Conversational AI and Knowledge Graphs for Social Robot Interaction

Wilcock, Graham

IEEE
2022-03-07

Wilcock , G & Jokinen , K 2022 , Conversational AI and Knowledge Graphs for Social Robot Interaction . in 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2022) . IEEE , pp. 1090-1094 , 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2022) , Sapporo , Japan , 07/03/2022 . https://doi.org/10.1109/HRI53351.2022.9889583

http://hdl.handle.net/10138/353846
https://doi.org/10.1109/HRI53351.2022.9889583

publishedVersion

Conversational AI and Knowledge Graphs for Social Robot Interaction

Graham Wilcock
CDM Interact
Helsinki, Finland
graham.wilcock@cdminteract.com

Kristiina Jokinen
Artificial Intelligence Research Center
National Institute of Advanced Industrial Science and Technology
Tokyo, Japan
kristiina.jokinen@aist.go.jp

***Abstract*—The paper describes an approach that combines work from three fields with previously separate research communities: social robotics, conversational AI, and graph databases. The aim is to develop a generic framework in which a variety of social robots can provide high-quality information to users by accessing semantically-rich knowledge graphs about multiple different domains. An example implementation uses a Furhat robot with Rasa open source conversational AI and knowledge graphs in Neo4j graph databases.**

***Index Terms*—social robots, conversational AI, knowledge graphs, graph databases**

I. Introduction

Ability to communicate in natural language has become more important in human-robot interactions with the recent increase in systems and applications that deal with digital data and assist humans with practical tasks. Rapid development in robot technology has produced a wide variety of robots ranging from advanced industrial robots like Kuka to humanoid social robots like Nao, Erica and Furhat, and cute robot pets like Aibo. At the same time, conversational AI has produced solutions which can take care of simple tasks and chat with users in a conversational manner, with commercial products like IBM Watson, Alexa, Siri, Cortana, and Google Home.

Interaction technology and robotics communities have long worked on natural spoken language-based human-robot interaction, but from different starting points: despite the success of spoken dialogue systems and robot agents in their respective research fields, integration of the two is still rare. We believe there are two main reasons for this: there has been very little interaction between the robotics and spoken dialogue communities until recently, and the de facto standard Robot Operating System (ROS) has paid very little attention to developments in state-of-the-art speech technology.

However, with maturing concepts and methods, integrated approaches to implementing and modelling conversational dialogues on robot platforms have started to gain interest. Some recent tutorials[1] and workshops[2] address these issues. Moreover, standards for interoperability and data sharing are also being developed for application purposes, e.g. in OMG (Object Management Group)[3] and in ISO.

The second author acknowledges that the study is based on results obtained from Project JPNP20006 commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

Sophisticated interaction models and robot implementations are critical when developing practical social robot applications and tools, e.g. robot companions or virtual coaches that are useful for assisting humans in their everyday tasks: providing information, entertainment, and contact, and supporting users in maintaining healthy and active life. An important factor is the embodied nature of robots, the physical embodiment and the context in which the interaction takes place. The social robot is not only a sophisticated computer with internet access and computing power, but an agent which interacts with the user like an autonomous partner, and has presence in the conversational setting [1]. It is important to recognize and generate engagement, and also build long-term relationships between humans and robots.

The contributions of the paper relate to integrating robot and dialogue technology with recent cutting-edge knowledgebase technology. First, we link conversational AI to social robotics platforms to enable more flexible interaction. Second, we enhance the robot dialogue capability by including knowledge graphs as the knowledge source for the robot to talk about a particular domain, and we use graph database queries to allow quick and efficient ways to query the knowledge. Finally, we demonstrate these advances on the Furhat robot platform using a small database of restaurants in the center of Tokyo.

The three fields of social robotics, conversational AI, and graph databases have developed in rather separate research communities. In this paper we discuss an approach in which the three areas intersect. The paper is structured as follows. Section II mentions some of our previous work in these fields. Section III summarizes key points in Rasa open source conversational AI, and describes how knowledgebase actions can be used in Rasa. Section IV describes how to use Rasa with knowledge graphs stored in Neo4j graph databases, and discusses ways that knowledge graphs can be used to give better dialogue responses and more intelligent interaction. Section V presents a working demo example of CityTalk with a Furhat social robot, Rasa conversational AI and knowledge graphs in a Neo4j graph database. Section VI concludes.

A. Background

Our aim is to develop a generic framework in which a variety of social robots can provide high-quality information to users by accessing semantically-rich knowledge graphs about multiple different domains. In particular, there is currently a lot of interest in exploiting knowledge graphs that can be enriched with semantic context information.

So far there has not been much visible work on connecting social robots to knowledgebase services or conversational AI systems like Rasa. We know of earlier work connecting Nao robots to IBM Watson[4] and our own previous work connecting Nao robots to PyDial or to basic care-giving tasks (Section II). We have now connected a Virtual Furhat social robot to Rasa conversational AI and knowledge graphs in Neo4j graph databases (Section V).

1) Social Robots: Social robots interact with humans and can take into acount social aspects of conversational interaction, such as understanding the user's intents, providing feedback on user utterances, enabling appropriate turn taking and providing emotionally appropriate affective responses. The following characteristic features of social robots are listed by the second author in [1]:

1) Use of multimodal signals (speech, gaze, gesture, body) to communicate with the human,
2) Observations of human multimodal behaviour (speech, gaze, gesture, body) to learn human intentions and to react appropriately,
3) Autonomous decisions of actions,
4) Independent moving in a 3-dimensional world,
5) Receiving and sharing information via Internet and IoT.

While there has been much work on the first four items, less attention has been paid to the last item, which this paper now addresses.

2) Conversational AI: Conversational AI has recently made rapid advances, not only in deep learning tasks to enable question answering but also integrating speech to explore end-to-end dialogue systems and including external semantic knowledge such as WikiData. Important issues to be addressed concern integration of spoken language in the chatbot instead of text, as well as the use of knowledge graphs to enable rich contextual information to be included in dialogue modelling. Context includes dialogue history and domain information as well as environmental information obtained via sensors and is relevant for maintaining coherence in long-term interactions.

Various dialogue engines exist, such as Amazon Alexa and Google Assistant, with different strengths and weaknesses. The choice between the alternatives usually boils down to the simple question of how the system uses dialogue knowledge to produce the most appropriate response to the user. We chose Rasa [2] since it has several appealing features [3], being open source and scalable, with a modular and extensible micro-services architecture. Moreover, it uses state-of-the-art NLU and processes user messages in a sequence of components (the NLU pipeline) which allows the components to be changed from the default sequence to a customized configuration. In a similar manner, there is also a possibility to customize the dialogue policies, i.e. sets of actions that the system takes at each step in a conversation according to the needs of the system.

3) Graph Databases: Graph databases [4] have recently moved from academic research into mainstream use by global companies and organisations. In graph databases, relationships between objects are themselves primary items in the database. This has made graph databases very successful in modelling and tracking social networks (friends of friends) and this in turn has led to breakthroughs in fraud detection (conspiracy circles). However, graph databases can be used to store many other types of knowledge.

For our purposes, the advances in graph databases are important because they provide powerful support for advanced knowledge graphs. Like graph databases, knowledge graphs have also moved from academic research into widespread use by companies and organisations, for example playing a key role in information technology for smart cities.

There are many alternative graph database systems, but we chose Neo4j from a leading graph database provider[5] that offers free cloud services for prototyping small-scale databases. Neo4j databases can be deployed both locally with Neo4j Desktop and in the cloud with Neo4j AuraDB.

II. Previous work

This section summarizes some of our previous work at the intersections of social robotics, conversational AI, and knowledge graphs.

A. CityTalk

An example of work at the intersection of social robotics and conversational AI is the CityTalk robot dialogue system developed by the first author [5]. CityTalk used Nao, a talking and gesturing humanoid social robot, to give information to tourists about local hotels and restaurants. For conversational AI, CityTalk used PyDial [6], a statistical dialogue system toolkit developed at Cambridge University. The experience of using PyDial with Nao is described by [7].

The version of CityTalk described by [5] and [7] lies at the intersection of social robotics and conversational AI, but it did not use knowledge graphs or graph databases. Like PyDial, it stored hotels and restaurants data in SQLite databases.

A new version of CityTalk was recently developed by the first author. PyDial was replaced by Rasa conversational AI, and the SQLite databases were replaced by knowledge graphs stored in a Neo4j graph database. These changes are described by [8]. The Nao robot was replaced by Rasa interface channels (Rasa X and Slack), so this work combined conversational AI and knowledge graphs, but not social robotics.

B. AIRC Robot Interaction System

The second author's work on the AIRC Robot Interaction System [9] combined social robotics and knowledge graphs. The system uses knowledge graphs to enable Nao robots to provide instructions to novice care-givers in the eldercare domain. The system supports communication between users and social robots on topics that concern the experience and knowledge of people in service industries, especially in elder-care services. Figure 1 is a screenshot from a short video[6] in which the robot explains basic care-giving tasks.

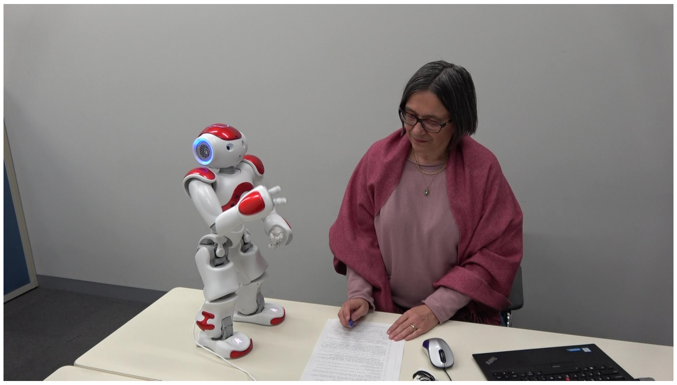

Fig. 1. A Nao robot in the AIRC Robot Interaction System.

The challenge for this type of social robot agents is that they need to be furnished with knowledge that enables them to communicate in a flexible and natural manner about a variety of tasks and care-giving activities. By systematizing and formalizing domain experts' knowledge into knowledgebases, it is possible to structure the knowledge and make the sometimes tacit knowledge explicit and available for reasoning.

While there are standards and best-practice instructions, it is essential to analyse the ways people act and interact in real situations, so as to capture not only a general view of the tasks, but also variation in individual behaviours. Furthermore, it is important to gather information on the knowledge that domain experts possess of the relevant facts, events, and objects, and thus be able to represent the knowledge of the community. This requires that knowledgebase creation must be easy and light, and the knowledge format must be flexible enough to be dynamically extended and modified by new knowledge.

III. Rasa and Conversational AI

This section summarizes some key information about Rasa conversational AI [2]. For up-to-date documentation refer to the online docs[7]. A practical guide to implementing Rasa applications with example code is given by [10].

A. Transformer-based dialogue processing in Rasa

Rasa open source conversational AI uses transformer-based dialogue processing for natural language understanding (NLU) and for deciding system responses (Dialogue Policy).

Rasa performs NLU in a pipeline that can be configured with multiple components depending on the requirements of the application. The key NLU tasks are classifying user intents and recognizing entities mentioned in user uttereances. Although these tasks are usually done by two separate components, Rasa provides a Dual Intent and Entity Transformer (DIET) component [11] that handles both tasks together.

Apart from NLU tasks, the key dialogue management task is to select the most appropriate system response depending on the context. Rasa provides a Transformer Embedding Dialogue Policy (TED) component [12] to handle this task.

The dialogue context includes the current dialogue state, which is held in short-term memory as a set of slots. Each slot holds the current value of some feature that is important for the dialogue task. Rasa supports form-filling, in which the system asks questions and the user provides answers until the required slots are all filled, but the TED dialogue policy aims to be flexible in case the user says something unexpected during form-filling. The system first responds to the unexpected input, and then resumes the form-filling questions.

B. Knowledgebase Actions in Rasa

Rasa applications can access world knowledge in external databases and APIs using developer-written Custom Actions. For example, a restaurants search custom action could access an SQLite database of restaurants and a hotels search custom action could access a MySQL database of hotels.

Rasa also provides actions for accessing knowledgebases that are loaded into memory from JSON files. We prefer these knowledgebase actions[8] because they offer a more domain-independent approach to accessing external world knowledge. Instead of domain-specific custom actions, these generic knowledgebase actions can make queries about many different types of objects in the knowledgebase. However, instead of loading the knowledgebase into memory, we use knowledge graphs held in external Neo4j graph databases.

The generic knowledgebase actions provided by Rasa can search the knowledgebase for all objects with a desired set of properties, and they can retrieve any requested properties of a given object. These two actions are very useful in practice, for example in restaurant dialogues users typically first ask for restaurants in a certain location serving a desired cuisine, and then ask about the price range or the address or phone number of one or two of the restaurants that were found.

Of course, the range of different kinds of information held in knowledgebases is huge. For example in healthcare applications, users may ask for recipe recommendations or nutrition information, or about suitable daily exercises, or seek information about best practices in care-giving.

IV. Neo4j and Knowledge Graphs

As mentioned in the Introduction, knowledge graphs are now in widespread use by major companies and organisations. For a comprehensive recent introduction to the theory and practice of knowledge graphs see [13].

A. An interface between Rasa and Neo4j

The knowledgebase actions provided by Rasa (described in Section III-B) can access in-memory knowledgebases loaded from JSON files. To enable access to knowledge graphs in an external Neo4j graph database, a new developer-written interface module between Rasa and Neo4j is required. A good starting point is the example code made available[9] by the authors of [10] for their example application.

An interface module was developed by the first author to enable the Rasa version of CityTalk (described in Section II-A) to query the restaurant and hotel knowledge graphs stored in the CityTalk Neo4j database. For queries to a graph database, the interface must not only enable access to the nodes and their properties, but must also be able to follow the relationship paths between the nodes in the graph.

B. Adding semantic metadata

Using external graph databases like Neo4j instead of in-memory knowledgebases allows scaling up to extremely large knowledge graphs. It also supports the addition of multiple layers of semantic metadata [14] which can be used to provide more intelligent responses in the dialogue.

One type of semantic metadata is a taxonomy, in which subtypes and supertypes are arranged in a semantic hierarchy. For example, in the CityTalk knowledge graph for restaurants a taxonomy of cuisines can show that French, German and Italian cuisines are subtypes of European cuisine, and Chinese, Japanese and Korean cuisines are subtypes of Asian cuisine. If a user asks for a specific cuisine that is not available in the requested location, the system can use the taxonomy to suggest possible alternatives that may be more appropriate than purely random suggestions. In the earlier version of CityTalk this was not possible as different cuisines were simply different strings with no hierarchical relationships.

Another type of metadata that can be added in a knowledge graph is data lineage, which records where the data came from. This can enable the system to evaluate the trustworthiness of the information, for example information from Wikipedia may be more trustworthy than an opinion expressed in a tweet. If a user asks the robot *How do you know that?*, the robot should be able to justify its responses by at least giving the source of the information.

The ability to add layers of metadata to knowledge graphs is important for the development of explainable AI. As well as using data lineage to justify responses, the use of taxonomies also opens up new possibilities for the robot to explain more clearly *why* it gave a particular response by describing the relevant taxonomic relationships explicitly to the user.

C. Related work

In earlier work based on the Semantic Web and RDF, tools such as OpenRobot Ontology (ORO) [15] have been developed and extensively used over the last 10 years to provide semantic understanding to robots. In Neo4j databases, knowledge graphs are represented using a property graph model, but RDF triples can be imported into Neo4j property graphs (and exported) using the Neosemantics[10] RDF and semantics plugin.

In related work on explainable reasoning and knowledge graphs, OpenDialKG [16] advances the DialKG Walker model that learns the symbolic transitions of dialogue contexts as structured traversals over the knowledge graph. By taking all the entities mentioned in the dialogue as starting nodes, and using supervized learning, it learns the reasoning path over the graph via graph attention.

GraphDial[11] is a recent virtual coaching initiative that seeks to implement a robot agent powered by knowledge graphs to assist in human-robot dialogue.

V. A Furhat Robot with Rasa and Neo4j

This section presents a proof-of-concept demo using the CityTalk application. The Nao robot is replaced by a Furhat social robot [17]. Nao speech recognition and synthesis are replaced by Google Cloud speech recognition and Amazon Polly speech synthesis. PyDial is replaced by Rasa. The Furhat SDK is connected via Furhat Remote API to a Rasa dialogue server, which is connected to a Neo4j graph database.

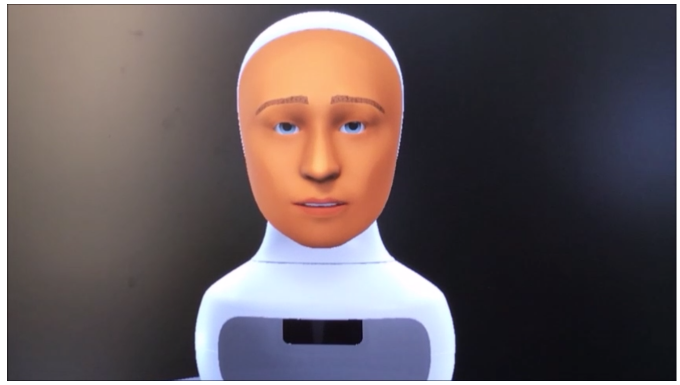

Fig. 2. Virtual Furhat with Rasa conversational AI and Neo4j graph database.

Figure 2 is a screenshot from a short video[12] in which Virtual Furhat engages with a user in spoken interaction managed by Rasa conversational AI and accesses the required information from knowledge graphs in a Neo4j database.

VI. Conclusion and future work

We have described Rasa open-source conversational AI and how it can be used with knowledge graphs in a Neo4j graph database. In particular we described how Rasa knowledgebase actions can be used, and how more advanced knowledge graphs can be developed by using a graph database. We also showed an example implementation in which a Furhat social robot uses Rasa conversational AI to access knowledge graphs in a Neo4j graph database.

As well as connecting to multiple back-end knowledge bases, Rasa conversational AI services can be accessed from multiple front-end channels (Facebook Messenger, Slack, Twilio, etc.). It would be interesting to develop “social robot channels” for Rasa conversational AI, if the social robotics and conversational AI communities would get together to support this kind of activity. For example, we could make a Furhat channel as the first social robot channel for Rasa.

Notes

[1] http://www.interspeech2020.org/index.php?m=content&c=index&a=show&catid=369&id=291.
[2] http://sap.ist.i.kyoto-u.ac.jp/ijcai2020/robotdial/.
[3] https://www.omg.org/.
[4] https://github.com/IBM/watson-nao-robot.
[5] https://neo4j.com/.
[6] https://www.youtube.com/watch?v=PBXLu5lfhGs.
[7] https://rasa.com/docs/rasa/.
[8] https://rasa.com/docs/action-server/knowledge-bases.
[9] https://github.com/PacktPublishing/Conversational-AI-with-RASA.
[10] https://neo4j.com/labs/neosemantics/.
[11] https://graphdial.nr.no.
[12] https://www.youtube.com/watch?v=lNYEakEaJhU.

References

[1] K. Jokinen, "Dialogue Models for Socially Intelligent Robots," in *Social Robotics. ICSR 2018*, S. S. Ge, J.-J. Cabibihan, M. A. Salichs, E. Broadbent, H. He, A. R. Wagner, and Á. Castro-González, Eds. Springer, 2018, pp. 127–138.

[2] T. Bocklisch, J. Faulkner, N. Pawlowski, and A. Nichol, "Rasa: Open source language understanding and dialogue management," 2017, arXiv:1712.05181.

[3] R. Chatterjee, "Why does RASA have an edge over other Chatbot Frameworks for implementing Conversational AI & Chatbot Solutions?" 2021, https://medium.com/co-learning-lounge/.

[4] I. Robinson, J. Webber, and E. Eifrem, *Graph Databases (2nd edition)*. O'Reilly Media, 2015.

[5] G. Wilcock, "CityTalk: Robots that talk to tourists and can switch domains during the dialogue," in *9th International Workshop on Spoken Dialogue Systems Technology*, L. F. D'Haro, R. E. Banchs, and H. Li, Eds. Springer, 2019, pp. 411–417.

[6] S. Ultes, L. M. Rojas Barahona, P.-H. Su, D. Vandyke, D. Kim, I. Casanueva, P. Budzianowski, N. Mrkšić, T.-H. Wen, M. Gasic, and S. Young, "PyDial: A multi-domain statistical dialogue system toolkit," in *Proceedings of ACL 2017, System Demonstrations*, Vancouver, Canada, July 2017, pp. 73–78.

[7] G. Wilcock, "Using a deep learning dialogue research toolkit in a multilingual multidomain practical application," in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*, Stockholm, Sweden, 2018, pp. 5880–5882.

[8] ——, "Recognising flexible intents and multiple domains in extended human-robot dialogues," in *Advances in Artificial Intelligence: Selected Papers from the Annual Conference of the Japanese Society for Artificial Intelligence (JSAI 2021)*, Y. Takama and et al., Eds. Springer, 2022.

[9] K. Jokinen, S. Nishimura, K. Watanabe, and T. Nishimura, "Human-Robot Dialogues for Explaining Activities," in *9th International Workshop on Spoken Dialogue Systems Technology*, L. F. D'Haro, R. E. Banchs, and H. Li, Eds. Springer, 2019, pp. 239–251.

[10] X. Kong and G. Wang, *Conversational AI with Rasa*. Packt Publishing, 2021.

[11] T. Bunk, D. Varshneya, V. Vlasov, and A. Nichol, "DIET: Lightweight language understanding for dialogue systems," 2020, arXiv:2004.09936.

[12] V. Vlasov, J. E. M. Mosig, and A. Nichol, "Dialogue transformers," 2019, arXiv:1910.00486.

[13] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. de Melo, C. Gutiérrez, S. Kirrane, J. E. Labra Gayo, R. Navigli, S. Neumaier, A.-C. Ngonga Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. F. Sequeda, S. Staab, and A. Zimmermann, *Knowledge Graphs*. Morgan & Claypool, 2021.

[14] J. Barrasa, A. E. Hodler, and J. Webber, *Knowledge Graphs: Data in Context for Responsive Businesses*. O'Reilly Media, 2021.

[15] S. Lemaignan, R. Ros, L. Mösenlechner, R. Alami, and M. Beetz, "ORO, a knowledge management platform for cognitive architectures in robotics," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 3548–3553.

[16] S. Moon, P. Shah, A. Kumar, and R. Subba, "Explainable conversational reasoning with attention-based walks over knowledge graphs," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 845–854.

[17] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström, "Furhat: A back-projected human-like robot head for multiparty human-machine interaction," in *Cognitive Behavioural Systems*, A. Esposito, A. M. Esposito, A. Vinciarelli, R. Hoffmann, and V. C. Müller, Eds. Springer, 2012, pp. 114–130.