Invited Talk: Natural Interactions and Knowledge in Context-Aware Virtual Coaching

Kristiina Jokinen
Artificial Intelligence Research Center (AIRC),
National Institute of Advanced Industrial Science and Technology
Japan

CCS CONCEPTS

• **Theory of computation → Semantics and reasoning**; **Logic**;
• **Computing methodologies → Artificial intelligence**.

KEYWORDS

Human Interaction, Dialogue Modelling, Graph Database

ACM Reference Format:
Kristiina Jokinen. 2021. Invited Talk: Natural Interactions and Knowledge in Context-Aware Virtual Coaching. In *KGR4XAI2021: The 1st International Workshop on Knowledge Graph Reasoning for Explainable Artificial Intelligence.* ACM, Online, 2 pages.

1 EXTENDED ABSTRACT

Rapid development in AI technologies has brought forward many interactive applications where natural interaction capability plays a significant role. Applications in such varied areas as care giving, healthcare, coaching, assisted living, etc. convey information, thoughts, and feelings, and an important requirement for them is to show awareness of the human partner, of the situation and the environment. For instance, we collaborate in a large EU-Japan project e-VITA [1], which focuses on these issues to support active healthy living.

Models for such context-aware robot interactions require situational awareness and understanding of human multimodal behaviour. Situational awareness refers to the knowledge of the partner and the world, understanding what is "going on." Consequently, ability to communicate smoothly in a given situation is important. Constructive Dialogue Modelling [2] defines dialogues with respect to communicative enablements which support building of the mutual context by grounding information, through the cycle of Contact, Perception, Understanding, and Reaction.

Much multimodal and multisensory data can be collected from the user and the environment, and it is important for the dialogue modelling to provide responses that are relevant and appropriate in the interaction context. Multimodal technologies allow studies on the understanding of human behaviour on the signal level: for instance, gaze patterns can be effectively used to predict turn taking [3] or estimate fluency in conversation [4], while hand and head gestures are important in the coordination of communication [5].

KGR4XAI2021, Dec. 6, 2021, Online

Our goal is to develop context-aware dialogue systems for interactive situations using AI techniques. The framework includes technologies for conversational AI and multimodal signal recognition, and an integrated architecture for a humanoid robot. In order to store and process multimodal dialogue data in a structured format, we use knowledge graphs.

As the basis for dialogue implementation, we use Rasa Conversational AI[1]. It is a popular open-source conversational AI framework which offers flexible use of state-of-the-art machine-learning techniques (transformer-based natural language and dialogue processing), and access to various knowledge sources. Our methodology consists of extending the dialogue model with a context model where the recognized entities are stored in the robot's memory and tagged according to the entity's status in the grounding process. Knowledge graphs are used to represent the current dialogue state and the relevant context, and the dialogue manager can query knowledgebases for the appropriate response based on the entities and their context, see Fig 1. Dialogue manager can also update knowledge graphs, which allows us to study how the robot's internal knowledge is constructed in interaction.

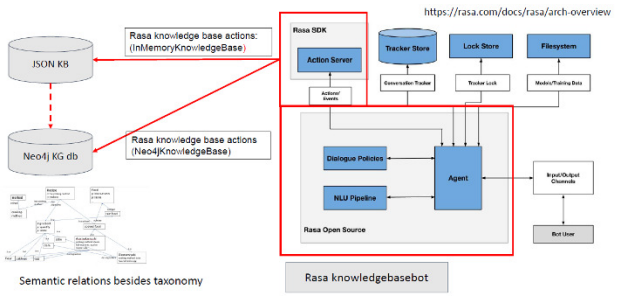

Figure 1: Dialogue modelling and knowledge bases with Rasa. Modified from https://rasa.com/docs/arch-overview

Knowledge graphs are a way to structure the domain knowledge. We use property graph models stored in a Neo4j graph database. Property graphs can provide a rich description of the domain, and they offer in an intuitive way to describe the important domain concepts and relations. Other types of databases, such as relational and document databases, can also be included in the system using custom actions. An interesting option is to produce knowledge graphs from textual sources like online books and articles by NLP techniques. However, this requires customization to the specific application domain to extract the information. Wikidata provides a good ready-made knowledge source with annotated semantic tags,

[1]https://rasa.com/docs/rasa/

although particular use cases may need to be added in order to match the task-specific requirements.

Finally, natural dialogues allow rich questions, long dialogues, references to previous answers, chatting, as well as personalized dialogues for different users with different needs in various situations and cultures. For this, we use user profiling: knowledge of the user preferences will guide the system to focus more on the particular type of information that is likely to be in the user's interest area and the user experience can be increased by the system remembering the previous interaction with the user. However, care should be taken in two aspects. First, user preferences may change, and suggestions based on the stored preferences may irritate the user if they do not follow her current preferences. The system should dynamically learn user preferences and modify its understanding of the user characteristics accordingly. Secondly, in long-term interactions, ethical and privacy issues are important [6]. The user must be aware of what is stored in the system memory, and the interactive system should ask permission to save information to support trust in the system truthful operation and reliability.

REFERENCES

[1] Jokinen, K., Homma, K., Matsumoto, Y., Fukuda, K. "Integration and Interaction of Trustworthy AI in a Virtual Coach". In: Matsushita, Takama (Eds.): Advances in Artificial Intelligence. Springer (2021).
e-VITA COACH: https://www.e-vita.coach/
[2] Jokinen, K. Constructive dialogue management – Speech interaction and rational agents. John Wiley (2009).
[3] Jokinen, K., Furukawa, H., Nishida, M., Yamamoto, S. Gaze and Turn-taking behaviour in Casual Conversational Interactions. ACM Trans on Interactive Intelligent Systems, Vol 3, Issue 2. (2013).
[4] Ijuin, K., Jokinen, K. Utterances in Social Robot Interactions — Correlation Analyses between Robot's Fluency and Participant's Impression. Proceedings of the 7th International ACM Conference on Human-Agent Interaction (HAI'19) (2019).
[5] Mori, T., Jokinen, K., Den, Y. On the Use of Gestures In Dialogue Breakdown Detection. Oriental COCOSDA, Singapore (2021).
[6] Jokinen, K., Wilcock, G. Do you remember me? Ethics in Social Robot Interactions. The 30th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2021) (2021).