in \mathcal{I}^3 in order to apply the pattern of inference (3). We apply a “semi-guided” procedure by selecting the candidate pair (\vec{a}, \vec{b}) then used to form the triplets, as explained in the next paragraph. Then we generate all the possible outputs \vec{d} that can be obtained from a selected pair and a compatible vector \vec{c} (i.e., a vector \vec{c} for which the proportion equation $\vec{a} : \vec{b} : \vec{c} : x$ has a solution), by applying the analogical inference (3). This forms \mathcal{O} .

The selection of pairs (\vec{a}, \vec{b}) makes sure that the number of ‘1’s in \vec{b} increases with respect to \vec{a} : this means that $\forall i, a_i = 1 \Rightarrow b_i = 1$ and $\exists j, a_j = 0$ and $b_j = 1$. The idea is that if when going from \vec{a} to \vec{b} a property is gained, it will also be gained when going from \vec{c} to \vec{d} (since 01 can be only completed by 01 for making a valid proportion). Moreover the pattern 1010 corresponding for \vec{b} and \vec{d} to the repeated loss of a property, is then forbidden. Still it does not mean that \vec{d} inherits all the properties of \vec{b} since 11 can be completed by 11 but also by 00 for obtaining a valid analogical proportion pattern. This last point shows that the analogical mechanism is not a pure property enhancement procedure, but it also manages to build a renewed set of outputs from the inspiring set, even if we privilege particular pairs and we forbid the use of pattern 1010 in the inference. Altogether, we favor a collection of outputs made of new enriched vectors, thus distinct and independent from the initial collection.

Starting with a set \mathcal{I} of size 100, the analogical inference yields a set \mathcal{O} of size approximately 150. Each of these two

⁴This choice may look a bit arbitrary, even if it substantially overlaps the 16 basic emotions advocated in (Ekman 1999). Note that here it is not so much a matter of emotions experienced by people, but just a text that is expected to communicate some feelings to the reader. The number (10) was chosen as being large enough for having variety, but not too large for limiting computation cost.

sets can be associated with a probability distribution. We are then in position to compute the dissimilarity or the divergence between the distributions $p_{\mathcal{I}}$ and $p_{\mathcal{O}}$ associated with \mathcal{I} and \mathcal{O} respectively.

Exploiting prompts The vectors in the sets \mathcal{I} and \mathcal{O} can be used as a basis for building prompts to be processed by a Large Language Model (LLM) for generating a text associated with each vector. Due to the one-to-one correspondence between the elements of the vectors and the presence or absence of an emotion, prompt construction is carried out by pre-defining a text containing ten variables to be instantiated. These variables are replaced by the strings “true” or “false” based on the binary value of the corresponding element. An example of a prompt is presented on page 7. Submitting the prompt to a language model subsequently generates a text associated with each vector.

The texts produced can be lemmatized and n -grams can be extracted (for fixed values of n) in order to compute a probability distribution of the n -grams for the texts generated from \mathcal{I} , and another probability distribution of the n -grams for the texts generated from \mathcal{O} , in order to test at the text level if some novelty can be detected in the texts issued from \mathcal{O} wrt those issued from \mathcal{I} .

For the text generation, we used LLaMA 3. LLaMA 3 is a latest-generation language model developed by Meta AI, designed to push back the boundaries of natural text generation thanks to a sophisticated architecture and a high number of parameters (8 billions). This model is part of a series of LLaMA (Large Language Model Meta AI) language models distinguished by their ability to understand and generate text in a coherent and contextually appropriate way.

As a warning, we want to emphasize that our aim was not to test the quality of the texts that LLaMA 3 can produce in any way. LLaMA 3 is just used here as a tool for producing short texts that are constrained to follow the story line of an unremarkable story—freely inspired from the story Queneau starts with, in *The Exercises in Style*—which includes or not a series of 10 emotions, in order to have two collections of texts associated with the collections of vectors in \mathcal{I} and \mathcal{O} respectively. The aim is to check if the potential novelty at the vector level is reflected or not at the text level. The story line is specified to LLaMA 3 through the following exchange:

user: I will give you a story and you’ll have to tell me the main route of the story in a paragraph (don’t jump lines keep it all together) okay?

assistant: Sounds like fun! I’m ready to hear your story. Please go ahead and share it, and I’ll do my best to summarize the main route of the story in one paragraph. Go!

user: An individual boards a public transportation vehicle. His figure blends into the mass of anonymous passengers. His gaze wanders, observing the familiar faces distantly. Suddenly, a peculiarity catches his attention: a man adorned with an unusual headgear, a felt hat of austere gray, perched majestically atop his head. And there, an irregularity, a disturbance in the usual landscape: a fly, intrepid explorer, ambles nonchalantly on the nape of this impassive man. A

shiver of curiosity runs through him. A question nags at his mind: should he intervene? A decision takes shape in the meanders of his thought. Does the man, this bearer of felt, deserve his attention? Is it a mere accident or a sign from destiny? An irrepressible impulse propels him to act. He rises, approaches with determination. A brief, incisive exchange, punctuated by abrupt gestures, suffices to draw the attention of the man, but in a markedly different register. A reprimand escapes his lips, cutting, imbued with thinly veiled arrogance. And the fly? It continues its journey, indifferent to this human theater, oblivious to its involuntary role in this fleeting scene of daily life

assistant: As an individual boards a public transportation vehicle, he becomes distracted by a peculiar sight: a man wearing an unusual gray felt hat and sporting an intrepid fly on his neck. The individual’s curiosity is piqued, and he begins to wonder if the situation warrants intervention. As he rises from his seat and approaches the man, he delivers a reprimanding remark laced with arrogance, drawing the attention of the hat-wearing individual in a completely different register than anticipated. Meanwhile, the fly remains oblivious to its role in this brief and fleeting drama that plays out amidst the daily routines of the anonymous passengers.

Thus the prompt corresponding to the initial vector $(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ is

user: Redo the main route you mentioned before transforming it in a way that dimension Joy is false, Fear is false, Anger is false, Tranquility is false, Nostalgia is false, Shame is false, Confusion is false, Curiosity is false, Love is false, Resignation is false. (don’t write any of this words in the paragraph (Joy, Fear, Anger, Tranquility, Nostalgia, Shame, Confusion, Curiosity, Love, Resignation) and don’t write (True or false) either).

Experimental results

First of all, the comparison of the collections \mathcal{I} and \mathcal{O} in the experiment yielded the following results in terms of Hellinger distance and of Kullback-Leibler divergence:⁵

$$H(p_{\mathcal{I}}, p_{\mathcal{O}}) = 0.8625; \quad KL(p_{\mathcal{I}} || p_{\mathcal{O}}) = 0.6778$$

As can be seen, the novelty of the output collection \mathcal{O} with respect to the inspiring set \mathcal{I} is clearly assessed by the high value of Hellinger distance (whose maximum is 1). This is confirmed by the divergence evaluation. However note that here the vectors \vec{d} in \mathcal{O} do not necessarily maximize their difference with the vectors $\vec{a}, \vec{b}, \vec{c}$ in \mathcal{I} from which they are obtained, in the sense of subsection .

The text produced by LLaMA 3 from the prompt associated with the initial vector $(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ is

assistant: As the individual steps onto the public transportation vehicle, his gaze drifts across the familiar faces, but one peculiarity stands out: a man clad in an austere gray felt hat, with an intrepid explorer of a different kind - a fly - perched

⁵Code is available here: <https://github.com/Sasufox/AnalogyStories>

nonchalantly on the nape of his neck. The individual's thoughts begin to weave a tapestry of inquiry, and he rises from his seat with purpose, approaching the hat-wearer with determination. A brief, incisive exchange ensues, punctuated by swift gestures, as the two individuals intersect in a fleeting dance of daily life.

This text is quite short, and is devoid of any particular emotion. As an example here is the text produced when Confusion, Curiosity, and Resignation are required (the prompt corresponds to vector (0, 0, 0, 0, 0, 1, 1, 0, 1)):

As he boards the public transportation vehicle, his gaze drifts across the sea of anonymous faces, his mind a maze of contradictions. A peculiar sight captures his attention: a man's austere gray hat, a fly's leisurely stroll on the nape of his neck. His thoughts twist and turn, unsure whether to intervene or simply observe. The air is thick with uncertainty as he approaches the man, his words laced with an unexpected tone that belies the calm exterior. Meanwhile, the fly continues its aimless flight, seemingly unaware of the complex emotions swirling around it, leaving only a lingering sense of disorientation in its wake.

As we can see the writing of the new text has been transformed, keeping the same story line. As expected, the collection of texts associated to \mathcal{O} differs from the collection associated to \mathcal{I} . This can be checked using Hellinger's distance (which has the advantage to be valued on a bounded scale) between the probability distributions of n -grams in the two collections. The results obtained for $n = 2$, $n = 3$, $n = 4$ are:

$$H(p_{\mathcal{I}}^{2-gram}, p_{\mathcal{O}}^{2-gram}) = 0.3947; \quad H(p_{\mathcal{I}}^{3-gram}, p_{\mathcal{O}}^{3-gram}) = 0.6279; \quad H(p_{\mathcal{I}}^{4-gram}, p_{\mathcal{O}}^{4-gram}) = 0.7690.$$

This confirms the novelty of texts in \mathcal{O} with respect to those in \mathcal{I} . However, it has been noted (Chen and Ding 2023) that advanced large language models obtain very good results on divergent semantic association tasks, suggesting a capacity for creativity. Since the production of texts by LLMs is based on a stochastic mechanism, one may measure the distance between stories generated from the same vector, in the following way. From a representative sample (of size 280) of the 1024 vectors, 100 texts are generated for each vector. For each of these collections, a text is taken as a reference and compared with the other texts. The Hellinger distance between the collection of reference texts and the collection of other texts can then be calculated. The results suggest that LLaMA generates a large number of n -grams that are not present simultaneously in two documents generated from the same vector. Thus the stochastic nature of LLM makes it difficult to determine exactly how much novelty is brought about by the use of a new prompt.

Final discussion

In the introduction we have distinguished three kinds of analogical reasoning: i) the one, sometimes termed 'conceptual blending' that brings together two situations belonging to different universes, as in the analogy between the atom and the solar system; ii) case-based reasoning that extrapolates to a new situation what was observed in a similar case; iii) the solving of analogical proportion-based equation that pro-

duces a fourth item from three others, as in analogy puzzles. The first kind corresponds to transformational analogy, the second kind can hardly be called creative since it's essentially a question of adapting something already known, the last kind is a form of exploratory analogy, where a universe (the one describable by means of the attributes used for representing the considered items) is explored.

Unlike a purely random path through the universe of Boolean features, this universe can be explored by means of the analogical proportion-based inference in a controlled way, for instance by favoring the accumulation of valuable properties. Generally speaking, the analogical proportions $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ make sure that the differences between \vec{a} and \vec{b} are preserved between \vec{c} and \vec{d} . The experiment has shown that even if the inspiring set (where $\vec{a}, \vec{b}, \vec{c}$ are taken) is made of items that are relatively close of each other, the novelty of \vec{d} can be established in terms of Hellinger distance. In case vectors would describe options at a meta level, the inference mechanism might be used for transformational creativity (viewed as exploratory creativity at the meta-level).

In our experiment, we have used Shannon entropy minimization in order to build an inspiring set that is not too disparate. That does not prevent the analogical inference mechanism to produce novel vectors clearly outside the inspiring set. It is clear that the choice of the inspiring set has an influence on the outputs that can be produced. The use of selection procedures on top of the inference mechanism in order to start with "inspiring" items is an important issue to further explore in future research. Other constraints could be considered for building the inspiring set.

The experiment has illustrated how the vectors generated can be the basis of prompts addressed to an LLM-based chatbot, leading to the production of "constrained" texts where the novelty inherited from the novelty of the vector generated is somewhat preserved. As already suggested in (Afantenos et al. 2024), analogical proportion-based inference can be used as a tool for proposing novel items. The combination of analogical reasoning and LLM-based capabilities to produce text constitutes a kind of creative device.

Queneau's *The Exercises in Style* provides a very good example of exploratory creativity, like his *A Hundred Thousand Billion Poems* (Queneau 1961) perfectly exemplifies the idea of combinational creativity, in the literary domain. However, the remarkable creativity of Queneau is at the conception level of these two books, i.e., in the device itself. Interestingly enough the h-creativity of Queneau for *The Exercises in Style* is perhaps debatable, since Desiderius Erasmus in his 1512 educational rhetoric textbook *De Copia* instructed on how to give variety to writing; thus in the 33rd chapter he gave 195 variations on the same (Latin) sentence whose meaning was "Your letter delighted me greatly." Queneau was also highly interested by one of the earliest known published volumes of computer poetry (Baudot 1964) where creativity was combinational and the combination of words, belonging to a restricted vocabulary, aleatory. This is shown by his commentaries gathered at the end of the book among many other reactions by writers and linguists. In particular, he wondered about the possibility for humans to discern

sentences produced by the machine from sentences written, e.g., Surrealists poets – a Turing-like test!

Conclusion

The paper has focused on the creative ability of analogical proportion-based inference. This paper has made several contributions: first, we have highlighted and illustrated the creative power of the analogical proportion-based inference. Second, the assessment of the novelty of what is produced has been confirmed in the experiment by the divergence measures applied to the generated prompts. Notably, Hellinger distance that takes its values in a bounded interval seems especially appropriate. Moreover, the experiment also showed the possibility of combining analogical inference and LLM-based chatbots, while preserving creative power.

Besides, several lines for further research would be worth investigating. First, the analogical mechanism used is only based on Boolean proportions. It might be of interest to use graded analogical proportions extended to numerical data (Dubois, Prade, and Richard 2016) or fuzzy logic (Colombo, D’Onofrio, and Portmann 2020), for, e.g., introducing a graded view of emotions in our experiment. In addition, in this paper we have focused on the assessment of the novelty of what is produced, but one may also check the quality of the texts obtained by estimating if the texts effectively contain words in relation with the emotions required by the prompts, using an appropriate lexicon (Mohammad 2018). Finally, one may like to compare the results obtained using analogically generated prompts with a direct request of generating texts on the basis of triplets of examples. In the absence of direct experimentation, any speculation regarding the results that would be obtained remains hypothetical, especially since the triplet of examples used can significantly influence the outcome. However, it should be noted that our methodology constrains the large language model in leveraging all terms located within a specific region of the embedding space associated with an emotion, while increasing diversity, as evidenced by the increase in the Hellinger distance.

The experimental process described in this paper should be viewed as a tool to aid creation rather than as a creative machinery in itself. Indeed, just as Queneau’s creativity in *Exercices de style* lies primarily in the idea of producing multiple versions of a mundane text—acting as a reference constraint while simultaneously employing different stylistic constraints—here, the user decides which parameters to vary in the texts to be generated. They then use analogical inference to explore the space of possibilities to determine the constraint specifying the parameters that will induce the prompt to be applied to the LLM.

Acknowledgments

This research was supported by the ANR project “Analogies: from theory to tools and applications” (AT2TA), ANR-22-CE23-0023.

References

- Aamodt, A., and Plaza, E. 1994. Case-based reasoning; foundational issues, methodological variations, and system approaches. *AICom*, 7, 39-59.
- Afantenos, S.; Prade, H.; Richard, G.; and Bernardes, L. 2024. Analogical proportions and creativity: A preliminary study. In Grace, K.; Llano, M. T.; Martins, P.; and Hedblom, M. M., eds., *Proc. 15th Int. Conf. on Computational Creativity (ICCC’24)*, Jönköping, 289–298.
- Barbot, N.; Miclet, L.; and Prade, H. 2019. Analogy between concepts. *Artif. Intell.* 275:487–539.
- Barré, J., and Poibeau, T. 2023. Beyond canonicity. modeling literary change in french novelist production: Canon vs. archive. In Sela, A.; Jannidis, F.; and Romanowska, I., eds., *Proc. Computational Humanities Research Conf., Paris, Dec. 6-8*, CEUR Workshop Proc., 3558, 814-830.
- Baudot, J. 1964. *La Machine à Ecrire mise en marche et programmée par Jean A. Baudot*. Editions du Jour, coll. les poètes du jour, Montréal. ‘Le premier recueil de vers libres rédigés par un ordinateur électronique’, 96p.
- Bayouhd, M.; Prade, H.; and Richard, G. 2012. Evaluation of analogical proportions through Kolmogorov complexity. *Knowl. Based Syst.* 29:20–30.
- Boden, M. A., ed. 1994. *Dimensions of Creativity*. MIT Press.
- Boden, M. A. 2003. *The Creative Mind - Myths and Mechanisms (2. ed.)*. Routledge. 1st ed., Basic Books, 1990.
- Bounhas, M.; Prade, H.; and Richard, G. 2017. Analogy-based classifiers for nominal or numerical data. *Int. J. Approx. Reason.* 91:36–55.
- Brown, D. G., and Ventura, D. 2022. Ethics, aesthetics and computational creativity. In Hedblom, M. M.; Kantosalo, A. A.; Confalonieri, R.; Kutz, O.; and Veale, T., eds., *Proc. 13th Int. Conf. on Computational Creativity (ICCC’22)*, Bozen-Bolzano, June 27 - July 1, 150–158.
- Buchanan, B. 2001. Creativity at the meta-level. *AI Magazine* 22:13–28.
- Bundy, A. 1994. What is the difference between real creativity and mere novelty? *Behavioral and Brain Sciences* 17:533–534.
- Chen, H., and Ding, N. 2023. Probing the creativity of large language models: Can models produce divergent semantic association? *CoRR* abs/2310.11158. Published in Findings of EMNLP 2023.
- Chieppe, P.; Sweetser, P.; and Newman, E. 2022. Bayesian modelling of the well-made surprise. In Hedblom, M. M.; Kantosalo, A. A.; Confalonieri, R.; Kutz, O.; and Veale, T., eds., *Proc. 13th Int. Conf. on Computational Creativity (ICCC’22)*, Bozen-Bolzano, Jun.27-Jul. 1, 126–135.
- Colombo, M.; D’Onofrio, S.; and Portmann, E. 2020. Integration of fuzzy logic in analogical reasoning: A prototype. In *Proc. 16th IEEE Int. Conf. Intel. Comp. Communic. and Proc. (ICCP’20)*, Cluj-Napoca, Sept. 3-5, 5–11.
- Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In Raedt, L. D.; Bessiere, C.;

- Dubois, D.; Doherty, P.; Frascioni, P.; Heintz, F.; and Lucas, P. J. F., eds., *Proc. 20th Europ. Conf. on Artificial Intelligence (ECAI'12), Montpellier, Aug. 27-31*, volume 242 of *Frontiers in Artif. Intellig.and Appl.*, 21–26. IOS Press.
- Colton, S.; Pease, A.; and Ritchie, G. D. 2001. The effect of input knowledge on creativity. In *Proc. Workshop on Creative Systems: Approaches to Creativity in AI and Cognitive Science*, at 4th Int. Conf. on Case-Based Reasoning, Vancouver, July 31.
- Cook, M., and Colton, S. 2011. Automated collage generation - With more intent. In Ventura *et al.*, D., ed., *Proc. 2nd Int. Conf. on Computational Creativity, Mexico City, April 27-29*, 1–3.
- Correa Beltran, W.; Prade, H.; and Richard, G. 2016. Constructive solving of raven's IQ tests with analogical proportions. *Int. J. Intell. Syst.* 31(11):1072–1103.
- de Bono, E. 1971. *The Use of Lateral Thinking*. Avon Books.
- de Mántaras, R. L. 1998. It don't mean A thing (if it ain't got that swing). In Prade, H., ed., *Proc. 13th Europ. Conf. on Artificial Intelligence, Brighton, Aug. 23-28*, 694–696. John Wiley and Sons.
- Dreistadt, R. 1969. The use of analogies and incubation in obtaining insights in creative problem solving. *The J. of Psychology* 71:159–175.
- Dubois, D.; Prade, H.; and Richard, G. 2016. Multiple-valued extensions of analogical proportions. *Fuzzy Sets Syst.* 292:193–202.
- Ekman, P. 1999. Basic emotions. In Dalglish, T., and Power, M. J., eds., *Handbook of cognition and emotion*. John Wiley Sons Ltd. 45–60.
- Falkenhainer, B.; Forbus, K. D.; and Gentner, D. 1989. The structure-mapping engine: algorithm and examples. *Artif. Intell.* 41(1):1–63.
- Fele, B.; Babic, J.; Pollak, S.; and Znidarsic, M. 2022. Evaluation of curriculum learning algorithms using computational creativity inspired metrics. In Hedblom, M. M.; Kantosallo, A. A.; Confalonieri, R.; Kutz, O.; and Veale, T., eds., *Proc. 13th Int. Conf. on Computational Creativity (ICCC'22), Bozen-Bolzano, June 27 - July 1*, 364–373.
- Fuchs, B.; Lieber, J.; Miclet, L.; Mille, A.; Napoli, A.; Prade, H.; and Richard, G. 2020. Case-based reasoning, analogy, and interpolation. In *A Guided Tour of Artificial Intelligence Research: Vol. I: Knowledge Representation, Reasoning and Learning*. Springer. 307–339.
- Gentner, D. 1983. Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7(2):155–170.
- Gust, H.; Kühnberger, K.; and Schmid, U. 2006. Metaphors and heuristic-driven theory projection (HDTP). *Theo. Comput. Sci.* 354(1):98–117.
- Halpern, D. 1986. Analogies as a critical thinking skill. In Berger, D.; Pezdek, K.; and Banks, W., eds., *Applications of Cognitive Psychology: Problem Solving, Education, and Computing*. 75–86, Routledge.
- Hesse, M. 1959. On defining analogy. *Proceedings of the Aristotelian Society* 60:79–100.
- Hummel, J. E., and Holyoak, K. J. 1997. Distributed representations of structure: a theory of analogical access and mapping. *Psychological Review* 104(3):427–466.
- Itti, L., and Baldi, P. 2009. Bayesian surprise attracts human attention. *Vision Research* 49:1295–1306.
- Kaufmann, A. 1969. L'imagination artificielle - (heuristique automatique). *Revue française d'Automatique, d'Informatique et de Recherche Opérationnelle (R.A.I.R.O.) - Recherche opérationnelle* 3(3):5–24.
- Kolodner, J. L. 1993. *Case-Based Reasoning*. Morgan Kaufmann.
- Kullback, S., and Leibler, R. A. 1951. On information and sufficiency. *Annals of Math. Stat.* 22:79–86.
- Kunda, M.; McGreggor, K.; and Goel, A. K. 2013. A computational model for solving problems from the Raven's Progressive Matrices intelligence test using iconic visual representations. *Cognitive Systems Research* 22:47–66.
- Maher, M. L. 2010. Evaluating creativity in humans, computers, and collectively intelligent systems. In Christensen, B. T.; Kristensen, T.; and Boztepe, S., eds., *Proc. 1st DESIRE Network Conf. on Creativity and Innovation in Design, Aarhus, Aug. 16-17, 22–28*. ACM.
- McCormack, J., and d'Inverno, M., eds. 2012. *Computers and Creativity*. Springer.
- McGregor, S. 2007. Algorithmic information theory and novelty generation. In *Proc. 4th Int. Joint Workshop on Computational Creativity, London, June 17-19*, 109–112.
- Miclet, L., and Prade, H. 2009. Handling analogical proportions in classical logic and fuzzy logics settings. In *Proc. 10th Eur. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09), Verona*, 638–650. Springer, LNCS 5590.
- Mohammad, S. 2018. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In Gurevych, I., and Miyao, Y., eds., *Proc. 56th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers)*, 174–184.
- Mondol, T., and Brown, D. G. 2021. Incorporating algorithmic information theory into fundamental concepts of computational creativity. In de Silva Garza, A. G.; Veale, T.; Aguilar, W.; and y Pérez, R. P., eds., *Proc. 12th Int. Conf. on Computational Creativity, (ICCC'21), Mexico City, (virtual), Sept. 14-18*, 173–181.
- Polya, G. 1954. *Mathematics and Plausible reasoning. Volume 1: Induction and Analogy in Mathematics*. Princeton University Press.
- Prade, H., and Richard, G. 2018. Analogical proportions: From equality to inequality. *Int. J. Approx. Reason.* 101:234–254.
- Queneau, R. 1947. *Exercices de Style*. Gallimard. Exercises in Style, tr. B. Wright, Gaberbocchus Press, 1958.
- Queneau, R. 1961. *Cent mille milliards de poèmes*. Gallimard. Afterword: F. LeLionnais. Trans.: S. Chapman, A Hundred Thousand Billion Poems.

- Ramalho, G. L., and Ganascia, J. 1994. Simulating creativity in jazz performance. In Hayes-Roth, B., and Korf, R. E., eds., *Proc. 12th National Conf. on Artificial Intelligence (AAAI'94)*, Seattle, Jul. 31 - Aug. 4, Volume 1, 108–113. AAAI Press / The MIT Press.
- Rhodes, M. 1961 An analysis of creativity *PhiDelta Kappan* 42:305–310.
- Ritchie, G. 2001. Assessing creativity. In *Proc. of the AISB Symp. on Artificial Intelligence and Creativity in Arts and Science, York*, 3–11.
- Ritchie, G. 2006. The transformational creativity hypothesis. *New Generation Computing* 24:241–266.
- Ritchie, G. 2007. Some empirical criteria for attributing creativity to a computer program. *Minds Mach.* 17:67–99.
- Ritchie, G. 2012. A closer look at creativity as search. In Maher, M. L.; Hammond, K. J.; Pease, A.; y Pérez, R. P.; Ventura, D.; and Wiggins, G. A., eds., *Proc. 3rd Int. Conf. on Computational Creativity (ICCC'12), Dublin, May 30 - June 1*, 41–48.
- Santo, L. E.; Wiggins, G. A.; and Cardoso, A. 2024. Towards a formal creativity theory: Preliminary results in novelty and transformativeness. *CoRR* abs/2405.02148.
- Schank, R., and Kass, A. 1990. Explanations, machine learning, and creativity. In Kodratoff, Y., and Michalski, R. S., eds., *Machine learning: an artificial intelligence approach, volume III*. Morgan Kaufmann Publishers, San Francisco. 31–59.
- Schmidhuber, J. 2010. Formal theory of creativity, fun, and intrinsic motivation (1990-2010). *IEEE Trans. on Autonomous Mental Development* 2(3):230–247.
- Schmidt, M.; Krumnack, U.; Gust, H.; and Kühnberger, K. 2014. Heuristic-driven theory projection: An overview. In Prade, H., and Richard, G., eds., *Computational Approaches to Analogical Reasoning: Current Trends*, Studies in Comput. Intelligence, 548, 163-194. Springer.
- Turner, M., and Fauconnier, G. 2003. *The Way We Think : Conceptual Blending and the Mind's Hidden Complexities*. Basic Books.
- Van de Cruys, T. 2020. Artificial intelligence and literature. In Marquis, P.; Papini, O.; and Prade, H., eds., *A Guided Tour of Artificial Intelligence Research: Volume III: Interfaces and Applications of Artificial Intelligence*. Springer. 487–501.
- Veloso, M. M. 1993. Prodigy/analogy: Analogical reasoning in general problem solving. In Wess, S.; Althoff, K.; and Richter, M. M., eds., *Topics in Case-Based Reasoning, 1st Europ. Workshop (EWCBR'93), Kaiserslautern, Nov. 1-5, Selected Papers*, LNCS, 837, 33-52. Springer.
- Wiggins, G. A., and Forth, J. 2018. Computational creativity and live algorithms. In McLean, A., and Dean, R. T., eds., *The Oxford Handbook of Algorithmic Music*. Oxford University Press.
- Wiggins, G. A. 2006a. A preliminary framework for description, analysis and comparison of creative systems. *Knowl.based Syst.*19(7):449–458.
- Wiggins, G. A. 2006b. Searching for computational creativity. *New Gener. Comput.* 24(3):209–222.
- Wiggins, G. A. 2019. A framework for description, analysis and comparison of creative systems. In Veale, T., and Cardoso, F. A., eds., *Computational Creativity - The Philosophy and Engineering of Autonomously Creative Systems*. Springer. 21–47.