Cognitive AI Systems Contribute to Improving Creativity Modeling and Measuring Tools

Faheem Hassan Zunjani(✉) and Ana-Maria Olteţeanu

Freie Universität, Berlin, Germany
{zunjani.faheem,ana-maria.olteteanu}@fu-berlin.de

Abstract. Cognitive science and cognitive psychology have long used creativity tests to measure and investigate the relationships between creativity, creative problem solving and other cognitive abilities. Implementing cognitive systems that can model and/or solve creativity tests can shed light on the cognitive process, and presents the possibility of building much more precise creativity measuring tools. This paper describes four cognitive AI systems related to the Remote Associates Test (RAT) and their contributions to creativity science. comRAT-C is a system that solves the RAT, correlating with human performance. comRAT-G reverse engineers this process to generate RAT queries with a high degree of parameter control. fRAT generates functional RAT queries, resurrecting a theoretical concept proposed by researchers many decades ago. The visual RAT takes advantage of the formal conceptualization necessary for computational implementation, to expand the RAT to the visual domain. All the cognitive systems and generated RAT queries have been successfully validated with human participants and have contributed in improving creativity modeling and measuring tools.

Keywords: Remote Associates Test · Human creativity · Visual associates · Computational creativity · Cognitive systems

1 Introduction

Creativity and creative problem solving, though not uniquely human traits, have contributed to technological, scientific and cultural advances that lay at the foundation of human civilization. They are still cognitive tools which make humans adaptable and able to progress in conditions in which not all knowledge or resources are available.

Various streams of research have been connecting creativity research and the computational sciences. One of these is computational creativity - which

Supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) through the Creative Cognitive Systems project OL 518/1-1 (CreaCogs).

J. M. Ferrández Vicente et al. (Eds.): IWINAC 2019, LNCS 11486, pp. 97–107, 2019.
https://doi.org/10.1007/978-3-030-19591-5_11

focuses on the development of computational creativity systems capable of various creativity feats - like poetry and painting [3] and on coming up with ways of evaluating computational creativity systems.

However, this research does not bring us closer to understanding human (cognitive) creativity. Cognitive creativity is generally explored in cognitive science and cognitive psychology. Computational implementations of hypotheses from the cognitive science community are very valuable, because they offer the chance to test cognitive theories via computational systems [7]. Some such systems have been implemented in the literature to study metaphor, insight, object replacement, etc. Furthermore, cognitive AI systems able to perform or model creativity tasks studied in the cognitive science domain may be used to later (i) build cognitive AI systems that act as a more natural interface for humans and (ii) improve creativity in human participants, by intuiting and supporting the weaknesses of their human counterparts.

This paper focuses on describing a research arc consisting of four cognitive AI systems, focused on one initial type of task - creative association. Besides providing a shortened eagle-eye overview of these systems, this paper's goal is to showcase how this type of research on cognitive AI systems contributes towards improving the tools used for modeling and measuring creativity.

The rest of the paper has been organized as follows. The first section describes the Remote Associates Test as a measure for the associativity factor in creativity, and briefly summarizes the principles of the CreaCogs framework. Section 2 describes the initial formalization of the RAT for adaptation to the CreaCogs framework, its implementation as the comRAT-C cognitive system and relation to human performance.

2 The Remote Associates Test and CreaCogs

Building AI systems which solve tasks that have an empirical creativity measure attached has not been explored systematically. With creativity being a multifaceted cognitive skill, many empirical creativity measures exist.

Among many other tests, one of the most popular and widely used tests to measure creativity in humans is the Remote Associates Test (RAT) proposed by Mednick and Mednick in 1971 [5]. Inspired from Mednick's belief that the creative process has an associative bias [4], the RAT was designed to measure the creativity of a participant based on their ability to draw remote associations. The RAT comprises a number of test queries where each query consists of 3 words and the participant is supposed to answer the word that is connected to all 3 of the query words. For example, for the "SWISS-COTTAGE-CAKE" query the answer would be "CHEESE" as "CHEESE" appears with each of the three query words forming a compound word in the English language.

CreaCogs [8,17] is a theoretical framework aiming to implement a set of creativity related abilities with a minimal set of processes and the same type of knowledge organization and cognitive architecture. This is in accord with main principles of cognitive architectures [7], in which one architecture should be flexible enough to solve a multiplicity of tasks within a certain domain. CreaCogs proposed processes [8] to solve associativity related tasks (treated here), creative use and inferences related to objects [18], and more complex insight problems using objects from the practical domain [9,16], using a multilayered type of knowledge organization. In the following, we focus on the impact of this knowledge organization on constructing cognitive AI systems related to the associative ability.

3 Computational Solver for the Compound RAT - comRAT-C

A computational solver of the RAT named comRAT-C was proposed [10] to test the CreaCogs abilities related to associative process. For this, the Remote Associates Test was formalized as a succession of queries, in which three words w_a, w_a and w_c are shown to the system, and an answer word is expected.

In order to test its knowledge organization hypothesis, comRAT-C sequentially extracted 2-grams from the Corpus of Contemporary American English (COCA: http://corpus.byu.edu/coca/) and used them to organize its knowledge base (RAT-KB) with three primary atomic structures: EXPRESSION, CONCEPT and LINK. Whenever comRAT-C arrives at a 2-gram EXPRESSION (two words representing two CONCEPTS), it checks for the presence of each of the two CONCEPTS in its existing KB. If both CONCEPTS are present, only a LINK between them is added. The LINK contains a numerical tag of the frequency of the 2-gram in the corpus attached to it, and represents a form of encoding the trace of cognitive association strength. If one or both of the words are unknown, they are added as CONCEPTS and a LINK between them is also added. When the comRAT-C is done constructing its RAT-KB, each CONCEPT then contains LINKS to all other CONCEPTS that it has been in an EXPRESSION with. This is then used to solve the Remote Associates Test.

comRAT-C performs a convergence process based on its associations. This process is meant as a hypothesis for a cognitive process humans may use when solving the task, and uses the strength of associations. It is known for example that some human participants, when given a query, reach the answer by what they perceive as insight: the answer seems to just "pop in" to their mind. The comRAT-C convergence process aims to replicate a form of associative search in the memory, in which the answer is revealed through the strength of the associations.

The probability of reaching an answer is calculated by comRAT-C based on association strength. The probability that $w_i^{'}$ is the answer is:

$$P[w_i^{'}] = 1/3 * P[w_i^{'}|w_a] * P[w_i^{'}|w_b] * P[w_i^{'}|w_c] \tag{1}$$

Figure 1 shows a visualization of the convergence process when comRAT-C searches for solutions to a query in RAT-KB. The words surrounded by green circles represent the query words. The blue ones are the words associated with LINKS to the query words. The yellow ones represent a 2-item convergence and the red one represents a 3-item convergence which is chosen as a plausible solution by comRAT-C.

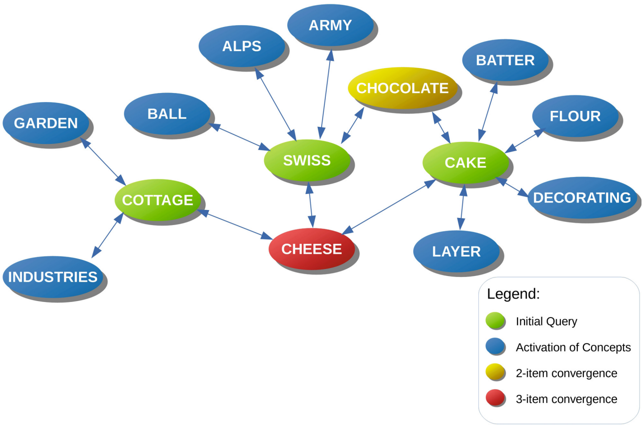

Fig. 1. Visual representation of the comRAT-C activation and convergence. (Color figure online)

The performance of comRAT-C was measured over the Bowden and Jung-Beeman normative dataset [2] of compound RAT queries. Out of 144 queries in the dataset, in 64 of them comRAT-C provided the exact answers to the queries. In more than 20 cases, comRAT-C provided an answer that could be considered as a plausible answer but it was not considered in the Bowden & Jung-Beeman dataset. In 97.92% of the cases in which comRAT-C had all the query words present in its KB, it gave the correct answer. It was interesting to note that in 30.36% of the cases comRAT-C gave the correct answer when it had only 2 query items in its KB demonstrating the robustness of the system.

In terms of its promise as a tool for future modeling, comRAT-C's process probability to solve a RAT query significantly correlated with the Accuracy of human participants, and correlated negatively with their reponse times (the higher the probability, the less time human participants take to solve a query).

4 Computational Generator of Compound RAT Queries - comRAT-G

The RAT was found be the second most used creativity test in a meta analysis study [1]. There are few repositories of normative data for the RAT, like Bowden and Jung-Beeman dataset of 144 RAT queries which was used in the validation of comRAT-C. These repositories still do not provide the ability to vary the query items on the basis of the frequency of occurrence of query words or answer words.

To deal with this bottleneck, after the successful validation of the comRAT-C, it was found that the knowledge organization used by it could not only be used to solve the RAT queries but also to generate them by reversing the convergence process. Hence, comRAT-G was created as a variant of comRAT-C that generates RAT queries using the same knowledge base (KB).

The intuition of the query creation process is the following: each query word in this type of knowledge organization can, in turn, be seen as a potential answer word. For example, in Fig. 1 – SWISS can be seen as an answer word for query ALPS, ARMY, CHOCOLATE.

The comRAT-G first iterated over the KB and provided nouns which were the potential answer words (w_{ans}) along with their linked query words (w_qs). Only the w_{ans}s that had at least three w_qs were selected by comRAT-G. This process yielded around 81500 unique (w_{ans}, w_q) combinations out of 9601 unique w_{ans}s. The probability of w_{ans} as answer for a query word w_q was calculated as:

$$P[w_{ans}|w_q] = \frac{fr(w_q w_{ans})}{\sum_{k=1}^{n}(w_q, w_k)} \tag{2}$$

To make constructing all the combinations computationally feasible, comRAT-G used Alan Tucker's combinatorics formula [20] and capped n at 100. With this, 17 million possible RAT queries were obtained. The probability of answering the generated queries was calculated as the conditional probability of w_{ans} being triggered by each of the query words (w_a, w_b, w_c) was calculated as in comRAT-C.

Table 1 shows a few examples from the generated queries. Frequency and frequency based probability of finding an answer was computed for all the queries.

Table 1. Example queries generated by comRAT-G.

w_a	w_b	w_c	w_{ans}	$P(w_{ans})$
box	panes	shades	window	0.4016
penalty	suit	toll	death	0.5243
paddle	roulette	steering	wheel	0.5582
checking	escrow	deficits	account	0.5626

To validate the generated queries, a study was conducted in which human performance in comRAT-G queries was compared to that in a normative dataset [2]. 113 native English speakers (72 females and 41 males) recruited from the University of Pittsburg and Figure-Eight (F8, https://figure-eight.com) participated in the study. The test consisted of 50 queries randomly selected from the Bowden & Jung-Beeman dataset and 50 queries randomly picked from the queries generated by comRAT-G.

An average significant correlation of $r = 0.54$ ($p < 0.0001$) was observed between the accuracy in comRAT-G queries and Bowden & Jung-Beeman queries. A highly significant large correlation of $r = 0.75$ ($p < 0.0001$) was observed between the response times for the two sets of queries. Cronbach's alpha on accuracy was 0.932 for Bowden & Jung-Beeman queries, 0.851 for comRAT-G queries and 0.936 for both the sets combined. The average and high correlations obtained between the performance of human participants on both the datasets validated the comRAT-G queries. The high Cronbach's alphas show that both the datasets are highly reliable and consistent with each other.

This showed the first implementation in which a cognitive AI system was used for the generation of new queries – displaying applications to improving creativity tests and other psychometric tools related to creativity science. Later on, this system was shown to provide a stronger control over factors and allow new empirical designs [13].

5 Functional RAT and Its Computational Generator - fRAT and comRAT-G_F

Worthen and Clark [21] argued that Mednick's original queries were not uniform, and could be broadly categorized into at least two types: structural and functional. "TOOTH-POTATO-HEART" (answer: "SWEET") is an example of a structural query since "SWEET" occurs together with each of the three query words syntactically in the English language. Whereas, in "DAISY-TULIP-VASE" (answer: "FLOWER"), "FLOWER" shares a language-independent functional relationship with the query words.

As the functional queries examples proposed by Worthen and Clark were lost in transport between two libraries, and this concept was thus never fully explored in the scientific community, [12] set as a goal the resurrection of the concept by computationally constructing a set of functional queries. This set of queries was to then allow comparisons between compound and functional queries for cognitive scientists.

comRAT-G_F [12] is thus a cognitive AI system that aims to computationally generate functional RAT (fRAT) queries. This approach modifies the comRAT-G (Sect. 4) to generate functional queries instead of compound (structural) ones. As Worthen and Clarke considered Palermo-Jenkins word association norms [19] a good source for validating Mednick's initial queries, comRAT-G_F uses a more modern dataset of free association, rhyme and word fragment norms [6] to extract the data for building the knowledge base. This dataset recorded words that were produced by human participants when presented with cue words. For example, when presented with "ABUNDANCE" as a cue word, participants came up with words like "FAMINE", "FOOD", "FULL".

For the creation of functional queries, items were extracted from the dataset and the knowledge base was organized in a structure similar to that of comRAT-G. Words with more than three associates were considered as possible answers and were used as the basis for generating potential functional queries in further stages. The number of subjects producing a target word from the University of South Florida association norms was used as a stand-in for frequency metric. Similarly, the number of participants producing a target word when given a cue over the total participants given that cue was used as a stand-in for probability.

Out of the 13,534,865 fRAT queries generated by comRAT-G_F, Table 2 shows a few examples. To evaluate the created queries, and establish human performance baselines on them, two studies were conducted.

Table 2. Example fRAT queries generated by comRAT-G_F.

w_a	w_b	w_c	w_p	Probability
exhausted	sleepy	weary	tired	0.7202
frame	photo	portrait	picture	0.6897
bassinet	crib	infant	baby	0.6916
daisy	tulip	vase	flower	0.6914
bulb	dark	dim	light	0.5530

In the first study, Figure-Eight users who had previously solved compound queries in a previous study were invited. In the test, 75 fRAT queries with probabilities of obtaining the answer word distributed over the range of 0 to 0.5 were presented to every participant in randomized order. The number of correct answers (accuracy) of every participant was computed and correlated to their performance in the comRAT-G study with compound queries. The accuracy scores on fRAT queries show a large positive significant correlation of $r = 0.55$ ($p < 0.005$) with accuracy on compound queries. For the response times, a medium sized correlation of $r = 0.41$ ($p < 0.05$) was observed.

The second study aimed to further investigate the relationships observed in the first study with a larger sample of 61 participants (44 females and 17 males). Participants were presented with 96 queries consisting of 48 fRAT queries and 48 compound RAT queries (24 comRAT-G queries and 24 Bowden & Jung-Beeman queries) in randomized order. A significant correlation of $r = 0.44$ ($p < 0.001$) was observed between the accuracy in fRAT and compound RAT queries. A highly significant and strong correlation of $r = 0.88$ ($p < 0.001$) was found between the response times for fRAT and compound RAT queries. A measure of reliability, Cronbach's alpha was found to be 0.79 for the accuracy on fRAT queries and 0.87 on compound RAT queries. Cronbach's alpha on response items of the correct answers also showed a high reliability, 0.90 for fRAT queries, compared to 0.96 for compound RAT queries.

The comRAT-G_F cognitive system was thus successful at creating functional queries of high reliability, that correlate in both accuracy and reponse times with compound queries. This offers a new point of measurement and psychometric tool for creativity science.

6 Visual Remote Associates Test - vRAT

The Remote Associates Test is a language based test, and has been widely used in literature to explore linguistic creativity and problem-solving in humans. However, solving complex insight problems may require forms of creativity beyond the linguistic - for example visual and/or spatial creativity. Though empiric evaluation methods for both visual and linguistic creativity exist, these are different: there is no one test through which visual and linguistic performance can be assessed cross-domain.

In order to provide such a tool, a visual adaptation of the RAT was attempted [11]. The formalization of the linguistic RAT supposes that three words are given (w_1, w_2 and w_3), and an answer word (w_{ans}) is to be found. This word is subsequently searched for with the knowledge organization boost from comRAT-C – by using associations, their strength and convergence. In adapting the RAT to the visual domain, the elements of this formalization were considered to more abstractly be sensory elements (e_1, e_2 and e_3), with a related answer element (e_{ans}) to be found.

Specifically for the visual domain, the three elements were considered to be visual elements (e.g. objects), with the answer a visual element as well, For example, such a visual query can be seen in Fig. 2, where visual elements in the three pictures – the chimney, the blacksmith in his workshop and the wood – are meant to elicit a visual association answer (fire).

Fig. 2. An example vRAT query ("chimney-blacksmith-wood"). Answer: "fire".

Two separate studies were conducted to investigate the human response to the created vRAT queries and comparability with linguistic RAT [15].

In Study 1 (n = 38), previous participants to a compound RAT study were presented with 46 vRAT queries in a randomized sequence. Study 2 investigated how people performed when they were administered with vRAT and linguistic RAT one after the other in the same session. Study 2 was administered via two platforms: F8 and Amazon Mechanical Turk (MTurk, http://mturk.com). Participants (n = 170) were presented with 46 vRAT queries (same as Study-1) followed by a set of 24 comRAT-G queries and 24 queries from the Bowden and Jung-Beeman dataset [2].

Table 3 shows the correlations of vRAT scores with linguistic queries from the comRAT-G and Bowden & Jung-Beeman (B-JB) datasets scored for both the studies. In Study-1, a significant correlation of 0.431 was observed between the *vRAT score* and *comRAT-G score*. The significance of this correlation was corroborated by Study-2 where the correlation between *vRAT score* and *comRAT-G score* was significant for each of the platforms independently and both of them combined.

This showed that the initial computational formalization of the RAT for a cognitive AI system can be used to create valid queries in the visual domain, and thus expand the reach of creativity evaluation tools.

Table 3. Correlations between the visual RAT and the linguistic RAT scores

Correlated with *vRAT score*	Study-1	Study-2		
		F8	MTurk	Combined
comRAT-G score	0.431^{**}	0.447^{*}	0.307^{***}	0.331^{***}
B-JB score	0.022	0.395^{*}	0.169^{*}	0.210^{**}
linguistic RAT score	0.202	0.465^{*}	0.266^{**}	0.302^{***}

Correlations significance level indicated as follows: 0.05 - *; 0.01 - **; 0.001 - ***.

7 Discussion and Perspectives

Various cognitive AI systems have been implemented and experimented with – they showed to contribute to improving creativity modeling and measuring tools.

In the creative association domain, comRAT-C was the first system to use the knowledge organization proposed under the CreaCogs framework. comRAT-C was successful in solving the compound RAT queries from the normative Bowden & Jung Beeman dataset and the system's performance significantly correlated with the performance of humans on compound RAT queries. comRAT-G was the first system to computationally generate RAT queries and contributed a large set of generated and validated queries to the existing normative data for RAT. This computational approach to generating new queries has shown to be fruitful by both correlating to human data on a compound normative dataset [14], and also allowing for more refined empirical designs with a better ability to record previously unexplored factors influencing creativity [13].

Before the comRAT-G_F system, functional queries were a theoretical idea [21], the examples of which have been lost to the influence of time and lack of digitalization. No normative data was available for functional RAT queries, thus no researchers could explore this idea further. comRAT-G_F generated a large sized dataset for functional RAT which was successfully evaluated with human participants.

Finally, even formalizing a task in the computational manner required by constructing a subsequent cognitive AI system can have an impact. The visual RAT test is the proof of how this formalization can help grasp a task with more precision, and deploy it in new fields. The visual RAT takes the Remote Associates Test beyond the linguistic domain, creating visual queries which can help gain cross-modal strength when investigating the associative factor in creativity.

In summary, computational cognitive approaches to creativity science can be very fruitful. The formalization and implementation of cognitive AI systems has had and can further have a deep impact on cognitive science tools and models.

As further work on the research arc regarding creativity and association, the authors plan to focus on (i) building a computational solver for the visual RAT, (ii) exploring multiple answer queries in the linguistic RAT, which were suspected to exist but first discovered computationally using comRAT-C and can be addressed in a systemic manner using comRAT-G, and on (iii) taking the RAT to other sensory domains.

References

1. Arden, R., Chavez, R.S., Grazioplene, R., Jung, R.E.: Neuroimaging creativity: a psychometric view. Behav. Brain Res. **214**(2), 143–156 (2010)
2. Bowden, E.M., Jung-Beeman, M.: Normative data for 144 compound remote associate problems. Behav. Res. Methods Instrum. Comput. **35**(4), 634–639 (2003)
3. Colton, S., Wiggins, G.A.: Computational creativity: the final frontier? In: Proceedings of the 20th European Conference on Artificial Intelligence, pp. 21–26. IOS Press (2012)
4. Mednick, S.: The associative basis of the creative process. Psychol. Rev. **69**(3), 220 (1962)
5. Mednick, S.A., Mednick, M.: Remote Associates Test: Examiner's Manual. Houghton Mifflin, Boston (1971)
6. Nelson, D.L., McEvoy, C.L., Schreiber, T.A.: The University of South Florida free association, rhyme, and word fragment norms. Behav. Res. Methods Instrum. Comput. **36**(3), 402–407 (2004)
7. Newell, A.: Unified Theories of Cognition. Harvard University Press, Cambridge (1994)
8. Olteţeanu, A.M.: Two general classes in creative problem-solving? An account based on the cognitive processes involved in the problem structure - representation structure relationship. In: Proceedings of the International Conference on Computational Creativity, January 2014. Institute of Cognitive Science, Osnabrück (2014)
9. Olteţeanu, A.M.: Proceedings of the Workshop on Computational Creativity, Concept Invention, and General Intelligence (C3GI2016), vol. 1767. CEUR-Ws, Osnabrück (2016)
10. Olteţeanu, A.M., Falomir, Z.: comRAT-C - A computational compound remote associates test solver based on language data and its comparison to human performance. Pattern Recogn. Lett. **67**, 81–90 (2015). https://doi.org/10.1016/j.patrec.2015.05.015
11. Olteţeanu, A.M., Gautam, B., Falomir, Z.: Towards a visual remote associates test and its computational solver. In: Proceedings of the Third International Workshop on Artificial Intelligence and Cognition 2015, vol. 1510, pp. 19–28. CEUR-Ws (2015)
12. Olteţeanu, A.M., Schottner, M., Schuberth, S.: Computationally resurrecting the functional remote associates test using cognitive word associates and principles from a computational solver. Knowl. Based Syst. **168**, 1–9 (2019)
13. Olteţeanu, A.M., Schultheis, H.: What determines creative association? Revealing two factors which separately influence the creative process when solving the remote associates test. J. Creat. Behav. (2017). https://doi.org/10.1002/jocb.177
14. Olteţeanu, A.M., Schultheis, H., Dyer, J.B.: Computationally constructing a repository of compound remote associates test items in American English with comRAT-G. Behav. Res. Methods Instrum. Comput. (2017). https://doi.org/10.3758/s13428-017-0965-8
15. Olteţeanu, A.M., Zunjani, F.H.: A visual remote associates test and its validation. Behav. Res. Methods (Submitted)

16. Olteteanu, A.M.: A cognitive systems framework for creative problem solving. Ph.D. thesis, Universität Bremen (2016)
17. Olteţeanu, A.-M.: From simple machines to eureka in four not-so-easy steps: towards creative visuospatial intelligence. In: Müller, V.C. (ed.) Fundamental Issues of Artificial Intelligence. SL, vol. 376, pp. 159–180. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-26485-1_11
18. Olteţeanu, A.M., Falomir, Z.: Object replacement and object composition in a creative cognitive system. Towards a computational solver of the alternative uses test. Cogn. Syst. Res. **39**, 15–32 (2016)
19. Palermo, D.S., Jenkins, J.J.: Word Association Norms: Grade School Through College (1964)
20. Tucker, A.: Applied Combinatorics. Wiley, Hoboken (2006)
21. Worthen, B.R., Clark, P.M.: Toward an improved measure of remote associational ability. J. Educ. Meas. **8**(2), 113–123 (1971)