

Automation and Optimisation of Humor Trait Generation in a Vocal Dialogue System

Matthieu Riou

CERI-LIA, Université d'Avignon
matthieu.riou@alumni.univ-avignon.fr

Stéphane Huet

CERI-LIA, Université d'Avignon
stephane.huet@univ-avignon.fr

Bassam Jabaian

CERI-LIA, Université d'Avignon
bassam.jabaian@univ-avignon.fr

Fabrice Lefèvre

CERI-LIA, Université d'Avignon
fabrice.lefevre@univ-avignon.fr

Abstract

This study pertains to our ongoing work about social artificial vocal interactive agents and their adaptation to users. In this regard, several possibilities to introduce humorous productions in a spoken dialogue system are investigated in order to enhance naturalness during interactions between the agent and the user. Our goal is twofold: automation and optimisation of the humor trait generation process. In this regard, a reinforcement learning scheme is proposed allowing to optimise the usage of humor modules in accordance with user preferences. Some simulated experiments are carried out to confirm that the trained policy used by the humor manager is able to converge to a predefined user profile. Then, some user trials are done to evaluate both the nature of the produced humor and its timely and proportionate usage.

1 Introduction

Interactive artificial agents, like spoken dialogue systems, can now support a broad range of applications such as technical support services or reservation systems (for flights, accommodation, restaurant, etc.). For a while, systems used patterns and rules to define their behaviours, e.g. (Rambow et al., 2001). Lately, stochastic-based models have replaced and improved this rule-based approach for all the components of dialogue systems (Young et al., 2013). These more advanced systems offer new possibilities, like a higher variability in their answers or higher flexibility to adapt to specific user preferences. Their downside is that they require a large amount of data to be trained.

In our ongoing work, we are interested in generating traits of humor in the outputs of a spoken dialogue system, to improve user experience and involvement. This paper investigates the opportunities offered by stochastic approaches to train artificial agents to produce humorous answers according to predetermined levels defining their nature and quantity.

Since humor has played an important role in cultural and social life of human beings, similar beneficial consequences can be expected in human-computer interfaces to improve their social competence (Niculescu et al., 2013). Several works attempted to document some theories or explanations on how humor works in general (Mulder and Nijholt, 2002; Bucaria, 2004; Goldwasser and Zhang, 2016) or even questioned the bare possibility to implement it in computers (Ritchie, 2009). Several studies have also been led to define computational rules for generating puns and riddles (Binsted and Ritchie, 1997; Ritchie, 2005; Hempelmann et al., 2006; Anthony Hong and Ong, 2009). Many of the solutions proposed in these papers have inspired our own modules described in the next section. However, while they only studied one phenomenon at a time, the approach presented here combines the existing possible computational ways to produce several humorous traits.

In this work, we are also interested in optimising this new capacity using reinforcement learning. As users can enjoy or reject agent's humor in general, and appreciate a specific type of humor, we want to allow the system to adapt its use of humor mechanisms, both in quantity and quality. We plan to follow an approach similar to the one used at the dialogue management level (Daubigney et al., 2012), so as to let the system choose, at each step, in an informed way, whether or not it is profitable to produce a humor-

ous utterance.

This study is, to the best of our knowledge, among the very first to both compile several means to produce humor in an automatic manner along with a process to train such capacity so as to adapt it to user preferences in the matter. It is worth mentioning also that humor is not considered in our work with the only objective to make laugh but more generally to ease the on-going interaction.

In this paper, Section 2 describes different options we investigated to introduce humor in an artificial agent. In Section 3 a framework to optimise the usage of humor in a goal-directed dialogue system is introduced. Then Section 4 presents the experimental work to evaluate this newly trained humorous system and we conclude in Section 5.

2 Humor Generation Modules

As our main objective here is to be able to optimise the usage of humor traits during dialogues, in this preliminary setup, only four humor generation modules are explored (Desfrançois, 2016). The natural language generation module is in charge of combining the humor modules' output with the next dialogue act selected by the dialogue manager in the smoothest way (several modalities are introduced as pre and post humor glues to standard system's outputs).

The **quote** module finds a humorous citation in relation to the user input, so as to remain close to the context of the dialogue. A corpus of purposeful citations collected on-line has been indexed and a query is built from the keywords available in the user's sentence. Quote results of the information retrieval system are ranked for each query according to their distance to a context vector. This distance is used as an indicator of the interest of using the quote in the situation. A threshold has been manually defined to discard cases where no quote is close enough to the user's input. A history is kept to avoid repeating the same quotes.

The **joke** module has a modus operandi similar to the quote module; it returns the jokes that are closest to the context and keeps a history to avoid repetitions. The indexed jokes are generally longer than the quotes and can be used alone, which presupposes a new user turn before returning to the main flow of the dialogue.

Self-derision is also considered, seen as a humorous signal of low-esteem intended to encourage further use of the system by admitting its er-

rors. The module outputs predefined humorous sentences like "Luckily I am not a human", "I will eventually file a complaint against the guy who programmed the system" when the dialogue manager can assess that it is in a poor situation.

3 Reinforcement Learning Paradigm for Humor Management

The introduction of the new modules described in the previous section makes a real difference in the system behaviour but their effects are complex to evaluate. The use of Reinforcement Learning (RL) techniques for the optimisation of this new capacity can be a good solution. Since each user can appreciate or reject the humor of the conversational agent, the system will be able to adapt its use of humorous mechanisms.

The dialogue manager used in this paper adapts a system presented in (Ferreira and Lefèvre, 2015). It is based on a dialogue management framework based on a Partially Observable Markov Decision Process (POMDP), the Hidden Information State (HIS) (Young et al., 2010). In this setup, the system maintains a distribution over possible dialogue states (the belief state) and uses it to generate an adequate answer. An RL algorithm, the KTDQ learning algorithm (Geist and Pietquin, 2010), is used to train the system by maximizing an expected cumulative discounted reward, according to two types of feedbacks.

The global feedback is given at the end of the dialogue by asking the user if the entire dialogue is a success or not. The social feedback is given at each turn to score the last response only. It is composed of two parts, the score given by the user to this last response, and the turn cost which penalises too long dialogues by adding a negative score for each turn taken. At the end of the dialogue, the policy is updated according to all the collected feedbacks.

In order to decide when to include a given humor trait with the four generation modules, a policy specific to humor is defined. For this purpose, a dialogue server is launched and consults at each turn of the dialogue system a humor manager. This manager is associated with a policy that is learned with simulated users (see Section 3.2). The following section describes the state space of the humor policy.

3.1 Humor State

The state space for humor is defined by five continuous parameters, each associated with Radial Basis Functions (RBF) to parameterize the policy (Daubigny et al., 2012): percentage of the system utterances with a quotation, a joke, a slip of the tongue, a self-derision assertion or without any trait of humor. All these parameters are defined in the $[0; 1]$ range and converted with RBF into 3 values, which results in a 15-dimension state.

The action space itself contains five different actions: four associated with the humor modules and one for avoiding humor. The number of actions could be extended in order to include a new type of humor module, and the framework has means to capitalise on the simpler policy and avoid starting the learning process from scratch again.

Rewards are defined by a simulated user during simulations from a linear interpolation of the dialogue final score $scoreFinalDialogue$ with respect to the goal, and the humor score $scoreHumor$ from the user’s point of view:

$$r_f = wDialogue \times scoreFinalDialogue + wHumor \times scoreHumor .$$

$wDialogue$ and $wHumor$ represent the weights of dialogue and humor scores respectively.

The humor score is derived from the satisfaction score computed from the simulated user:

$$scoreHumor = satisfaction \times MaxReward .$$

Depending on the percentage of humor matches made during the dialogue and its profile, satisfaction is calculated thanks to the number of humorous actions coherent with the user’s preferences. $MaxReward$ is the maximum reward that can be obtained during a dialogue, and so $scoreHumor$ is the reward obtained in a particular dialogue.

3.2 Humor Simulator

For this study an agenda-based user simulator has been extended to take into account humor traits generated by the system. At this point it was not possible to simulate a real appreciation of the quality and pertinence of the generated humor. Hence, only the type and quantity of humor were taken into account. For that purpose, a user’s profile was defined, supposed to represent acceptable quantity of each type of possible humor of a specific user.

Then, the user simulator was able to reward the simulated dialogues by weighting their success in accordance with its defined profile. From all the possible profiles a few mean profiles were defined as gold standards with a moderate but diversified level of humor, and used for the field trials.

4 Experimental Study

4.1 Task Description

Experiments presented in this paper concern a chit-chat dialogue system framed in a goal-oriented dialogue task. In this context, users discuss with the system about an image (out of a small predefined set of 6), and they tried jointly to discover the message conveyed by the image, as described in (Chaminade, 2017). In order to use a goal-oriented system for such a task, the principle which has been followed was to construct, as the system’s back-end, a database containing several hundreds of possible combinations of characteristics of the image, each associated with a hypothesis of the conveyed message. During its interaction with the system, it is expected that the user progressively gives elements from the image, which matches entities in the database. In return, the system selects a small subset of possible entities to inform the user or ultimately provides a pre-defined message to give as a plausible explanation for the image’s purpose. Thus, the user can speak rather freely about the image before arguing briefly about the message. No argumentation is possible from the system’s side, it can only propose a canned message and the discussion is expected to last only around one minute at most.

The task-dependent knowledge base used in the experiments is derived from INT task description (Chaminade, 2017), as well as from a generic dialogue information. The semantics of the task is represented by 16 different act types, 9 slots and 51 values. The lexical forms (53) used to model act types were manually elaborated.

4.2 Humor Manager Configuration

The dialogue system is built accordingly to the proposition of (Young et al., 2010). More recent system architectures are available (most notably based on end-to-end recurrent neural networks) but it is not yet possible to bootstrap them without a training data set, which is our situation here. The system used hereafter has been trained in joint learning of semantic parser based on a zero-shot

learning algorithm combined with the Q-learner RL approach to learn the dialogue manager policy. The humor manager policy is trained alongside the already trained dialogue policy. Ultimately it is foreseeable to train the two policies jointly but to reduce the state space we chose to fix the dialogue policy when learning the humor manager policies.

4.3 Results

To confirm the humor policy learning process, four profiles were defined and implemented in a user simulator. Each profile sets out a ratio of usage expected from each of the module: “uniform” (all modules and do-nothing are uniformly possible); “light” (all modules can only intervene less than 40% of the time, and are uniformly possible between each other); “jokes” (mainly jokes and do-nothing are possible, 30% and 40% respectively, and the others have low probabilities: 10%); “none” (no humor is allowed).

The user simulator tries to enforce these ratios in policy behaviours, but is also subject to the variability of the whole dialogue process and the availability of the various humor modules at each turn. Therefore the simulated trainings have been carried out to check how close the learned policy could be to the initial profiles. Each profile was used in 50 training simulations, of 500 epochs each. On a simulated test of 200 dialogue examples, without exploration, both “none” and “uniform” policies’ distribution are almost identical to the definition of the profiles (exactly for “none” and less 1 point of difference for each type of humor for “uniform”). “Jokes” and “light” presents variations between 1 and 10 points.

The previous experiment only allowed us to confirm the adequacy between a user preference and a trained policy for humor usage. In a second step the whole system with humor generation mechanisms was tested in user trials. To this end, two profiles have been selected, “light” and “uniform,” and their policy used in a system for a test of 124 dialogues each. 14 participants were recruited to evaluate the system with humor (they all tested the two profiles). They all already experimented the system in its baseline version, i.e. without humor. After each dialogue the users were prompted to answer a short survey with 6 questions :

- Success: “Was the task successful?” (0/1)

- System Understandability: “Was the system easy to understand?” [0,5]
- System Understanding: “Did the system understand you well?” [0,5]
- Humor Identification: “Do you think you identified when the system was making humor?” [0-5]
- Humor Impact: “Do you think humor had a favorable impact on your system perception?” [0-5]
- Humor Quantity: “Are you satisfied with the amount of humor produced by the system?” [0-5]

Table 1 shows the average results of the tests in user trials. Lines 1 to 3 display the overall number of tests, the success rate and the average cumulative rewards of the dialogue manager policy, respectively. The remaining lines provide the subjective scores made for each of the last 5 questions (task success is the first question). The integration of humor with the “light” or “uniform” profiles leads to very competitive success rates (83 % and 89 %) compared to what is observed without humor (86 %). These results, confirmed by the close values measured for the cumulative rewards for the three setups, show that humor does not disturb too much the dialogue system, especially since those differences were not significant¹. There is also no significant¹ differences concerning the system understandability.

Interestingly, the judgments made over the understanding ability of the system are significantly¹ higher for the profiles with humor than the baseline, supporting the interest of introducing this social competence. Let us note that the use of the humor generation modules was easily identified by users (4.4 and 4.7/5). Finally, the last two questions with respect to the impact and quantity of humor show that judges do not really have a preference between the “light” and “uniform” profiles.

5 Conclusion

In this paper, several possibilities to integrate mechanisms to produce humorous utterances in an interactive artificial agent were introduced. A

¹Statistical significances were analyzed with a two-tailed Welch’s t-test. Results were considered statistically significant with a p-value < 0.001.

Model	No humor	“Light” profile	“Uniform” profile
Tests (#)	72	124	124
Success rate (%)	86	83	89
Average cumulative reward	10.2	9.8	10.7
System understandability	4.4	4.6	4.5
System understanding	2.6	3.3	3.5
Humor identification	—	4.4	4.7
Humor favorable impact	—	3.2	3.0
Humor quantity	—	3.5	3.5

Table 1: Evaluation of several profiles for humor generation policy.

two-step process has been devised. First, regular-enough humorous mechanisms have been identified, formalised and automated. Second, those mechanisms have been implemented in a dialogue system and their usage optimised by means of reinforcement learning and on-line adaptation learning approaches. To evaluate the social competence increase of artificial agents endowed with humor, evaluations with real users have been conducted. They allowed us to confirm that dialogue success rate is maintained at a comparable level while the system was generally judged more pleasant.

We have many other challenges ahead. The humor generation modules are in their initial states, and the user trials have been very instructive in highlighting several ways of improvements that will be pursued. Likewise, the optimisation process is currently limited to the nature and quantity of the generated humor. We are investigating an enlargement of the humor state so as to encompass more contextual information enabling the policy to react with greater opportunity.

Acknowledgments

This work has been partially carried out within the Labex BLRI (ANR-11-LABX-0036). The authors would like to thank Thomas Desfrancois for his contribution to this work.

References

- Bryan Anthony Hong and Ethel Ong. 2009. [Automatically Extracting Word Relationships as Templates for Pun Generation](#). In *NAACL HLT Workshop on Computational Approaches to Linguistic Creativity*, pages 24–31, Boulder, Colorado. Association for Computational Linguistics.
- Chiara Bucaria. 2004. [Lexical and syntactic ambiguity as a source of humor: The case of newspaper headlines](#). *Humor*, 17(3):279–310.
- Thierry Chaminade. 2017. [An experimental approach to study the physiology of natural social interactions](#). *Interaction Studies*, 18(2):254–276.
- Lucie Daubigny, Matthieu Geist, Senthilkumar Chandramohan, and Olivier Pietquin. 2012. A comprehensive reinforcement learning framework for dialogue management optimization. *Selected Topics in Signal Processing*, 6(8):891–902.
- Thomas Desfrancois. 2016. [Apprentissage automatique d’humour pour les système de dialogues vocaux \(automatic learning of humor production for the vocal dialogue systems\)](#). Master Thesis.
- Emmanuel Ferreira and Fabrice Lefèvre. 2015. Reinforcement-learning based dialogue system for human-robot interactions with socially-inspired rewards. *Computer Speech & Language*, 34(1):256–274.
- Matthieu Geist and Olivier Pietquin. 2010. Managing uncertainty within value function approximation in reinforcement learning. In *Active Learning and Experimental Design workshop (collocated with AIS-TATS 2010)*, volume 92.
- Dan Goldwasser and Xiao Zhang. 2016. [Understanding Satirical Articles Using Common-Sense](#). *Transactions of the Association for Computational Linguistics*, 4:537–549.
- Christian F Hempelmann, Victor Raskin, and Katrina E Triezenberg. 2006. [Computer, Tell Me a Joke ... but Please Make it Funny: Computational Humor with Ontological Semantics](#). *FLAIRS Conference*.
- Matthijs P Mulder and Antinus Nijholt. 2002. [Humour Research: State of the Art](#). Technical report, University of Twente.
- Andreea Niculescu, Betsy van Dijk, Anton Nijholt, Haizhou Li, and Swee Lan See. 2013. [Making social robots more attractive: The effects of voice pitch, humor and empathy](#). *International Journal of Social Robotics*, 5(2):171–191.
- Kim Binsted and Graeme Ritchie. 1997. [Computational rules for generating punning riddles](#). *Humor - International Journal of Humor Research*, 10(1).

- Owen Rambow, Srinivas Bangalore, and Marilyn Walker. 2001. Natural language generation in dialog systems. In *HLT*.
- Graeme Ritchie. 2005. [Computational Mechanisms for Pun Generation](#). In *Proceedings of the 10th European Natural Language Generation Workshop*.
- Graeme Ritchie. 2009. [Can Computers Create Humor?](#) *AI Magazine*, 30(3):71.
- Steve Young, Milica Gašić, Simon Keizer, François Mairesse, Jost Schatzmann, Blaise Thomson, and Kai Yu. 2010. The hidden information state model: A practical framework for pomdp-based spoken dialogue management. *Computer Speech and Language*, 24(2):150–174.
- Steve Young, Milica Gašić, Blaise Thomson, and Jason D. Williams. 2013. POMDP-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.