

## Impulse: How do we arrive at a decent spoken dialogue system?

ALJOSCHA BURCHARDT, JAN NEHRING, DFKI Speech and Language Technology Lab, Germany

Additional Key Words and Phrases: Spoken dialogue systems, interdisciplinarity

### 1 WHAT DO WE KNOW ALREADY?

Several of today's applications of speech-based technologies just use a different modality, but are not dramatically different from their written counterpart. If we take voice search, e.g., there are obviously some advantages such that it can be hands-free, is possible without a display, e.g., in extreme lightning situations etc. And there are disadvantages as it is, e.g., not possible to utilize lightweight input support via type-ahead and it is especially challenging to present the search results in spoken language. But the search process itself is not fundamentally different. But when it comes to the communicative aspect of language, speech-based systems can make a difference and what comes to mind immediately are spoken dialogue systems.

In a very strong simplification, AI (Artificial Intelligence) systems are either *modeled* in some symbolic framework or *learnt from data* using some form of machine learning (ML).<sup>1</sup> In terms of performance, ML has recently often outperformed the symbolic approaches while suffering from lack of control and transparency. In the realm of dialogue processing, we typically find a mix of technologies. The periphery (speech input and output) is usually data-driven as is the recognition of user intents. The dialogue itself is usually modeled using one of several existing frameworks that are also the basis of our personal assistants. That is why the dialogues are typically centered around simple tasks such as accessing an FAQ list or switching the light on where it is comparably easy to measure task completion, i.e., success of the dialogue. Some systems that have learnt simple dialogues like asking for opening hours do exist, but we are very far from the open domain dialogue capacities of a four year old child.

On the other hand chatbots have more to offer than fulfilling a certain task in a short time. We performed user studies that suggest that chatbots can generate a more playful and interesting user experience than traditional interfaces. When the chatbot designer also targets other goals beyond task completion, such as user satisfaction, fun or curiosity, we believe that chatbots can generate a fun and rich experience.

### 2 HOW DO WE STUDY THE PHENOMENON?

We predominantly built chatbots for industry applications. In these projects the goal of creating an interesting character for the chatbot was often subordinate to other project goals. Nevertheless, in the user studies we performed, even these rather serious chatbots have been perceived as more interesting, innovative and generally attractive than traditional user interfaces by the participants. Methodologically, we performed classic user experience evaluation using questionnaires such as AttrakDiff [1] and the User Experience Questionnaire [2], sometimes accompanied by structured interviews.

In a cooperation with the University of Arts / Berlin (UdK) we created chatbots for the mere purpose of entertainment and although we could not perform a formal user evaluation, the feedback to these chatbots was very positive.

---

<sup>1</sup>The author is aware that this is not a scientifically clean partition.

---

Author's address: Aljoscha Burchardt, Jan Nehring, Aljoscha.Burchardt@dfki.de, DFKI Speech and Language Technology Lab, Alt-Moabit 91c, Berlin, Germany, 10559.

### 3 WHAT WE WOULD LIKE TO KNOW?

Human dialogue or – more generally – conversation has been studied in a huge variety of research areas including philosophy, cognitive psychology, (psycho)linguistics, sociology etc. Our question would be how we can inform and drive the design process of future dialogue systems starting from the insights and frameworks used in these fields of study: How can we model a (good) dialogue? What brings a dialogue further? How can we make sure that the partners understand each other (grounding)? What is the final goal of a dialogue? How can we evaluate and measure the success of a dialogue?

How can chatbots generate a rich and entertaining user experience? Of which dimensions does this user experience consist? And how do we measure these dimensions? If we had such an evaluation framework, we could evaluate different chatbot implementations to find out which one provides the best user experience beyond task success.

The most central question, however, would be how we can translate all these insights from different areas into either formal models or data in order to incorporate them in AI systems as sketched above. This will require a high degree of interdisciplinary cooperation.

### 4 WHAT DO WE WANT TO LEARN FROM DIFFERENT DISCIPLINES?

We think the different disciplines would first and foremost need to find a common language. Not only on the meta level talking about the phenomenon and their different perspectives and research questions, but also when describing the phenomenon of study itself, e.g., when describing what kind of knowledge a system (dialogue partner) would need about the world, the current situation or the (human) dialogue partner.

### 5 WHAT DO WE WANT TO TEACH OTHER DISCIPLINES?

We would try to make people see the world through "digital lenses": Everything a weak AI system can do has to be taught to it before in some way. For a translation system, this is comparably easy, we feed it with (human translated) texts. For a summarization system, it is already much more complicated. We need to feed it with texts and their summaries. There is much more variety (and conversely, sparseness) in this data, but we can imagine how it would work. What data and knowledge structures do we need to do for a dialogue system that goes beyond the current state of the art?

### REFERENCES

- [1] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003: Interaktion in Bewegung*, G. Szwillus and J. Ziegler (Eds.). B. G. Teubner, Stuttgart, 187–196.
- [2] Bettina Laugwitz, Martin Schrepp, and Theo Held. 2006. Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In *Mensch und Computer 2006: Mensch und Computer im Strukturwandel*, Andreas M. Heinecke and Hansjürgen Paul (Eds.). Oldenbourg Verlag, München, 125–134.