



**HAL**  
open science

## JOKER CHATTERBOT

Guillaume Dubuisson Duplessis, Vincent Letard, Anne-Laure Ligozat, Sophie Rosset

► **To cite this version:**

Guillaume Dubuisson Duplessis, Vincent Letard, Anne-Laure Ligozat, Sophie Rosset. JOKER CHATTERBOT: RE-WOCHAT 2016 - SHARED TASK CHATBOT DESCRIPTION REPORT. Workshop on Collecting and Generating Resources for Chatbots and Conversational Agents - Development and Evaluation, May 2016, Portorož, Slovenia. hal-01310204

**HAL Id: hal-01310204**

**<https://hal.science/hal-01310204>**

Submitted on 4 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ***JOKER CHATTERBOT***

## **RE-WOCHAT 2016 - SHARED TASK**

### **CHATBOT DESCRIPTION REPORT**

<b>Guillaume Dubuisson Duplessis</b> LIMSI, CNRS, Université Paris-Saclay, F-91405 Orsay gdubuisson@limsi.fr	<b>Vincent Letard</b> LIMSI, CNRS, Univ. Paris-Sud, Université Paris-Saclay, F- 91405 Orsay letard@limsi.fr	<b>Anne-Laure Ligozat</b> LIMSI, CNRS, ENSIIE, Université Paris-Saclay, F-91405 Orsay annlor@limsi.fr	<b>Sophie Rosset</b> LIMSI, CNRS, Université Paris- Saclay, F-91405 Orsay rosset@limsi.fr
--	--	---	--

#### **Abstract**

The Joker chatterbot is an example-based system that uses a database of indexed dialogue examples automatically built from a television drama subtitle corpus to manage social open-domain dialogue.

#### **1 Joker Chatterbot General Description**

The Joker chatterbot is part of the Joker project which aims at building a generic intelligent user interface providing a multimodal dialogue system with social communication skills including humor and other social behaviors (Devillers et al. 2015). This project is primarily interested in entertaining interactions occurring in a social environment (e.g., a cafeteria).

The Joker chatterbot targets dyadic social open-domain conversations between a human and the system. It is based on a conversational strategy that has been automatically authored from a corpus of television drama subtitles (Dubuisson Duplessis et al., 2016). It aims at generating surprising and entertaining contributions in response to a human utterance in order to maintain the human participation in the dialogue.

Dialogues with the chatterbot consist in social open-domain conversations carried out in English through a text-based web interface. Dialogues are structured as a sequence of pairs of utterances: one from the human followed by the response of the system.

#### **2 Joker Chatterbot Technical Description**

The conversational strategy behind the chatterbot belongs to the category of example-based dialogue modeling that aims at using a database of semantically indexed dialogue examples to manage dialogue (Lee et al., 2009). One feature of our approach is the complete automation of the conversation strategy authoring process from the creation of the database of dialogue examples based on a corpus of dialogues to the conversational management process, see (Dubuisson Duplessis et al., 2016) for a complete account.

The main purpose of the dialogue management process is to select an appropriate response from a database of dialogue examples given the human utterance, and adapt it by taking into account the current dialogue. To this end, our approach discerns three main steps: (1) the selection of candidate system responses from the database of examples, (2) the selection of the most appropriate response, and (3) the transformation of the selected response by taking into account the human utterance and the history of dialogue.

Our dialogue example database is built from subtitles coming from a variety of television drama genres including comedy and sci-fi series (e.g., “The Big Bang Theory”, “Malcolm In the Middle”, “Real Humans”). 274,227 dialogue examples have been extracted from a corpus of 432 subtitle files.

### 3 Sample Dialogue Sessions

Table 1 presents a sample dialogue session of real user interaction with the Joker chatterbot. It has been collected in the context of the RE-WOCHAT shared task via a web interface. This session shows the emergence of a dialogue that can be qualified as roughly coherent, despite the fact that our system focuses on the local coherence of dialogue. Indeed, it only takes into account the last user utterance to select its response.

This example shows a certain ability of the system to select surprising utterances (4, 6, 12). It also points out the selection of incoherent utterances such as 8 and 14. Additional examples and discussions about the capabilities of the system can be found in (Dubuisson Duplessis et al., 2016).

Table 1: Sample Session with the Joker Chatterbot

1	USER: hi there !
2	SYSTEM: hi
3	USER: how are you joker ?
4	SYSTEM: i 've got to get back to work
5	USER: yeah i 've got to get back to work also
6	SYSTEM: i 'll go with you
7	USER: yeah sure
8	SYSTEM: hey
9	USER: how are you feeling today ?
10	SYSTEM: good why
11	USER: i just want to know
12	SYSTEM: if you guys are organizing something
13	USER: no no , we are not organizing anything
14	SYSTEM: printing a story that has become a topic of conversation

### 4 Additional Information

This chatterbot system has been previously used to collect a corpus of 41 open-domain textual dialogues with 27 human participants (Dubuisson Duplessis et al., 2016). This corpus is freely available at the URL: <https://ucar.limsi.fr>. An evaluation of the system was carried out involving objective measures and subjective self-reports (Dubuisson Duplessis et al., 2016). This has made it possible to discuss the general capabilities of the system (regarding understandability, repetition and coherence of system responses) as well as the subjective evaluations of human participants (in terms of amusement, engagement and enjoyability).

The current version of the system is available online at the URL: <https://webjoker.limsi.fr/RE-WOCHAT>.

#### Acknowledgments

This work was partly funded by the JOKER project and supported by ERA-Net CHIST-ERA, and the “Agence Nationale pour la Recherche” (ANR, France).

#### References

- Devillers, L., Rosset, S., Dubuisson Duplessis, G., Sehili, M., Béchade, L., Delaborde, A., Gossart, C., Letard, V., Yang, F., Yemez, Y., Türker, B., Sezgin, M., El Haddad, K., Dupont, S., Luzzati, D., Estève, Y., Gilmartin, E., and Campbell, N. (2015). Multimodal data collection of human-robot humorous interactions in the joker project. In 6th International Conference on Affective Computing and Intelligent Interaction (ACII).
- Dubuisson Duplessis, G., Letard, V., Ligozat, A.-L., Rosset, S. (2016). Purely Corpus-based Automatic Conversation Authoring. In 10<sup>th</sup> International Conference on Language Resources and Evaluation (LREC).
- Lee, C., Jung, S., Kim, S., and Lee, G. G. (2009). Example-based dialog modeling for practical multi-domain dialog system. *Speech Communication*, 51(5):466–484.