

Computational Creativity Conceptualisation Grounded on ICCC Papers

Senja Pollak¹, Biljana Mileva Boshkoska^{1,3}, Dragana Miljkovic¹, Geraint A. Wiggins², Nada Lavrač^{1,4}

¹Dept. of Knowledge Technology, Jožef Stefan Institute, Ljubljana, Slovenia

²Computational Creativity Lab, Queen Mary University of London, London E1 4NS, UK

³ Faculty of Information Studies, Novo Mesto, Slovenia

⁴ University of Nova Gorica, Nova Gorica, Slovenia

{senja.pollak,biljana.mileva,dragana.miljkovic,nada.lavrac}@ijs.si; geraint.wiggins@qmul.ac.uk

Abstract

In information science, it is considered that domain conceptualization can be realized by (one or several) ontologies. This paper presents a method of semi-automated domain conceptualization, where the domain of interest is Computational Creativity (CC). Grounded on papers, which were published in six consecutive years since 2010 in the Proceedings of International Conferences on Computational Creativity (ICCC), this paper proposes a tentative conceptualization of the CC domain. Some additional properties of the CC domain are studied, analysed by means of fully mechanical or semi-automated information extraction and dependency analysis techniques. This approach affords an interesting opportunity for automated historiography of a research field.

Introduction

As a sub-field of Artificial Intelligence research, *Computational Creativity* (CC) is concerned with engineering software that exhibits behaviours which would reasonably be deemed creative (Wiggins, 2006; Colton and Wiggins, 2012). A part of CC research addresses *Concept Creation Technology*, concerned with engineering software that exhibits creative conceptualization behaviour.

For a given domain, whose conceptual space (Boden, 2004) is closed, pre-defined and yet unexplored, it is interesting to study computational means for automated (or semi-automated) domain conceptualization. In the current research, we use the term *conceptualization* in alignment with its standard use in information science: a *conceptualization* is defined as “an abstract (simplified) view of some selected part of the world, containing the objects, concepts, and other entities that are presumed of interest for some particular purpose and the relationships between them.” Domain conceptualization is, in information science, frequently realized by manually defining (one or several) ontologies formally describing the domain of interest (Gruber, 1993; Smith, 2003).

Manual construction of ontologies represents a significant investment of human resources when used for modelling a new domain. Therefore, methods for (semi-)automated extraction of domain knowledge from unstructured texts were developed, including automated taxonomy construction described by Velardi, Faralli, and Navigli (2013).

While an ontology is a “formal, explicit specification of a shared conceptualization” (Gruber, 1993), represented as a set of domain concepts and the relationships between

them, a so-called *topic ontology* is a set of domain topics or concepts—formed of related documents—represented by the most characteristic topic keywords and related by the *subconcept-of* relationship (Fortuna, Grobelnik, and Mladenić, 2007). The task addressed in this paper is semi-automated construction of a topic ontology from documents in the area of computational creativity.

CC domain conceptualization has not been substantially addressed in the CC literature. Jordanous and Keller (2012) used automated natural language processing methods and a statistical measure of association to identify words related to creativity (in general, not specifically CC). They clustered the words into semantically-related groups by using a lexical similarity measure, resulting in an ontology of creativity¹). Others presented extraction of creativity concepts related to, e.g., sub-fields of creativity (Agres et al., 2015) and creativity evaluation (van der Velde et al., 2015).

This approach to the study of collections of documents opens the prospect of an automated historiography of the field of computational creativity, an idea which constitutes a satisfyingly recursive application of the research outputs of that area of interest; a similar exercise has been undertaken within the Association for Computational Linguistics (Anderson, McFarland, and Jurafsky, 2012). Here, we illustrate, with real computational examples, the kinds of analysis (e.g., diachronic comparisons of conceptualisation) that would be used for such studies.

The current paper presents a method of semi-automated domain conceptualization, where the domain of interest is Computational Creativity (CC). The paper proposes a conceptualization of CC grounded in papers published in the Proceedings of International Conferences on Computational Creativity (ICCC). Some additional properties of the CC domain are studied, obtained by means of information extraction and dependency analysis techniques. The experimental data is presented, followed by the results of CC domain conceptualization and time dependency analysis.

The Computational Creativity Domain

This section describes the data used in the experiments, together with initial domain understanding achieved through automated terminology extraction.

¹<http://purl.org/creativity/ontology>

ICCC proceedings data The documents were taken from the ICCC proceedings published between 2010 and 2015, inclusive. In total, we considered 247 articles from all six proceedings, containing the following numbers of articles: 2010: 43, '11: 30, '12: 44, '13: 40, '14: 49, '15: 41.

The papers, in PDF, were first converted to plain text. We omitted the references, but added the information about the conference year (however, for time dependency analysis, presented in the last section of this paper, the version of the corpus including references was used).

Automated CC terminology extraction The goal of terminology extraction is to automatically extract relevant terms for a given domain, represented by a given corpus. We used terminology extraction method LUIZ-CF (Pollak et al., 2012), a modified version of the LUIZ term recognition tool (Vintar, 2010). LUIZ-CF is implemented as a workflow in the ClowdFlows environment.

Term extraction consists of two steps: extracting the noun phrase candidates based on morphosyntactic patterns; followed by weighting and ranking of the candidates based on their termhood value, for single word and multi-word terms. The termhood value is computed based on comparison of relative frequencies of lemmas of a term in the domain corpus (here, the ICCC proceedings) compared to a reference corpus: for English, the frequencies of the British National Corpus are used. The extracted terms are ranked by termhood value on a scale between 1 and 0. In addition to default stop words, we eliminated also the names of ICCC PC members, leading to the exclusion of some of the paper authors from the term list. The top ranked candidates are listed in Table 1, followed by a list of top ranked multi-word terms from the same term list given in Table 2. The term extraction method is explained in details in (Pollak et al., 2012). The term extraction workflow is available in ClowdFlows². The extracted terms may be considered as an initial computational creativity vocabulary for building a dictionary of computational creativity, which is planned in future work.

CC domain conceptualization with OntoGen

A tool named OntoGen³ (Fortuna, Grobelnik, and Mladenić, 2007) was used to build a topic ontology for CC domain conceptualization. OntoGen is a semi-automatic and data-driven ontology editor. Semi-automatic means that the system is an interactive tool that aids the user during the topic ontology construction process. Data-driven means that most of the aid provided by the system is based on the underlying text data (document corpus) provided by the user. The system combines text mining techniques with an efficient user interface and was already validated in several applications, including its application to inductive logic programming conceptualization Lavrač et al. (2010).

OntoGen accepts texts in various formats. We chose the named line document format, where each line represents one document, starting with the document ID and the conference edition (2010, 2011, 2012, 2013, 2014 or 2015) as a category. OntoGen performs basic lemmatization and stop word removal (and accepts additional user-defined stop word lists)

²<http://clowdflows.org/workflow/7219/>

³<http://ontogen.ijs.si/>

Table 1: Top 15 terms from the ICCC corpus.

Score	Term
1.000000	[creativity]
1.000000	[computational creativity]
0.862623	[system]
0.247012	[creative system]
0.182190	[process]
0.174810	[model]
0.141525	[image]
0.126607	[concept]
0.102306	[creative process]
0.101973	[word]
0.099952	[evaluation]
0.099844	[conceptual space]
0.081564	[domain]
0.080851	[generation]
0.073521	[story]

Table 2: Top 15 multi-word terms from the ICCC corpus.

Score	Term
1.000000	[computational creativity]
0.247012	[creative system]
0.102306	[creative process]
0.099844	[conceptual space]
0.030894	[computational model]
0.021593	[computational system]
0.018712	[fitness function]
0.018064	[jaguar knight]
0.012638	[genetic algorithm]
0.012078	[human creativity]
0.011300	[poetry generation]
0.011171	[story generation]
0.010011	[neural network]
0.009657	[creative domain]
0.009299	[transformational creativity]

and constructs Bag-of-Words (BoW) vector representations of documents, weighted by the TF-IDF weights (Salton and Buckley, 1988), where TF-IDF stands for Term Frequency-Inverse Document Frequency.

OntoGen is illustrated in Fig. 1. The “unsupervised concept suggestion” functionality is a central part of the system: for a given concept (e.g., the central concept “computational creativity” represented by all the documents of the ICCC domain), a list of sub-concepts is suggested by k -means clustering (Jain, Murty, and Flynn, 1999) and Latent Semantic Indexing (Deerwester et al., 1990) techniques. If the user does not want to affect the conceptualization outcome, only parameter k needs to be chosen to determine the number of concepts, i.e., the number of categories in which the documents will be clustered. “Keywords” (automatically assigned names of clusters) are the words that are the most descriptive for the content of the concepts instances (articles), i.e., words with the highest weights in the document centroid vectors (Fortuna, Mladenić, and Grobelnik, 2006). The main OntoGen window represents the ontology visuali-

sation in which each concept is represented by top three keywords unless manually edited, while the Concept hierarchy window (on the upper left corner) offers a quick overview of all the concepts with their position in the concept hierarchy that can be also directly manipulated.

An alternative view is over the Concepts' documents, where documents of each concept (document cluster) are visualized. Fig. 2 shows documents of the selected concept, i.e. the one represented by keywords "music, chord, improvisation", which could be reasonably be called "Musical creativity". In the similarity graph (at the bottom of the figure), the red dots represent documents belonging to the selected concept, while blue dots the documents not belonging to the concept. The graph inspection functionality can be used for selecting documents to be manually inspected and eventually removed or added to the concept. SVM keywords (see left bottom corner of the figure) are composed from words most distinctive for the selected concept concerning its sibling concepts in the hierarchy (obviously not available for the root concept). SVM keywords are explained more detailed in Fortuna, Grobelnik, and Mladeni (2006).

An important additional functionality of OntoGen is a supervised method for adding concepts. It is based on SVM active learning method Fortuna, Grobelnik, and Mladeni (2006). The user supervision is provided first by a query describing the concept that the user has in mind and followed by a sequence of questions whether a particular instance (document) belongs to the concept and the user can select Yes or No. The questions are chosen from the instances on the border between being relevant to the query or not and are therefore most informative to the system. The system refines the suggested concept after each reply from the user and the user can decide when to stop the process based on how satisfied he is with the suggestions. After the concept is constructed it is added to the ontology as a sub-concept of the selected concept.

Automated CC conceptualization First we performed k -means clustering for $k = 5$. In the topic ontology (Fig. 1), OntoGen uses the first three automatically extracted keywords as concept/topic descriptors.

By inspecting the keywords, we manually named the concepts of the automatically generated topic ontology as follows: "Musical creativity", "Visual creativity", "Linguistic creativity", "Creativity in Games" and "Conceptual creativity" (see Table 3). While some of the categories are quite uniform regarding the keywords (e.g., "music, chord, improvisation, melodies, harmonize, composition, accompaniment, pitch, emotions, beat" for the concept category that we named "Musical creativity"), other categories are more noisy, e.g., "image, story, actions, painting, character, agents, narrative, artists, robot, darci" do not denote a uniform category. We decided to name this category "Visual creativity", but it obviously contains documents from other topics as well, such as narratives generation.

From a set of ten automatically generated keywords⁴ characterizing each of the five document clusters (Table 3), we manually selected three keywords believed best to describe the cluster of papers belonging to each concept (in italics).

⁴Words with the highest weights in the document centroid vectors (Fortuna, Mladenič, and Grobelnik, 2006)

These keywords were added to the concept labels of Fig. 3.

In the next section, aiming at a more elaborated version of the CC ontology (see Fig. 4), we use the concept moving facility of OntoGen, by which we moved e.g., the concept "Narrative" from "Visual" to "Lexical", together with other techniques for manipulating the initial ontology.

Semi-automated CC domain conceptualization This section describes improved CC conceptualization, created by manipulating the initial ontology using different OntoGen functionalities. The main concepts were further divided and when forming meaningful concepts, the categories were added as sub-concepts (see e.g. the sub-concepts of Linguistic creativity in Fig. 4). As already mentioned, some (sub-)concepts were moved, e.g., the "narrative" category was moved from Visual to Linguistic creativity. Some concepts were added by query and active learning. On the first level, this is the case for the category Evaluation, which was a recurrent topic in other categories and we used query and active learning to form an independent category. We also used it, e.g., for creating the category "Recipes" as a sub-concept of lexical creativity. We also used the OntoGen function to (de)select the documents being categorized to one concept category.

Fig. 2 shows the documents belonging to a category. It is very interesting to inspect some outliers (documents similar to documents in the category not being classified to this category). In the concept document graph, we identified some of the outliers, represented by blue dots in the similarity line of red dots. An example is article 2014_44, entitled "Arts, News, and Poetry The Art of Framing", by Gross, Toivanen, Laane and Toivonen. This paper was not classified in the Linguistic creativity category, but was identified as similar to the documents of that category. The article indeed refers to linguistic creativity (poetry) but also to generated pictures as well as pictures painted by an artist. We manually added this document also to the category of linguistic creativity.

The result of this experiment is shown in Fig. 4. On the first level we distinguish between Musical, Visual, Linguistic creativity, Games and creativity, Conceptual creativity as well as newly created category of Evaluation. On lower levels, we added e.g. Narratives, Poetry, Recipes and Lexical creativity for Linguistic creativity, where the latter comprises e.g., humour, neologisms, etc.

Each concept is represented by descriptive Keywords (see keywords for six first level concepts in Table 4) from which we selected three keywords (in italics) to represent the concept in the visual ontology (Fig. 4). The ontology can be considered as a draft to be collaboratively improved by the CC community. Further, since the concepts are grounded in the documents, the top ranked documents might be considered as interesting reading for newcomers to CC. The bibliography can be created for concepts of different levels (as an example see three articles per selected topic):

Narratives:

- A System for Evaluating Novelty in Computer Generated Narratives (Pérez y Pérez et al., 2011)
- Kill the Dragon and Rescue the Princess: Designing a Plan-based Multiagent Story Generator (Laclaustra et al., 2014)

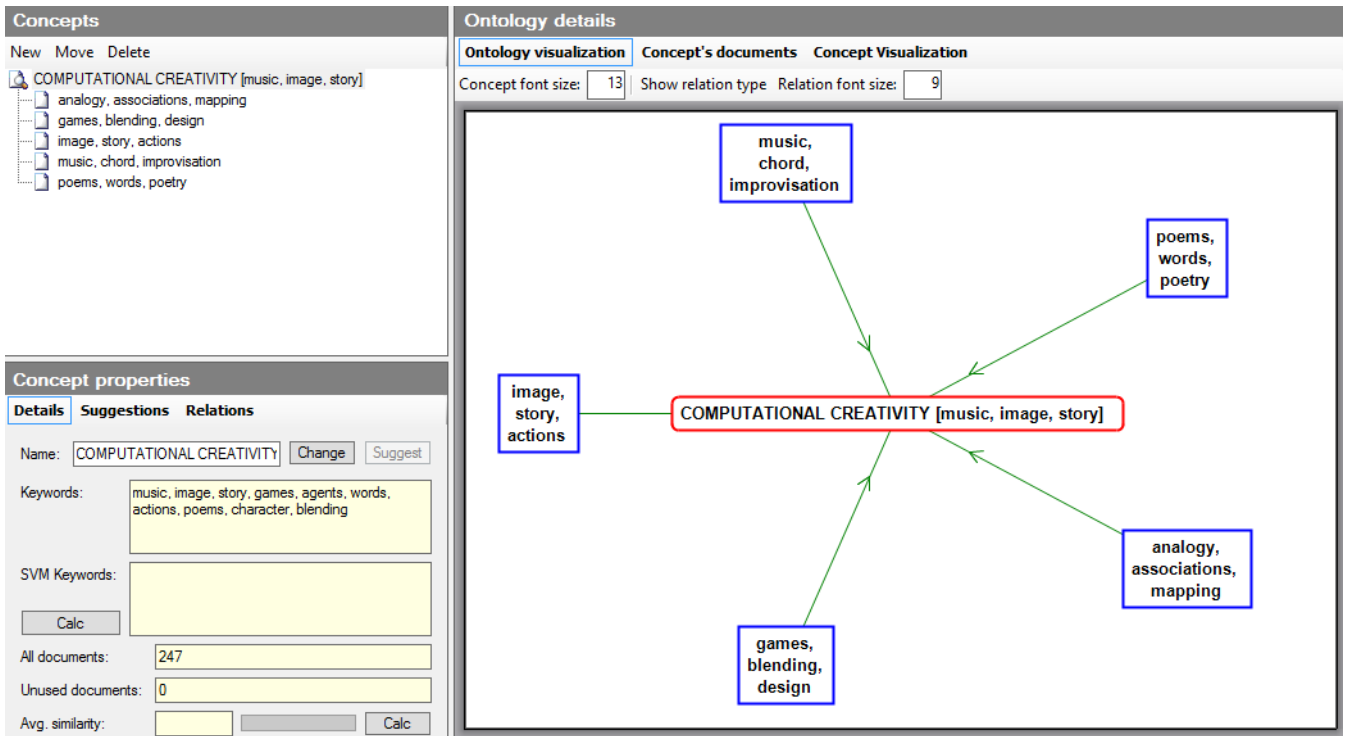


Figure 1: Automatically generated conceptualization of the CC domain.

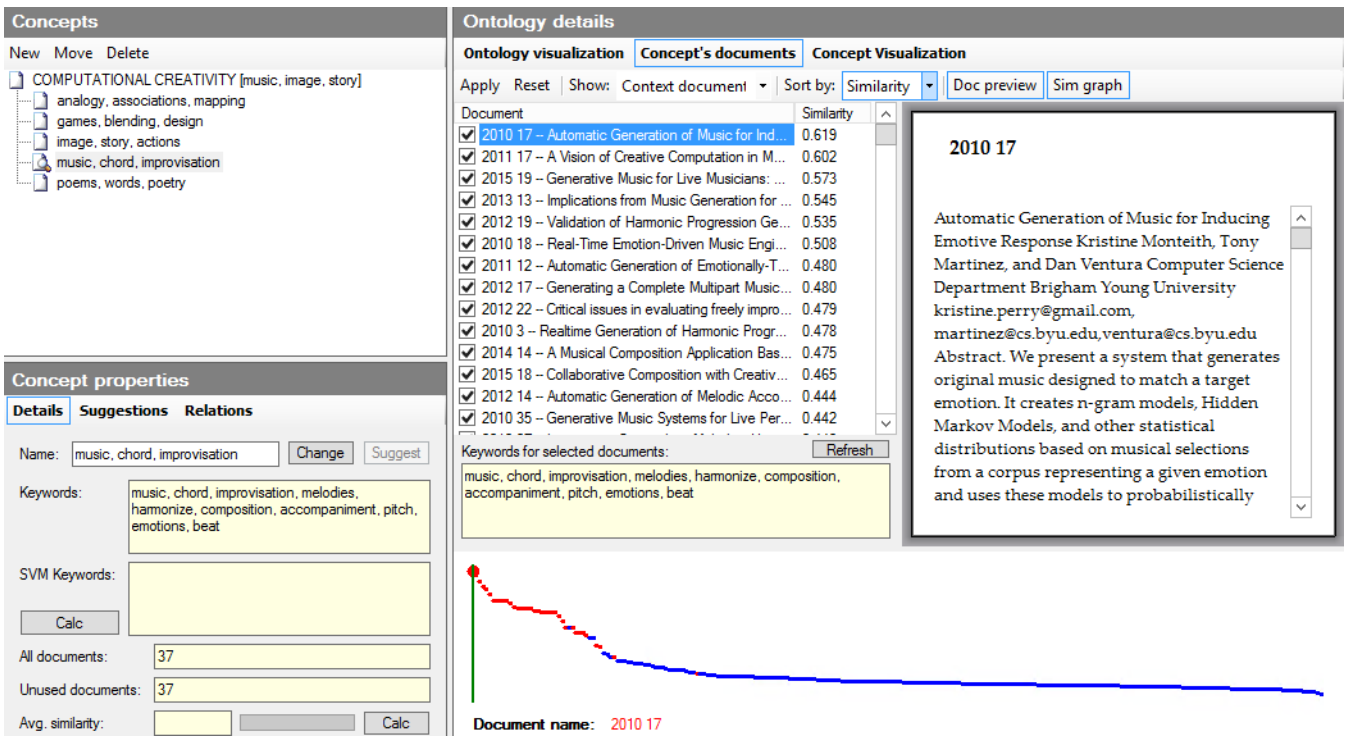


Figure 2: Document view of the automatically generated conceptualization of the CC domain.

Table 3: Automatically generated concepts (concept names were manually determined) and keywords.

Concept	Automatically extracted keywords
Music	<i>music</i> , chord, <i>improvisation</i> , melodies, harmonize, <i>composition</i> , accompaniment, pitch, emotions, beat
Visual	<i>image</i> , story, actions, <i>painting</i> , character, agents, narrative, <i>artists</i> , robot, darci
Linguistic	<i>poems</i> , words, poetry, artefacts, <i>story</i> , evaluating, creativity_system, predict, <i>text</i> , creativity
Games	<i>games</i> , blending, <i>design</i> , analogy, <i>player</i> , conceptual, games_design, angelina, ontology, agents
Conceptual	<i>analogy</i> , associations, <i>mapping</i> , graphs, objective, problem, fractal, domain, <i>representation</i> , relationship
Comp. creativity	<i>music</i> , <i>image</i> , <i>story</i> , games, agents, words, actions, poems, character, blending

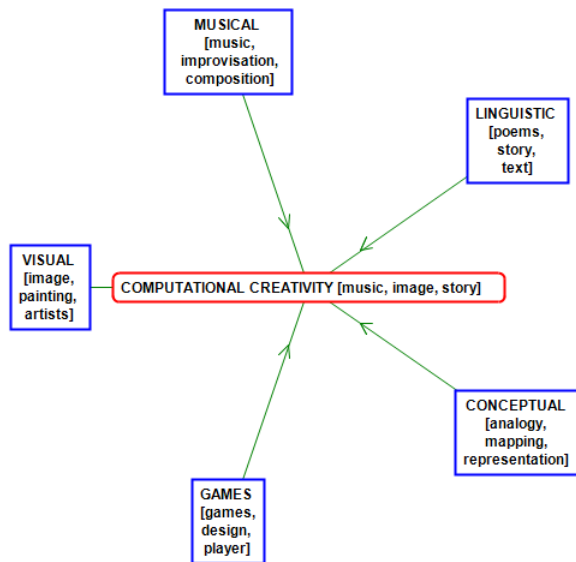


Figure 3: CC conceptualization through automated clustering, concept naming and manual keywords selection.

- Creativity in Story Generation From the Ground Up: Non-deterministic Simulation driven by Narrative (León and Gervás, 2014)

Games:

- Computational Game Creativity (Liapis, Yannakakis, and Togelius, 2014)
- Ludus Ex Machina: Building A 3D Game Designer That Competes Alongside Humans Michael (Cook and Colton, 2014)
- Knowledge-Level Creativity in Game Design (Smith and Mateas, 2011)

Music:

- Automatic Generation of Music for Inducing Emotive Response (Monteith, Martinez, and Ventura, 2010)
- A Vision of Creative Computation in Music Performance Roger (Dannenberg, 2011)
- Generative Music for Live Musicians: An Unnatural Selection (Eigenfeldt, 2015)

Analysis of CC domain development in time

In this section, we investigate temporal aspect of ICCC proceedings. First, we use OntoGen by which we first categorize the articles into editions and then observe the words

distinguishing these categories. Second, the frequency analysis of terms was used to identify terms that are characteristic only for first or last editions. Last but not least, we use the copulas for investing the time dependency between the content of different conference editions.

Distinctive keywords by years using OntoGen OntoGen (Fortuna, Grobelnik, and Mladenić, 2007) is capable of supervised categorization of documents in predefined categories. For this experiment, we used conference years as categories and extracted characteristic keywords for each year. The first set of *descriptive* keywords (KeyW in Table 5) is extracted using document centroid vectors, while the second set of *distinctive* keywords (SVM in Table 5) is extracted from the SVM classification model dividing documents in the topic from the neighbouring ones (Fortuna, Mladenić, and Grobelnik, 2006). Table 5 shows both sets of words for each year.

Unsurprisingly, descriptive keywords overlap different years: words recurring most across years are “creativity, design, modelling, system”. More interesting are the distinctive keywords: ICCC-2015 might be characterized by “bots”, ICCC-2014 and ICCC-2011 by “games”, 2014 by “metaphors” and “stereotypes”, ICCC-2012 by “melodies” and “associations”, and ICCC-2010 by “analogies”.

Categorization by year can also be used for specific topics: in the ontology in Fig. 4, we can split a selected topic into year categories and observe the distinctive (SVM) keywords. The papers representing the concept of Musical creativity in 2015 contain words such as “musebot, pc, unnatural...”, but in 2010, “chord, improvisation, jazz...”.

Terminology distribution by years The frequency distributions of the top 1,000 terms obtained by the terminology extraction process described earlier in the paper are an indicator to detect terms that recently occurred or were present only in the early editions. Examples of terms that appear in 2014 and 2015 and not before are: “game jam, co-creative system, concept invention, generative software, curation coefficient, procedural generation, player goal, network analysis, simulation model” (30 terms in total). In contrast, the terms that were used in 2010 and 2011 are: “fractal representation, fractal feature, basel problem, sensory system, fractal algorithm”.

Copula-based analysis of dependencies between ICCC proceedings

This section describes measuring the detected dependencies between different years of ICCC proceedings. As in the previous section, we counted the frequencies of automatically extracted terms in the ICCC pro-

Table 4: Categories and keywords of the first layer of the semi-automatically constructed CC ontology.

Category	Automatically extracted keywords
Musical	<i>music, chord, improvisation, melodies, harmonize, composition, accompaniment, pitch, emotions, beat</i>
Visual	<i>image, painting, darci, artifacts, collage, adjectives, associations, rendered, colored, artists</i>
Linguistic	<i>story, poems, actions, character, words, agents, narrative, artefacts, poetry, evaluating</i>
Games	<i>games, design, player, games_design, angelina, agents, code, jam, filter, gameplay</i>
Conceptual	<i>analogy, blending, mapping, conceptual, objective, associations, team, graphs, concepts, domain</i>
Evaluation	<i>music, poems, improvisation, evaluating, interactive, poetry, creativity system, musician, participants, behavioural</i>
Comp. creativity	<i>music, image, story, games, agents, words, actions, poems, character, blending</i>

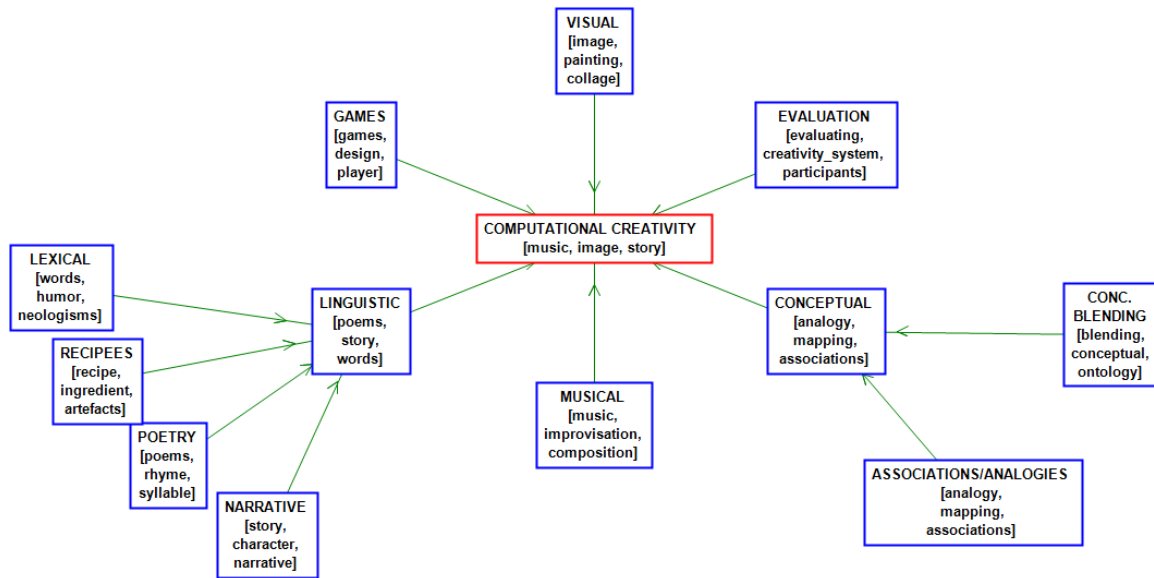


Figure 4: Semi-automatically generated conceptualization of the CC domain, with concept naming and subconcept creation.

Table 5: Keywords and distinctive (SVM) words by year.

Year	Category	Words
2015	KeyW:	creativity, generated, image, work, human, blending, design, based, music, words
	SVM:	<i>blending, humor, bots, choice, vectors, musician, jam, conceptual, colour, participants</i>
2014	KeyW:	creativity, computer, process, modelling, evaluating, words, agents, story, domain, based
	SVM:	<i>games, agents, story, artists, adjectives, ontology, domain, motifs, poems, actions</i>
2013	KeyW:	modelling, process, figure, image, design, performs, based, levels, interactive, concepts
	SVM:	<i>robot, metaphor, motion, surprising, evolved, image, composition, mechanism, stereotypes, fictional</i>
2012	KeyW:	creativity, system, computer, music, evaluating, user, human, figure, work, set
	SVM:	<i>melodies, associations, accompaniment, character, template, shape, player, monotone, text, cluster</i>
2011	KeyW:	creativity, system, story, design, modelling, results, games, music, set, problem
	SVM:	<i>story, games, movements, playing, graphs, games_design, darci, actions, identical, strategies</i>
2010	KeyW:	generated, system, user, set, idea, design, emotions, analogy, developments, based
	SVM:	<i>analogy, chord, emotions, improvisation, genes, filter, lives, team, jazz, songs</i>

ceedings of each year from 2010 until 2015. This information was used as input for the copula-based dependency analysis between six ICC proceedings, described below.

The scatter plot of the terms that occur in 2010 in comparison with terms that occur in years 2011 to 2015 are given

in Fig. 5. The number of occurrences of the specific term in 2010 are given on the x axis, while the number of occurrences of the same term in other years is represented on the y axis. The scatter plot graphically shows a positive dependence between different years as data points are clus-

Table 6: Results from bi-variate copulas for all terms.

No.	Copula type	Coupling of domains	θ
1	Best Clayton	2010-2012	3.1941
2	Best Frank	2010-2012	9.1507
3	Worst Clayton	2010-2015	2.4010
4	Worst Frank	2010-2015	7.3955

Table 7: Two Archimedean copulas.

Copula type	$C_\theta(u, v)$
Clayton	$[\max(u^{-\theta} + v^{-\theta} - 1, 0)]^{-1/\theta}$
Frank	$-\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right)$

tered in a band running from lower left to upper right. In the next step, we quantify the dependency between the different years. The most commonly used measure for dependency is correlation. The correlation between two variables (e.g., the ICCC proceedings of two distinctive years) can be measured by means of the Pearsons correlation coefficient. It is a dimensionless quantitative measure of statistical relationships between two (or more) variables. It measures the degree of linear correlation; however the two variables may have different functional dependency. For this reason we apply the copula functions as a tool for studying and measuring the dependences of random variables (Sklar, 1959).

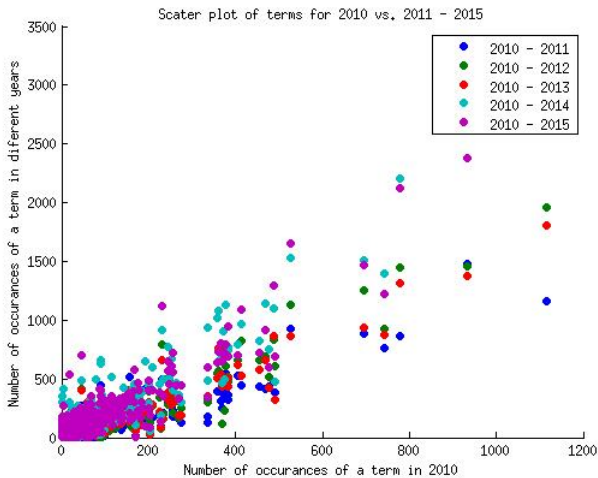


Figure 5: The scatter plot of the terms that occur in 2010 in comparison with terms that occur in years 2011 to 2015.

Copulas are functions that formulate the multivariate distribution in such a way that various general types of dependences including the non-linear one may be captured. We focus on two families of bi-variate Archimedean copulas: Clayton and Frank. Their usage is mainly motivated by their convenient properties, such as symmetry and associativity. Their mathematical forms are presented in Table 7. The parameter θ is estimated from the data. Higher values of θ

mean higher dependence between the two variables.

We explored the dependencies between pairs of proceedings. For this purpose we built Clayton and Frank bi-variate copulas. The results of best copulas (the most dependent pairs of proceedings) and the worst copulas (the least dependent pairs) are provided in Table 6. It can be observed that the ICCC-2010 and ICCC-2015 proceedings are contents-wise the least connected, while the most dependent proceedings are those from 2010 and 2012, where both conferences were organized in Europe.

Conclusions

In the paper, we have presented the conceptualization of the computational creativity domain by semi-automated topic ontology construction based on the corpus of ICCC proceedings. We analysed automatically extracted keywords and subconcepts for CC domains (Visual, Musical, Linguistic, Conceptual creativity, Creativity and games and Evaluation). In addition we analysed characteristics of different editions of CC conferences and used copulas to measure the dependency between proceedings of different editions. As result of this research, we make available for further research a) the ICCC proceedings corpus in .txt format with and without reference sections, b) automatically extracted ICCC terminology that can be used for future efforts in creating a CC glossary, c) fully automated topic ontology with automatic keywords extraction, as well the semi-automated CC ontology, which is the result of manual manipulation of the automatic ontology. Ontologies are available in .png and .rdf formats. All the resources are available publicly⁵

In future, as these techniques develop to full automation, and the amount of data increases with successive conferences, it will be possible to construct a timeline of the conceptual development of the field of computational creativity, using the objective analysis of the literature. It will be possible to trace the rise and fall of trends, and their success or failure, and to identify the development of core CC science, as proposed by (Lakatos, 1970). This activity will be unique in science, and will support an unprecedented level of unbiased philosophical reflection on the field of computational creativity.

Acknowledgements

We acknowledge the support of the Slovenian Research Agency and European projects Prosecco (grant nb. 600653) and ConCreTe, which acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET grant number 611733.

References

- Agres, K.; McGregor, S.; Purver, M.; and Wiggins, G. 2015. Conceptualizing creativity: From distributional semantics to conceptual spaces. In *In the Sixth International Conference on Computational Creativity, ICCC 2015*.

⁵http://kt.ijs.si/senja_pollak/CC_resources/.

- Anderson, A.; McFarland, D.; and Jurafsky, D. 2012. Towards a computational history of the acl: 1980-2008. In *Proceedings of the ACL-2012 Special Workshop on Rediscovering 50 Years of Discoveries*, ACL '12, 13–21. Stroudsburg, PA: ACL.
- Boden, M. A. 2004. *The Creative Mind: Myths and Mechanisms*. Routledge.
- Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In de Raedt, L.; Bessiere, C.; Dubois, D.; and Doherty, P., eds., *Proc. ECAI Frontiers*.
- Cook, M., and Colton, S. 2014. Ludus ex machina: Building a 3d game designer that competes alongside humans. In *Proceedings of the Fifth International Conference on Computational Creativity, ICCCC2014*, 54–62. ACC.
- Dannenberg, R. B. 2011. A vision of creative computation in music performance. In *Proceedings of the Second International Conference on Computational Creativity*, 84–89.
- Deerwester, S.; Dumais, S. T.; Furnas, G. W.; Landauer, T. K.; and Harshman, R. 1990. Indexing by latent semantic analysis. *Journal of the American Society for Information Science* 41(6):391–407.
- Eigenfeldt, A. 2015. Generative music for live musicians: An unnatural selection. In *Proceedings of the Sixth International Conference on Computational Creativity (ICCC 2015)*, 142–149. Park City, Utah: Brigham Young Uni.
- Fortuna, B.; Grobelnik, M.; and Mladenčić, D. 2007. Ontogen: Semi-automatic ontology editor. In *Human Computer Interface (Part II) (HCI 2007)*, LNCS 4558, volume 4558, 309–318.
- Fortuna, B.; Grobelnik, M.; and Mladeni, D. 2006. Semi-automatic data-driven ontology construction system. In *PASCAL EPprints (2006)* <http://eprints.pascal-network.org/perl/oai2>. Working Group Summary 15.
- Fortuna, B.; Mladenčić, D.; and Grobelnik, M. 2006. *Semantics, Web and Mining: Joint International Workshops, EWMF 2005 and KDO 2005, Porto, Portugal, October 3-7, 2005, Revised Selected Papers*. Berlin, Heidelberg: Springer Berlin Heidelberg. chapter Semi-automatic Construction of Topic Ontologies, 121–131.
- Gruber, T. R. 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition* 5(2):199–220.
- Jain, A. K.; Murty, M. N.; and Flynn, P. J. 1999. Data clustering: a review. *ACM Computing Surveys* 31(3):264–323.
- Jordanous, A., and Keller, B. 2012. Weaving creativity into the semantic web: a language-processing approach. *Proceeding of the Third International Conference on Computational Creativity* 216–220.
- Laclaustra, I. M.; Ledesma, J. L.; Mendez, G.; and Gervás, P. 2014. Kill the dragon and rescue the princess: Designing a plan-based multi-agent story generator. In *Proceedings of the Fifth International Conference on Computational Creativity*. Ljubljana, Slovenia: Jožef Stefan Institute, Ljubljana, Slovenia.
- Lakatos, I. 1970. Falsification and the methodology of scientific research programmes. In Lakatos, I., and Musgrave, A., eds., *Criticism and the Growth of Knowledge*. Cambridge, UK: Cambridge University Press. 91–196.
- Lavrač, N.; Grčar, M.; Fortuna, B.; and Velardi, P. 2010. *Computational Social Network Analysis: Trends, Tools and Research Advances*. London: Springer London. chapter Exploratory Analysis of the Social Network of Researchers in Inductive Logic Programming, 135–154.
- León, C., and Gervás, P. 2014. Creativity in story generation from the ground up: Non-deterministic simulation driven by narrative. In *5th International Conference on Computational Creativity, ICCCC 2014*.
- Liapis, A.; Yannakakis, G. N.; and Togelius, J. 2014. Computational game creativity. In *Proceedings of the Fifth International Conference on Computational Creativity*. Ljubljana, Slovenia: Josef Stefan Institute.
- Monteith, K.; Martinez, T.; and Ventura, D. 2010. Automatic generation of music for inducing emotive response. In *Proceedings of the International Conference on Computational Creativity*, 140–149. Lisbon, Portugal: Department of Informatics Engineering, University of Coimbra.
- Pérez y Pérez, R.; Ortiz, O.; Luna, W.; Negrete, S.; Castellanos, V.; Alosa, E. P.; and Ávila, R. 2011. A System for Evaluating Novelty in Computer Generated Narratives. In Ventura, D.; Gervás, P.; Harrell, D. F.; Maher, M. L.; Pease, A.; and Wiggins, G., eds., *Proceedings of the Second International Conference on Computational Creativity*, 63–68.
- Pollak, S.; Vavpetič, A.; Kranjc, J.; Lavrač, N.; and Špela Vintar. 2012. NLP workflow for on-line definition extraction from English and Slovene text corpora. In Jancsary, J., ed., *Proceedings of KONVENS 2012*, 53–60. ÖGAI. Main track: oral presentations.
- Salton, G., and Buckley, C. 1988. Term-weighting approaches in automatic text retrieval. *Information Processing & Management* 24(5):513–523.
- Sklar, A. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* 8:229231.
- Smith, A. M., and Mateas, M. 2011. Knowledge-level creativity in game design. In *In Proc. of the 2nd International Conference in Computational Creativity, (ICCC 2011)*.
- Smith, B. 2003. Chapter 11: Ontology. In Floridi, L., ed., *Blackwell Guide to the Philosophy of Computing and Information*, volume 7250. Blackwell. 155–166.
- van der Velde, F.; Wolf, R. A.; Schmettow, M.; and Nazareth, D. S. 2015. A semantic map for evaluating creativity. In *Sixth Interantional Conference on Computational Creativity (ICCC 2015)*. Park City, Utah, USA: Brigham Young University.
- Velardi, P.; Faralli, S.; and Navigli, R. 2013. Ontolearn reloaded: A graph-based algorithm for taxonomy induction. *Computational Linguistics* 39(3):665–707.
- Vintar, Š. 2010. Bilingual term recognition revisited the bag-of-equivalents term alignment approach and its evaluation. *Terminology* 16:141–159.
- Wiggins, G. A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Journal of Knowledge Based Systems* 19(7):449–458.