Long- and short-term common ground for tailored explanations in Voice User Interfaces

Jorge Del-Bosque-Trevino, Julian Hough, Matthew Purver
{j.delbosque;j.hough;m.purver}@qmul.ac.uk
Queen Mary University of London
London, UK

ABSTRACT

While Voice User Interfaces (VUI) are becoming increasingly embedded into everyday life, their ability to tailor their output to individual users is limited. Research in VUIs has explored the use of static user models to encode general preferences; and, separately, dynamic models of dialogue context or short-term common ground have been used to inform natural language generation decisions. Neither of these alone is enough to provide a VUI with the ability to dynamically explain concepts. This paper highlights the need to use both, and thus develop new interactive models of tailored explanations.

INTRODUCTION: TAILORED EXPLANATIONS

Voice User Interfaces (VUIs) such as Alexa, Cortana, Google Assistant and Siri are good at narrow tasks like responding to direct questions like "What is the cheapest flight from London to Mexico?". However, if three individuals were to ask the question "What is the International Space Station?" to all four devices, they would get the same answer. If, in contrast, this question was posed to a human, the response given would be an explanation which took the question-asker into account, in terms

of both their background and likely initial understanding, and their ongoing interactive behaviour, including displays of understanding and non-understanding. This position paper outlines the need to develop new models of tailored explanations within dialogue that should incorporate both the use of long-term common ground with users in the sense of [1] and short-term grounding process and dialogue context in the sense of [2, 5].

HOW DO HUMANS EXPLAIN?

Humans generate explanations so that the hearer can understand, both by using information about the hearer's knowledge, and also by dynamically adapting and building their explanations on the fly in real time [7]. To do this, explainers commonly exploit mechanisms of analogy and metaphor [6]– that is, the use of one conceptual structure to organise the understanding of an entirely different concept [13] [4]. Metaphor can give rise to conceptual blending [3], a cognitive operation in which conceptual ingredients are combined to create a new structure with its own emerging meanings [13]. However, these powerful mechanisms must be rooted in the user's background and understanding in order to be meaningful.

ACHIEVING MEANINGFULNESS THROUGH THE USE OF COMMON GROUND

The concept of common ground as presented by Stalnaker [12] builds on notions of common knowledge [9], mutual knowledge or belief [11] and joint knowledge [10], and is described more recently by Clark [1] as two peoples' sum of their mutual, common, or joint knowledge, beliefs and suppositions. Long-term common ground which can be drawn on includes nationality, residence, education, occupation, employment, hobby, language, religion, politics, ethnicity, subculture, cohort and gender [1]. Speakers use this knowledge to design the production of their utterances to fit the audience [8]. A VUI cannot acquire or develop common ground in the same way humans do, but it could acquire functionally equivalent information useful for producing effective explanations. Allowing a functional common ground to be built up over time by a system and a user over repeated interactions, along with mechanisms for updating this long-term common ground through short-term dialogue grounding mechanisms [2], would allow VUIs to generate explanations tailored for individuals which can also be adapted on the fly.

CONCLUSION

We propose building models of tailored explanations that can not only adapt to a user based on common ground built up in a given interaction, but also draw on the long-term common ground built up by the system with that user over repeated interactions. Using this common ground would allow effective metaphor creation in explanations, allowing VUIs to achieve more meaningful and educational user experiences.

REFERENCES

[1] Herbert H. Clark. 1996. Common Ground. In *Using Language*. Vol. 9. Cambridge University Press, New York, Chapter 4. Common, 92–121.

[2] Herbert H Clark, Susan E Brennan, et al. 1991. Grounding in communication. *Perspectives on socially shared cognition* 13, 1991 (1991), 127–149.

[3] Gilles Fauconnier and Mark Turner. 1996. Blending as a central process of grammar. In *Conceptual Structure, Discourse, and Language*, Adele Goldberg (Ed.). Cambridge University Press, Cambridge, Chapter Blending a. https://www.researchgate.net/publication/255999754

[4] Dedre Gentner. 1983. Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science* 7 (1983), 155–170. https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog0702{_}3

[5] Jonathan Ginzburg. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press.

[6] John E. Hummel, John Licato, and Selmer Bringsjord. 2014. Analogy, explanation, and proof. *Frontiers in Human Neuroscience* 8 (2014), 867. https://doi.org/10.3389/fnhum.2014.00867

[7] Frank C. Keil. 2006. Explanation and Understanding. *Annual Review of Psychology* 57, 1 (2006), 227–254. https://doi.org/10.1146/annurev.psych.57.102904.190100 arXiv:https://doi.org/10.1146/annurev.psych.57.102904.190100 PMID: 16318595.

[8] Boaz Keysar and William S Horton. 1996. When do speakers take into account common ground? *Cognition* 59, 1 (1996), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1

[9] David Lewis. 1969. Convention: a philosophical study.

[10] Jhon McCarthy. 1990. Formalization of two puzzles involving knowledge. *Formalizing Common Sense: Papers by John McCarthy* (1990), 158–166.

[11] Stephen R. Schiffer. 1972. Meaning. (1972).

[12] R Stalnaker. 1978. Assertion. In (P. Cole, ed.) Syntax and Semantics. *Pragmatics* 9 (1978).

[13] Tony Veale. 2012. Pimp my Ride: Customizing language with metaphor, analogy and blending. In *Exploding The Creativity Myth: The Computational Foundations of Linguistic Creativity*. Bloomsbury Publishing, London, Chapter Pimp my Ri, 87–108.