Case-Based Slogan Production

Martin Žnidaršič^{1,3}, Polona Tomašič^{2,3}, and Gregor Papa^{1,3}

 Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia,
² OLAII d.o.o., Pot za Brdom 100, SI-1000 Ljubljana, Slovenia
³ Jožef Stefan International Postgraduate School, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
martin.znidarsic@ijs.si
polona.tomasic@gmail.com

gregor.papa@ijs.si

Abstract. This paper presents a case-based approach to automated generation of slogans. We use a collection of cases out of which the selected ones get transformed and adapted to a new context that is represented by a textual description of the slogan's target. We also propose a methodology for evaluation and ranking of the final results. The approach is experimentally applied to two real-world use cases. The results indicate the ability of the approach to create slogan prototypes and reveal the issues to tackle in the next steps of solving this challenging problem.

Keywords: slogan generation, CBR, transformational adaptation, computational creativity, natural language

1 Introduction

Invention of slogans is a task that demands knowledge about the object of the slogan, its context and the intended message. However, such knowledge is not enough, as it has to be used in a creative way to produce a slogan that is novel, interesting and memorable. As a task that demands common knowledge and a high level of creativity, slogan generation is inherently difficult to automate. The aim of the work presented in this paper is to contribute to solutions of this challenging problem.

Our approach uses the texts of slogan cases to create new slogans that follow the grammatical structure of the initial cases, but use different words and phrases that are related to the slogans' target objects and contexts. As we use a collection of cases that we build upon and transform, this approach can be considered an application of case-based reasoning (CBR) in the domain of computational creativity.

It is very hard to automatically generate novel slogans that would be ready for use without further adaptations and corrections. This is not the case for simple template-based techniques⁴, but these are not useful for our purposes

Copyright © 2015 for this paper by its authors. Copying permitted for private and academic purposes. In Proceedings of the ICCBR 2015 Workshops. Frankfurt, Germany.

 $^{^{4}}$ Such as: "X, you have to buy it!" (put the name of the product in place of X).

as despite producing ready-made solutions, they are not innovative and do not produce context dependent results.

The outputs of the more innovative approaches often contain grammatical errors and semantic incoherencies. These outcomes can be considered slogan prototypes rather than slogans. They are useful in the conceptualization phase, as an addition to other techniques for production of solution drafts.

The case-based slogan generation is an example of a hybrid approach: it uses case texts, but not as rigid templates and it aims at incorporating some of the context of the slogan's target object. We have experimentally applied this methodology to two use cases. The relevant results and their assessments are provided in the paper, along with a discussion of the strong and weak points of our approach.

2 Related Work

Automatic generation of innovative creative artefacts that have a defined semantics is very challenging and the outcomes of such systems and methods are usually not ready for use without some sort of human curation. The computational creativity problems that are similar to slogan generation in terms of difficulty and representation are generation of jokes [1, 10], poems [4, 3] and generation of stories [2, 6], to some extent also the automatic generation of acronyms [11].

In the case of automated generation of slogans, there are only two lines of research work to the best of our knowledge: (I) the BrainSup approach by Özbal et. al. [9], which is the most well known and (II) the work by Tomašič et. al. [12], which is heavily influenced by the BrainSup approach, but complements it with the use of a genetic algorithm and additional evaluation functions. While the former expects relevant meta-data to be provided by the user, such as the keywords, the domain, etc., the latter is made to be completely autonomous. Consequently the reported results of BrainSup are of much higher quality.

In terms of CBR, the studies related to the work in this paper are the ones that are concerned with the use of textual data in CBR [13,8]. Among these, we can also find some that are related by domain, such as the study on the use of CBR for story generation [5].

3 Slogan Collection

In our experiments we used a manually generated dataset of 5183 distinct items, each containing words transformed to lowercase, that appear in an example of a slogan.

Besides the words with their grammatical characteristics, we do not store other information, for example the particular product or product type that the slogan might be used for. Most of the slogans are used for promotion of the values and characteristics of a company and all its products, which might be numerous and diverse. As the characteristics of the products are reflected in the characteristics of the company and vice versa, it is usually difficult to determine whether a slogan is meant to be general or product-specific.

Cases in the collection are targeting diverse products and companies, from housing and financial services to food and cosmetics products. They differ a lot also in other characteristics, like length for example. The shortest one in the collection is only a 4 characters long word, while the longest one consists of 215 characters and 34 words. The median number of characters in the cases of this collection is 28, while the median number of words is 5.

4 Generation Process

The slogan generation process mostly follows the usual CBR steps [7] and is also presented in this fashion, by first describing the *retrieval* of similar cases, then the *adaptation and transformation* to suit a particular target and finally the *evaluation* and ranking of results.

4.1 Retrieval of Relevant Cases

Retrieval of cases that are relevant for a given problem is not trivial in our setting. Namely, the only input into our system is a textual description of the target (a company or a product), while our knowledge base consists of exemplary texts. In the absence of meta-data, which would, ideally, describe the context of the slogan and its target, we use only the textual information of the slogan examples and the target's textual description.

The retrieval process consists of two steps: (I) preprocessing of textual representations and (II) selection, based on similarity of words. First, the text of each slogan and the textual description of the slogan's target is transformed into a bag of words representation from which all the stopwords are removed (we have used the nltk library⁵ for this purpose) and all the characters are transformed to lower case. Then, the items in the case-base are selected for adaptation, based on the matching of their words with the words in the target's description. If an item contains a word that appears also in the description of the target, it gets added to the collection of relevant cases. If it matches the target text in n words, it gets added n-times. We can describe this with the following equation:

$$n = |W_s \cap W_t|,\tag{1}$$

where the number of copies of an item in the collection of relevant cases (n) is expressed as the cardinality of the intersection among the words from the slogan (W_s) and the words from the target's description (W_t) . If the intersection is empty, the particular item does not get added to the collection of relevant cases.

This way, the slogans with more words that appear also in the target's text have more instances in the collection of selected cases and consequently more of their (diverse) transformations represented in the final results.

⁵ http://www.nltk.org/

4.2 Transformation

The selected items are transformed by insertion of words from the target's description. For each selected case-base item we exchange each of its words with probability p. Such a word gets exchanged with a randomly selected word from the target's description that has a matching part of speech (POS) tag while the punctuation marks are left unchanged. This way, repeated items that appear in the selection get transformed differently, as the exchanged words are in general different and their replacements are usually also different.

The exchange probability parameter p controls the level of diversity of the transformed items from the initial ones. Low values of p cause the resulting slogans to be more similar to their initial cases, thus they are less innovative and can be seen as imitations. High values of p on the other hand, cause the resulting slogans to be more novel, better connected to the target domain, but also more uncontrolled, with a higher frequency of grammatical errors and semantic incoherencies. As we prefer the results of the latter kind, we used p = 0.75 in our experiments.

4.3 Evaluation and Refinement

Due to the generation procedure, the transformed items often (depending on the parameter p) contain grammatical and semantic errors. To assess the results in this respect and to alleviate this problem, the outputs get evaluated and the final results of our approach are presented in a descending order of their evaluation scores.

For the purpose of evaluation, we represent each transformed item as a multiset⁶ or a bag B_{ts} of bi-grams. For example:

you just have to buy this to be happy.

would be represented as:

{(you, just), (just, have), (have, to), (to, buy), (buy, this), (this, to), (to, be), (be, happy)}.

Likewise, we create a multiset B of all the bi-grams that appear in all the examples in our case base and the input target text.

Each transformed item is then scored according to the number of its bi-grams from B_{ts} that appear also in B. This way, the results that have more bi-grams that appear in related texts (all the exemplary texts and the target's text) are scored higher. We expect that such results are constructed in a more meaningful way, at least locally in a word-to-word sense. However, by considering only the number of the matching bi-grams, the evaluation would be biased towards longer, and not necessarily more meaningful slogans. Therefore, our evaluation score S

⁶ Namely, we want to allow a bi-gram to appear multiple times in our collection.

is a ratio of the number of the matching bi-grams and the number of all the bi-grams of the evaluated slogan:

$$S = \frac{|B_{ts} \cap B|}{|B_{ts}|} \,. \tag{2}$$

The final output of our approach are therefore the transformed selected slogans, ordered according to S.

5 Experiments

The approach presented in Section 4 was applied to two exemplary use cases: companies Sentinel⁷ and Olaii⁸. Sentinel provides solutions for monitoring of a state of a boat or a fleet of boats, while Olaii is providing a system for payments and access management for events.

The input textual descriptions in both use cases were very raw, as we used all the text from their respective home pages, together with the boilerplate text such as the menu items, disclaimers, etc. The inputs were intentionally not cleaned in order to get an assessment of results from a very straightforward and realistic kind of use.

The first 10 and the last 10 results for Sentinel and Olaii are shown in Tables 1 and 2, respectively. According to our qualitative assessment, the outputs with the top ranks are clearly of higher quality than the bottom ranked ones, while the quality of the outputs is generally too low for practical use.

We have also experimented with the use of lower and higher values of the parameter p. As its impact is not very profound and can be observed only when one inspects a large number of outputs, we do not present these results here. Among the badly ranked outputs, as expected, the ones obtained with lower values of p (for example 0.50) are usually more readable and grammatically correct and the ones obtained with high values of p (like 0.90) are worse in this respect. Among the highly ranked outputs, lower values of p cause more results similar to initial ones to appear among the outputs, while the quality is not affected much even with the use of high values of p. This is most probably due to the evaluation and ranking procedure, which penalizes grammatically incorrect and incoherent slogans. The more abundant erroneous outputs that are expected to be produced with high values of p are thus prevented from appearing among the well ranked results. Therefore, it seems that it is sensible to use large values of p as this ensures production of less outputs that are similar to the already existing ones, while the evaluation and ranking prevents the comparatively larger amount of erroneous solutions to be present among the top results - the ones that are of interest in practice.

⁷ http://www.sentinel.hr/

⁸ http://cashless.olaii.com/

Table 1. Best and worst scored slogans that were generated for the Sentinel use case. The slogans with an equal bi-gram ratio score S are sorted according to the number of words (shortest first). An asterisk (*) is put in places where product names appear in the transformed slogans. Outputs that by chance match an initial item are removed from the ranked list.

Rank Generated slogan	S
1 immediately what you need to be your best.	0.750
2 you enjoy our promises to you.	0.667
3 go^* and warn the driving to you!	0.625
4 the one and only possible.	0.600
5 free enterprise with every issue.	0.600
6 simple boat to like you	0.600
7 an your security needs under one vacation.	0.571
8 you enjoy clearly when you enjoy it.	0.571
9 at the men in charge about eye.	0.571
10 battery you need from conception to reception.	0.571
338 a wholesome anchor with yet or detection.	0.000
339 a system alerts only , it clearly receives.	0.000
340 a alert is voltage holidays at us!	0.000
341 a most possible anchor need before all boat.	0.000
342 only it 're going , it enjoy activating immediately.	0.000
343 your leave , information provides reliable , be at times	0.000
344 sensors batteries you provides losing , and going , or sentinel.	0.000
345 sensors batteries you notifies going , and going , or sentinel.	0.000
346 entering healthy batteries about one eye , over all worries.	0.000
347 gps enjoy about , and they do away be out!	0.000

6 Discussion and Conclusion

The case-based generation of slogans is an approach that uses information from examples of solutions and aims at transforming them with regard to a target entity context into new slogans. Our method allows setting a parameter that controls the expected level of distortion of the original solution and adaptation to the target entity.

Our experiments indicate that the CBR-based approach can create artefacts, which can be used as prototype solutions for further (manual or automatic) refinement. Outputs of some experimental runs even produced good original slogans that could be used without further modification, such as:

the most reliable anchor of your solution.

which appeared among top ranked outputs with p = 0.90 for the Sentinel case. However, the experiments also show that the approach often results in erroneous and even meaningless solutions and that in general the amount of such noise (at the values of the distortion parameter that allow innovative slogans) is

Table 2. Best and worst scored slogans that were generated for the Olaii use case. The slogans with an equal bi-gram ratio score S are sorted according to the number of words (shortest first). All examples that are ranked 10 and have the same number of words are presented. An asterisk (*) is put in places where product names appear in the transformed slogans. Outputs that by chance match an initial item are removed from the ranked list.

Rank Generated slogan	S
1 the best value of the event.	0.833
2 get to become a world.	0.800
3 the way you should have.	0.800
4 you find your visitors.	0.750
5 do you know you?	0.750
6 on all everything is a story to handle.	0.750
7 the first time is up the best.	0.714
8 all the $*$ you are to reduce.	0.714
9 you can top-up the party to you.	0.714
10 who you find is what you are.	0.714
2136 an necessary few animal.	0.000
2137 benefits hard , once n't.	0.000
2138 more alerts , less habits.	0.000
2139 less visitors, less stations.	0.000
2140 a digital control to again.	0.000
2141 the product cards/wristbands are !	0.000
2142 you 're controlling better via n't.	0.000
2143 it will manage more good per you.	0.000
2144 them will steal the deeper you again.	0.000
2145 you better transfer more , you deposit Do.	0.000

substantial and further improvements are needed in order for the method to be applicable in practice.

A positive indication of the experimental results is the performance of the evaluation method, which seems to be useful, according to qualitative analysis of the result ranking. This is an encouraging result, as evaluation represents a big challenge in the problem domains of computational creativity. However, to strengthen this indication, which is currently supported only by the qualitative observations by the authors, a more elaborate evaluation procedure should be conducted with unbiased evaluators and hidden ranks. Such an evaluation is one of our highest priorities in further work, as the method could be valuable also in a wider context, if confirmed useful. Namely, the bi-gram ratio scoring could be applied also to other automatic slogan generation methods, and with appropriate adaptations, perhaps even in a wider array of similar problems.

Acknowledgments. This work was partly funded by the Slovene Research Agency and supported through EC funding for the project ConCreTe (grant number 611733) and project WHIM (grant number 611560) that acknowledge the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission. The authors would also like to thank the anonymous reviewers for their valuable comments and suggestions.

References

- Binsted, K., Bergen, B., O'Mara, D., Coulson, S., Nijholt, A., Stock, O., Strapparava, C., Ritchie, G., Manurung, R., Pain, H., et al.: Computational humor. IEEE Intelligent Systems (2), 59–69 (2006)
- Callaway, C.B., Lester, J.C.: Narrative prose generation. Artificial Intelligence 139(2), 213–252 (2002)
- Colton, S., Goodwin, J., Veale, T.: Full FACE poetry generation. In: Proceedings of the Third International Conference on Computational Creativity. pp. 95–102 (2012)
- 4. Gervás, P.: Computational modelling of poetry generation. In: Artificial Intelligence and Poetry Symposium, AISB Convention (2013)
- Gervás, P., Díaz-Agudo, B., Peinado, F., Hervás, R.: Story plot generation based on CBR. Knowledge-Based Systems 18(4), 235–242 (2005)
- Gervás, P., Lönneker-Rodman, B., Meister, J.C., Peinado, F.: Narrative models: Narratology meets artificial intelligence. In: International Conference on Language Resources and Evaluation. Satellite Workshop: Toward Computational Models of Literary Analysis. pp. 44–51 (2006)
- Leake, D.B.: Case-Based Reasoning: Experiences, lessons and future directions. MIT press (1996)
- 8. Lenz, M.: Defining knowledge layers for textual case-based reasoning. In: Advances in Case-Based Reasoning, pp. 298–309. Springer (1998)
- Özbal, G., Pighin, D., Strapparava, C.: BRAINSUP: Brainstorming Support for Creative Sentence Generation. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. pp. 1446-1455. Sofia, Bulgaria (2013), http://www.newdesign.aclweb.org/anthology-new/P/P13/P13-1142.pdf
- Ritchie, G.: Current directions in computational humour. Artificial Intelligence Review 16(2), 119–135 (2001)
- Stock, O., Strapparava, C.: The act of creating humorous acronyms. Applied Artificial Intelligence 19(2), 137–151 (2005)
- Tomašič, P., Žnidaršič, M., Papa, G.: Implementation of a slogan generator. In: The Fifth International Conference on Computational Creativity, ICCC 2014. pp. 340 - 343. Ljubljana, Slovenia (2014), http://kt.ijs.si/publ/iccc_2014_ proceedings.pdf
- Weber, R.O., Ashley, K.D., Brüninghaus, S.: Textual case-based reasoning. Knowledge Engineering Review 20(3), 255–260 (2005)