Towards using Cognitive Word Associates to Create Functional Remote Associates Test Problems

Ana-Maria Oltețeanu
Cognitive Systems, Bremen Spatial Cognition Center
Universität Bremen
Bremen, Germany
Email: amoodu@informatik.uni-bremen.de

Abstract—**Human creativity is assessed with a variety of tools, one of which is the Remote Associates Test. Two linguistic variants of this test exist: the compound and the functional Remote Associates Test. While normative data and solvers for the *compound* RAT exist, data sets of *functional* items are very rarely encountered in the literature. Such data sets would allow (i) a deeper understanding and the simulation of the cognitive associative processes used in creativity tasks and (ii) the comparison of performance and process between the two types of queries. In this paper, an approach to knowledge acquisition and computational generation of functional Remote Associates Test items is explored. Possibilities of cognitive evaluation are discussed.**

Keywords-**creativity; Remote Associates Test; knowledge acquisition; word associates; functional Remote Associates Test; creative problem solving; cognitive systems**

I. INTRODUCTION

Human creativity is assessed in a variety of ways, with a variety of creativity tests and tools [Duncker, 1945], [Wallach and Kogan, 1965], [Kim, 2006], [Qiu et al., 2008]. One such tool is the Remote Associates Test [Mednick and Mednick, 1971] - henceforth addressed as RAT. This test aims to measure a participant's ability to make associations and is normally presented in language (though visual variants have also been proposed [Oltețeanu et al., 2015]). As part of this test, three words are given to the participant, and a fourth word which is related to all the given three needs to be found; for example OPERA, HAND and DISH are given, and SOAP can be an answer.

Though appreciating Mednick's remote association theory [Mednick, 1962], some authors [Worthen and Clark, 1971] have voiced the opinion that Mednick and Mednick's operationalization of the concept - the RAT query items [Mednick and Mednick, 1971] - might not be fully appropriate for the evaluation of the ability to associate, as:

(i) (at least) two distinct types of items were posited to be part of the test and
(ii) not all the items in the test could be found in Palermo-Jenkins word association norms [Palermo and Jenkins, 1964].

Regarding point (i), the two types of items posited to appear as a mix in Mednick's set of queries were (a) items which associate in language, and (b) items which associate based on non-linguistic relationships (or relationships which go beyond language). Worthen and Clark have called these: (a) *structural* and (b) *functional* items. For example, items like BLACK and MAGIC would associate in language, while items like BIRD and EGG would have a functional relationship. This makes the above mentioned RAT query a structural one, as SOAP OPERA, HAND SOAP and DISH SOAP are linguistic associates.

Normative data does exist for queries involving items which associate with the answer in language [Bowden and Jung-Beeman, 2003] – though Bowden and Jung-Beeman call these items *compound* items, because of their ability to form a compound linguistic structure with the answer. Computational solvers like comRAT-C [Oltețeanu and Falomir, 2015] have been shown to answer such queries well, and even correlate with human performance, given knowledge extracted from language corpuses. However, datasets of *functional* items do not exist, nor does normative data on such datasets, to our knowledge. Worthen and Clark redesigned a set of 20 of Mednick's queries as functional items, but this set of queries is hard to retrieve, if it was not completely lost[1].

Also, the second issue mentioned by [Worthen and Clark, 1971], pertaining to the relation between RAT queries and word association norms, has not yet been solved in the creative problem solving literature. In order to deem which RAT problems would be functional, Worthen and Clark analysed how many and which of the queries proposed by Mednick and Mednick had corresponding Palermo-Jenkins word association norms. In their view, RAT items for which such norms could not be found might be inappropriate. Without wanting to go as far, the authors of this paper consider word associate norms to still be relevant to functional RAT items. Also, a dataset of functional items would be useful in providing a comparison point for human performance on compound problems. This could help further refine theories about the process through which (these and other types of) creative tasks are solved.

Thus, in this paper, an approach to cognitive knowledge acquisition and computational generation of func-

---

[1]The former National Auxiliary Publications Service (NAPS) items have been sent to the Library of Congress. The Library of Congress mentions not having received Worthen and Clark's functional items collection from the NAPS - personal correspondence 21st of July 2016. Further correspondence inquiries at the authors' universities has not yet revealed any clues as to whether the items have been preserved.

tional RAT items is put forward, using word associate norms as a centerpoint; however, these are not Palermo-Jenkins associates but the modern set of University of South Florida association norms by [Nelson et al., 2004]. The rest of this paper is organized as follows. Preliminary work on which this paper is based is described in Section II, including (i) the comRAT-C approach of [Olteţeanu and Falomir, 2015] which will be extended here, (ii) the difference between functional and compound items and (iii) the role of word associates from [Nelson et al., 2004]. The procedure of generating queries based on word associates is explained in Section III. The results of applying this procedure are presented in Section IV. These results, current limitations and the further implications of the technique are discussed in Section V. Finally, a conclusion is drawn and further work is proposed in Section VI.

II. PRELIMINARY WORK

comRAT-C [Olteţeanu and Falomir, 2015] is a computational solver of the compound RAT. comRAT-C uses language knowledge and a particular type of knowledge organization to solve compound RAT queries. As language knowledge, n-grams from a language corpus[2] are organized as expressions by comRAT. Expressions are considered to be links between two concepts. comRAT-C thus develops a network of concepts, with the links between them symbolizing an existing expression or compound word in which the words have co-occured, as shown in Figure 1. For example, words SWISS and ALPS are words that have co-occured.

After building this knowledge base network, comRAT-C uses convergence over its knowledge to solve RAT queries. For example, as shown in Figure 1, comRAT-C would be given the query COTTAGE, SWISS and CAKE (items in green) to solve. comRAT-C aims to solve the query by activating the neighborhood nodes (items in blue) of each of the seed concepts (items in green). In this process, comRAT converges upon common items - that is items which co-occur with two (in yellow) or three (in red) of the initial words. The first found 3-item convergence wins. Short of that, 2-item convergences can also be offered as answers. The links can further be weighted using frequency of expressions, thus calculating the probability that a certain answer would occur. comRAT-C can offer multiple answers to compound queries, and it correlates to human performance in solving said queries, as measured by response times and accuracy per query in [Bowden and Jung-Beeman, 2003].

In order to tackle a *functional* form of the RAT in a similar manner as the approach by [Olteţeanu and Falomir, 2015], a source of functional, rather than linguistic connections was required. As [Worthen and Clark, 1971] refer to Palermo-Jenkins word association norms as a source for evaluating the RAT items, we decided to use a dataset for word association

norms in order to extract the required knowledge for functional items. The dataset of [Nelson et al., 2004] provides free association, rhyme and word fragment norms. Free association measures the likelihood that a word may cue another word with minimal contextual constraints in effect [Nelson et al., 2000]. Free association probabilities thus have been shown to provide a relative measure of forward strength, with other associates competing for production [Nelson et al., 2005]. Thus, as can be seen in Table I, which shows data fragments relevant to the task at hand from [Nelson et al., 2004]'s Appendix A, a cue word like ABUNDANCE can yield target words like FOOD, which was mentioned by 5 of the 152 participant that were given this word as a cue.

In [Olteţeanu and Falomir, 2015], two-grams like SWISS ALPS were encoded as the two concept nodes of the expression, with the link between them being marked by the frequency of occurrence of this expression in the corpus. The probability that the answer would be generated, given that a certain query item was given, was calculated as the ratio of favourable times (response appearing in conjunction with item) over the number of times that query item was given. Similarly, the number of subjects producing the target from the University of South Florida association norms (Table I, column 4) can be used here in lieu of frequency; the forward strength, calculated as the number of participants producing the target when given a specific cue over the sample size of participants which were given that cue, can be used as the probability.

III. GENERATING QUERIES BASED ON WORD ASSOCIATES

The approach for generating RAT queries has been the following: the type of knowledge organization of comRAT-C, akin to a semantic network, arranges the concepts which have co-occurred in the same expression or word as linked around each other. While this type of knowledge organization can easily be used to check for convergence starting from 3 initial words, it can also be used to generate queries: each item with more than 3 links connected to it is potentially an answer to such a query. An example of this is shown in Fig. 2. Here, the word SWISS, which was previously a query word, is a potential answer for queries like ALPS, CHOCOLATE, CHEESE.

Using this conceptual insight on knowledge organization and the word associates dataset, a set of functional RAT queries are generated based on word associates, in 3 steps.

First, items corresponding to expressions, concepts and frequencies are extracted from the word associates dataset. Secondly, an answered centred view is applied, organizing items in terms of the answers and their potential query items. A database table is made with each answer and all its potential question items, together with the frequency and probability metrics. Finally, queries are generated for the existing answers using their potential query items, via a combinatorics algorithm from Alan Tucker
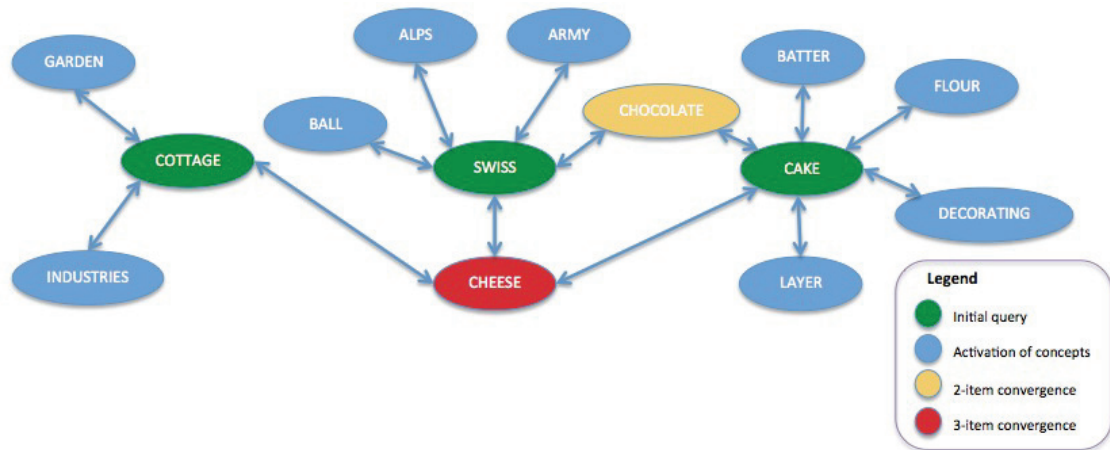
Figure 1. Visual depiction of the comRAT-C knowledge organization (with limited links for visibility), and of activation during a query.

Cue	Target	Sample size	Subjects producing target	Forward strength
Abundance	Famine	152	2	0.013
Abundance	Food	152	5	0.033
Abundance	Full	152	3	0.02
Abundance	Lack	152	4	0.026
Abundance	Large	152	2	0.013
Abundance	Little	152	2	0.013
Abundance	Lots	152	34	0.224
Abundance	Many	152	12	0.079
Abundance	Money	152	6	0.039

[Tucker, 2006]. Because of its relation to comRAT-C, the compound RAT solver, and its emphasis on functional queries generation, we called the system implementing this approach $comRAT - G_F$.
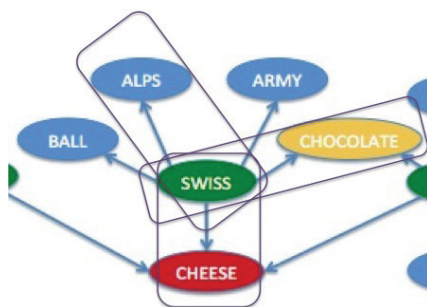

Figure 2. The step from solving to generating queries, taking advantage of the existing knowledge organization structure: SWISS, previously a query word, is the answer to triples of items from its own existing links.

IV. RESULTS

The University of Florida free associations dataset contains 72186 items, thus we obtained the same number of concept to concept pairs (the equivalent of expressions in the comRAT-C system). Items holding less than three associates were removed in the phase of centering potential queries around potential answers. Thus 62137 answer-query item pairs remained after this phase. The query creation phase based on this generated 13 534 865 items. In the following we provide a short analysis of the results.

Table II presents 21 examples of such generated queries. As it can be seen, the queries created with $comRAT - G_F$ look like what Worthen and Clark would call functional queries, in the sense that most of them seem to have a relationship which goes beyond language. Sometimes, the generated queries mix relationships that are purely functional with relationships which are also language relationships, but go beyond language. An example of this is Q4, where the given query items are DAISY, TULIP, VASE and the answer is FLOWER. The relationship within the item pair (DAISY, FLOWER) and the item pair (TULIP, FLOWER) is functional: it defines the category of the query item through the answer, thus constituting an *is a* relationship. Meanwhile, the relationship within the item pair (VASE, FLOWER) is both a (i) linguistic relationship, because of the compound FLOWER VASE, and (ii) a functional relationship, as flowers go in vases, thus the relationship of these items appears in reality, not just in language.

Some compound queries are still part of the results - for example Q6 is (also) compound - BANK ACCOUNT, BANK TELLER, BANK VAULT. Thus queries the items of which can be found in word associates lists are not necessarily

all functional.

Some relationships between the answers and the query items are of synonymy. However this synonymy can span different semantic domains. For example, in Q9, a relation of synonymy exists between each of the 3 query items and the answer ATTENDANCE ≡ PRESENT(CE), CONTEMPORARY ≡ PRESENT (TIME) and GIFT ≡ PRESENT. However, being present at an event, being of the present time and giving a gift to someone are different semantic domains, despite the apparent similar form of the answer which is a synonym for each three of the query item instances. The three versions of the word PRESENT retrieved by the query are thus homonyms to each other.

Multiple mappings of queries to answer are generated in this way. For example, queries 15 and 19 have the same answer, with different probabilities, constructed as a function of forward strength[3]. This allows future explorations of whether the different probabilities allow for faster response times in the various queries, when these queries are given to human participants.

The queries thus created demonstrate a wide range of frequency and probabilities. The smallest probability of answering the query, calculated as the average of the forward strength of the 3 items equally influencing the outcome, is 0.0101. A set of 5 queries have this probability, including BEAD,IRON,STYLE – answer DRESS. The largest probability is 0.8223 - query SALMON,TROUT, TUNA – answer FISH. The distribution of items by probability is numerically described in Table III.

Frequency is calculated as the number of people providing the target when being given the cue. The smallest frequency is 2, for $query\_word - answer\_word$ pairs like DECEPTION-SPY, BEAD-DRESS, LEAN-STRONG, etc. The largest frequencies are in the 134-165 range, for example: 136 - EXHAUSTED-TIRED and ROW-BOAT, 146 - HUSBAND-WIFE, ATTEMPT-TRY, 163 EAST-WEST, 165 WEEP-CRY. The distribution of these is numerically described in Table IV.

V. DISCUSSION

As seen above, the current approach can generate functional RAT items, validating the use of word associates knowledge for the creation of such queries. Some queries thus generated can indeed also stand as compound items. However, this can be remedied with relative ease in the case in which functional associates which have no linguistic relationship are required (though this criterion would be tougher than the ones imposed by Worthen and Clark): compound items alone could be generated, then recognized and extracted from the functional items database.

The main limitation which seems to influence the quality of the queries is related to the suitability of the University of South Florida association norms dataset [Nelson et al., 2004] for the task at hand. This dataset of

[3]Each of the three query items is considered to be equally contributing when calculating this probability, but different assumptions can also be modelled.

word associates is very useful for, amongst others, ensuring experimental control in a variety of tasks, and word associates can definitely be used as a data source when generated functional items (as shown above). However, the fact that this particular dataset only collected one target per cue word for each participant makes the closeness of the cue-target pairs be relatively high. Thus such queries might be easier than ones created by, say, using targets that have been ellicited as the third word after having given the cue word. The only measure we currently have for "harder" queries or more distant associates are the targets produced with a lower frequency - i.e. by a smaller number of people.

Our experiments endorse the approach itself, and the use of free associates for creating functional RAT problems. However, access to (or the collection of) a dataset of word associates in which more than one target word is allowed for each participant could potentially create much more interesting queries, that is queries in which the items are more remote conceptually.

Different ways of improving the quality of the results come to mind, other than by using (or previously collecting) a multi-target dataset. As previously observed, some relations are of synonymy. Removing synonymy-based relations is a possible approach to making some of the queries harder. Some relations of synonymy might make certain queries very likely to be answered - for example queries in which the synonymy happens with all three query words in the same semantic domain, like Q1 in Table II. On the other hand, removing all synonymy relations might not allow queries which link into different semantic domains, like Q9, which seems like a good an interesting query, to be generated. An approach aiming to constrict synonymy would have to account for the semantic domain component of the synonymy relationships.

A second way of improving the quality of our generated queries would be to use the list of word associates a second time to make sure that no cue-target relationships appear between the pairs (w_1, w_2), (w_2, w_3),(w_3, w_1) (or their reverses). This would ensure that items are remote from each other, using what the system currently has in its knowledge base as close items. Thus, queries like Q2 would never be generated, because of the close relationship between words $photo$ and $frame$, assuming that this relationship is captured by our dataset.

Despite such current limitations, we consider this approach to have shown that datasets of functional RAT problems can be created computationally. The approach explored here provides the following advantages:

(a) A set of functional items where frequency and probability of elliciting the answer (as a function of forward strength) are known can help check various hypotheses about how word association or the relationship between known concepts affect the solving of a creativity task like the RAT;

(b) Multiple queries that have the same answer are obtained, which allows examining whether some such queries ellicit different performance from human par-

Table II
EXAMPLE OF FUNCTIONAL REMOTE ASSOCIATES PROBLEMS GENERATED USING THIS APPROACH. w_{ans} STANDS FOR THE ANSWER WORD,
WHILE $w_1 - w_3$ STAND FOR THE QUERY ITEMS. ITEMS ARRANGED IN DECREASING PROBABILITY ORDER.

Query no.	w_1	w_2	w_3	w_{ans}	Probability
1	exhausted	sleepy	weary	tired	0.7202
2	frame	photo	portrait	picture	0.6897
3	bassinet	crib	infant	baby	0.6916
4	daisy	tulip	vase	flower	0.6914
5	bulb	dark	dim	light	0.5530
6	account	teller	vault	bank	0.4301
7	cashew	rat	squirrel	nut	0.3518
8	comet	limit	velocity	speed	0.2301
9	attendance	contemporary	gift	present	0.2301
10	capability	function	leadership	ability	0.1101
11	plenty	quantity	site	lot	0.0701
12	car	piston	steam	engine	0.0701
13	A	rate	test	grade	0.0501
14	agent	deception	FBI	spy	0.0163
15	earthquake	war	weakness	fear	0.0114
16	admire	jewel	ocean	beautiful	0.0111
17	cougar	go	learn	fast	0.0110
18	burn	flash	pants	down	0.0110
19	exam	flee	warn	fear	0.0110
20	condition	croak	doctor	dead	0.0105
21	case	fact	threshold	point	0.0103

Table III
DISTRIBUTION OF FUNCTIONAL RAT GENERATED ITEMS BY PROBABILITY

Probability	< 0.1	$0.1 - 0.2$	$0.2 - 0.3$	$0.3 - 0.4$	$0.4 - 0.5$	$0.5 - 0.6$	$0.6 - 0.7$	$0.7 - 0.8$	$>= 0.8$
No.	10 800 063	1 884 946	668 914	142 324	31 711	5 826	975	105	1

Table IV
DISTRIBUTION OF FUNCTIONAL RAT GENERATED ITEMS BY FREQUENCY

Frequency	< 10	$10 - 30$	$30 - 50$	$50 - 70$	$70 - 90$	$90 - 110$	$110 - 130$	$130 - 150$	$>= 150$
$fr(w_1, w_{ans})$	10 187 792	2 253 943	530 668	235 974	170 569	108 480	37 277	9 729	433
$fr(w_2, w_{ans})$	10 167 757	2 289 819	524 780	238 535	156 880	99 398	42 443	13 736	1517
$fr(w_3, w_{ans})$	10 150 287	2 281 637	540 365	258 690	164 917	83 566	41 647	11 613	2143

ticipants, independent of the answer, and whether this is related to association strength;

(c) A further look at this dataset might also yield sets of queries with multiple answers. In this case, the reason an answer might be preferred over another will be interesting to study and determine.

VI. CONCLUSION AND FUTURE WORK

This paper has shown that functional Remote Associates Test items can be created computationally using word associate datasets. Further work is necessary in order to assess and evaluate the queries thus created, and refine the cognitive knowledge acquisition approach.

As future work, we propose:

- Checking whether a comRAT-F system armed with this knowledge of word associates could solve the functional items in [Worthen and Clark, 1971]'s test, if the items are retrieved, or over different sets of functional items, as available;
- Evaluating the difficulty and interestingness of such queries with human participants;

- Improving the queries by acquiring and using multi-target word associates; crowd-sourcing multiple targets to the initial cues might provide for more interesting queries and more distant associations;
- Improving the queries by making sure no relations between query words exist;
- Comparing the performance of human participants in functional RAT problems to that in compound RAT problems.

REFERENCES

[Bowden and Jung-Beeman, 2003] Bowden, E. M. and Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers*, 35(4):634–639.

[Duncker, 1945] Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58(5, Whole No.270).

[Kim, 2006] Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity research journal*, 18(1):3–14.

[Mednick, 1962] Mednick, S. (1962). The associative basis of the creative process. *Psychological review*, 69(3):220.

[Mednick and Mednick, 1971] Mednick, S. A. and Mednick, M. (1971). *Remote associates test: Examiner's manual*. Houghton Mifflin.

[Nelson et al., 2005] Nelson, D. L., Dyrdal, G. M., and Goodmon, L. B. (2005). What is preexisting strength? Predicting free association probabilities, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review*, 12(4):711–719.

[Nelson et al., 2000] Nelson, D. L., McEvoy, C. L., and Dennis, S. (2000). What is free association and what does it measure? *Memory & cognition*, 28(6):887–899.

[Nelson et al., 2004] Nelson, D. L., McEvoy, C. L., and Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, 36(3):402–407.

[Olteţeanu and Falomir, 2015] Olteţeanu, A.-M. and Falomir, Z. (2015). comRAT-C - A computational compound Remote Associates Test solver based on language data and its comparison to human performance. *Pattern Recognition Letters*, 67:81–90.

[Olteţeanu et al., 2015] Olteţeanu, A.-M., Gautam, B., and Falomir, Z. (2015). Towards a visual remote associates test and its computational solver. In *Proceedings of the Third International Workshop on Artificial Intelligence and Cognition 2015*, volume 1510, pages 19–28. CEUR-Ws.

[Palermo and Jenkins, 1964] Palermo, D. S. and Jenkins, J. J. (1964). Word association norms: Grade school through college.

[Qiu et al., 2008] Qiu, J., Li, H., Yang, D., Luo, Y., Li, Y., Wu, Z., and Zhang, Q. (2008). The neural basis of insight problem solving: An event-related potential study. *Brain and cognition*, 68(1):100–106.

[Tucker, 2006] Tucker, A. (2006). *Applied combinatorics*. John Wiley & Sons, Inc.

[Wallach and Kogan, 1965] Wallach, M. A. and Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity-intelligence distinction*. Holt, Rinehart & Winston.

[Worthen and Clark, 1971] Worthen, B. R. and Clark, P. M. (1971). Toward an improved measure of remote associational ability. *Journal of Educational Measurement*, 8(2):113–123.