Online Journal of Applied Knowledge Management

A Publication of the International Institute for Applied Knowledge Management

Volume 6, Issue 1, 2018

Some design aspects of a cognitive user interface

Markus Huber, Brandenburg Technical University Cottbus-Senftenberg and InnoTec21 GmbH, Germany, markus.huber@b-tu.de

Matthias Wolf, Brandenburg Technical University Cottbus-Senftenberg, Germany, matthias.wolff@b-tu.de

Werner Meyer, Brandenburg Technical University Cottbus-Senftenberg, Germany, werner.meyer@b-tu.de

Oliver Jokisch, Leipzig University of Telecommunications (HfTL), Germany, jokisch@hft-leipzig.de

Kati Nowack, Brandenburg Technical University Cottbus-Senftenberg, Germany, kati.nowack@b-tu.de

Abstract

We report on research conducted as part of the Universal Cognitive User Interface (UCUI) project, which aims at developing a universal, autarkic module for intuitive interaction with technical devices. First, we present an empirical study of image schemas as basic building blocks of human knowledge. Image schemas have been studied extensively in cognitive linguistics, but insufficiently in the context of human-computer-interaction (HCI). Some image schemas are developed early at pre-verbal stages (e.g., up-down) and may, thus, exert greater influence on human knowledge than later developed image schemas (e.g., centre-periphery). To investigate this for HCI contexts, we applied a speech interaction task using a Wizard of Oz paradigm. Our results show that users apply early image schemas more frequently than late image schemas. They should, therefore, be given preference in interface designs. In the second part of this contribution we therefore focus on the appropriate representation and processing of semantics. We introduce novel theoretical work including feature-values-relations and Petri net transducers, and discuss their impact on behaviour control of cognitive systems. In addition, we illustrate some details of the implementation regarding learning strategies and the graphical user interface.

Keywords: Cognitive system, intuitive interaction, image schema theory, feature-values-relation, Petri net transducer, and behaviour control.

Introduction

Our collaborative project *Universal Cognitive User Interface* (UCUI) aims at developing a universal hardware module that enables intuitive interaction with technical devices whilst safeguarding privacy. User interfaces and modalities have been already surveyed within the knowledge management (KM) context, in particular from the perspective of mobile and multimodal applications (Karolić, 2013; Silber-Varod & Geri, 2014). We extend previous investigations by adapting our interface design to allow for successful speech interaction. Taking a design-for-all approach, speech interaction is of particular relevance for both technically inexperienced users as well as visually impaired people.

At first, we report on an experimental investigation of image schemas as basic aspects of human knowledge that could help to support the development of intuitive interface designs. Such interface designs should enable effective interaction with a system based on the subconscious application of basic prior knowledge (Mohs, Hurtienne, Kindsmüller, Israel, Meyer, et al., 2006; Turner, 2008). Second, we report on the technical side dealing with the representation and processing of semantics. We briefly introduce feature-values-relations and Petri net transducers, and describe our approach of behaviour control for cognitive systems. Third, we present some results from our implementation of the mentioned hardware module.

The transition from technology-centred devices toward human-centred devices requires the implementation of basic models for a variety of human experiences. One such fundamental aspect of human experience, and the focus of this investigation, is the impact of image schemas on human knowledge and language. Image schemas (e.g., up-down, centre-periphery) are basic pre-conceptual, universal patterns of human experience that integrate information from multiple perceptual modalities such as visual, acoustic and haptic information (Lakoff & Johnson, 2011). They serve to structure human knowledge, behaviour and experience. As basic building blocks of human knowledge generated in earliest childhood interactions, image schemas should be available to all potential users. Whilst research seems to support this (Hurtienne et al., 2010), studies on concept formation suggest that some image schemas occur earlier in infancy compared to other image schemas (Mandler & Pagán Cánovas, 2014). According to Mandler and Pagán Cánovas (2014), path schemas such as up-down, container, location, blockage, into, out of, open are basic image schemas occurring early in infancy at pre-verbal stages. On the contrary, centre-periphery, scale, balance, cycle and other process schemas, near-far, multiplicity-unity as well as attributional image schemas (e.g., big-small, warm-cold) are built upon these basic primitives and, thus, should exert less influence on the development of human thought and knowledge (Mandler & Pagán Cánovas, 2014).

To investigate whether developmental occurrence of image schemas influences their appliance in speech interaction with technical devices, we applied a Wizard of Oz paradigm using a special tailored software framework (Huber, Meyer, Nowack, & Geßler, 2016; Huber & Jokisch, 2017).

Method

Participants

Forty-three adults (20 men, 23 women; mean age 29.2 years, SD 9.9 years) participated in the speech interaction study. They received a payment of 25 Euros. All were German native speakers. The study was conducted in accordance with ethical principles stated in the Declaration of Helsinki (1964). Informed consent was obtained for experimentation with human participants.

Materials and Procedures

To assess technical experience (TE), participants were asked to indicate the frequency with which they used a variety of technical devices at home (e.g., smartphone, laptop) and in public (e.g., ticket machine, self-service banking). For seventeen items, participants had to select one of the following response options: 0 – I don't know this device, 1 – never, 2 – rarely, 3 – once per month, 4 – once a week, 5 – almost daily, 6 – once per day and 7 – more than once per day. The mean of all responses was then calculated for each participant with a higher score indicating greater TE. The range of TE in our population sample was 2.35 – 4.24 (mean TE score: 3.34). Both age and gender were not correlated with TE scores (all p's > 0.05).

The speech interaction task took place in a quiet and moderately illuminated room. All participants were asked to complete the questionnaire ascertaining their technical experience as well as demographic data. Participants were then asked to stand in front of a multi-touch panel (NEC MultiSync® P701, size of the display: 70 inches, display resolution: 1920 x 1080). To encourage natural speech interaction with a technical device, a Wizard of Oz design was applied: participants thought they were interacting with an autonomous technical device that was actually operated by an unseen experimenter. They were presented with twenty test scenarios investigating image schemas underlying free speech interaction with a heating device ((Figure 1). They were not informed about image schemas in any way. Participants were simply informed that they would be presented with various scenarios that could also occur in their everyday life when interacting with technical devices. They were told that there were no right or wrong answers, since they were to test the functionality of a newly developed heating device. Participants were asked to respond to the various scenarios, as they would also do at home and to simply tell the heating device which changes, if any, they would like the device to carry out. Upon participants' speech responses, the next scenario was presented. All stimuli were presented on a black background. Responses were collected on the same computer used for stimulus presentation. The experiment was run in German; all examples have been translated into English.

Figure 1. Examples of individually presented test scenarios for the speech interaction task

Design, Data Selection, Cleaning, and Reduction

Age, gender, and technical experience were considered as continuous independent between-subjects variables. As within group factor, developmental occurrence of image schemas (basic versus later) was manipulated as independent variable. For every participant, speech responses were recorded as dependent variables. Items were presented randomly.

A content analysis was carried out to identify image schemas underlying utterances (see Table 1). Image schemas underlying speech and gesture responses were analysed by two independent coders. To ensure reliability of coding, one coder coded the entire speech and gesture dataset. A second coder that was blind to the study hypotheses coded 25% of both datasets. In the speech interaction study, the two coders agreed 88% of the time (Cohen's kappa = 0.75, $p < 0.001$). Disagreements were resolved by discussion. Frequencies were analysed using the IBM SPSS Statistics Version 24.

Table 1. Coding of Speech Utterances into Image Schemas

Image Schema	**Example speech utterance**
Basic	
Path: up-down	(to move/ drive) up, down, to rise, to lower, higher, lower
Path: verticality	to set (German: *stelle ein; stelle auf*)
Path: horizontality	From...to, along, back/ front, backwards, forward
Container	to put in, to hold, in the room, in the house, in daytime
Later	
Balance	to regulate, to adjust
Attribute: warm-cold	make it warmer/ colder, to heat (up/ down), to temper
Process (cycle, iteration)	to turn (up/ down), to deactivate
Near-far	Near xx degrees, next week, next day
Multiplicity-unity	whole day, whole week, a few (degrees)
Scale	to reduce, to cut, to increase/ decrease, how much

Pearson's correlational analyses were conducted to examine associations between age, gender, technical experience and frequencies of image schemas. Repeated-measures analyses of variance (ANOVA) and analyses of covariance (ANCOVA) were conducted to compare frequencies of early and late image schemas after testing for normal distribution (Shapiro-Wilk test). All effects are reported as significant at $p < .05$.

Results and Intermediate Discussion

On average, participants applied 1.8 (SD = 0.3) image schemas in each test scenario. In total, 10 different image schemas (up-down, verticality, horizontality, balance, scale, warm-cold, process, near-far, multiplicity-unit and container) were identified in all scenarios. Each participant used on average 6.7 different image schemas (SD = 1.00) in speech interaction. The four-image schemas up-down, verticality, horizontality and container were regarded as early image schemas. Whilst the six-image schemas balance, scale, warm-cold, process (including cycle), near-far and multiplicity-unity were regarded as later image schemas (Mandler & Pagán Cánovas, 2014). Figure 2 depicts mean frequencies in percent for early and late image schemas. A repeated-measures (developmental occurrence: basic versus later) ANOVA was conducted for frequencies after testing for normal distribution (Kolmogorov-Smirnov test). In line with the first hypothesis, there was a statistically significant effect of developmental occurrence on frequencies ($F(1,42)$ =

67.494, $p < 0.001$, $\eta^2 = 0.616$). There were no significant effects for gender, age and technical experience (all p's > 0.05) on frequencies. Thus, the frequency with which specific image schemas were linked to speech utterances varied with developmental occurrence of image schemas but not with individual differences in gender, age, and technical experience. In general, basic image schemas were employed significantly more often (M = 67.81 %, SD = 14.22) than image schemas occurring later in development (M = 32.19 %, SD = 14.22).

In line with Mandler and Pagán Cánovas (2014), basic image schemas exert not only greater influence on the development of human thought, the results indicate that they are also more frequently applied in speech interaction with technical devices.

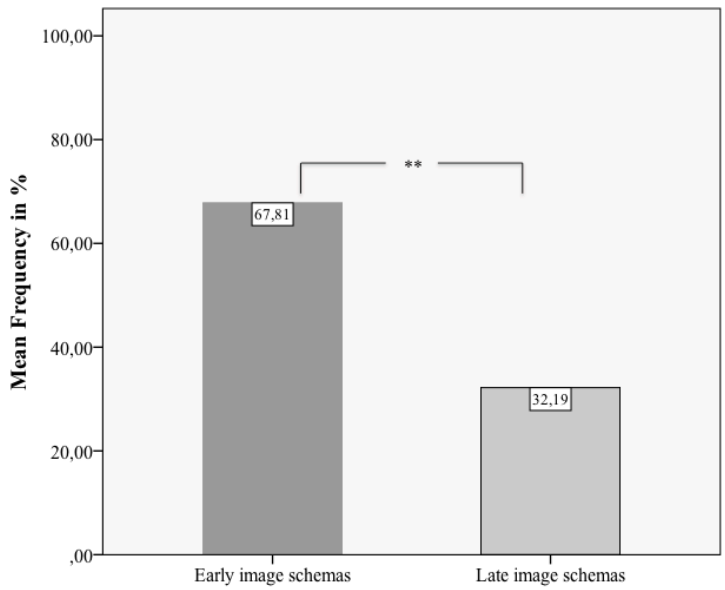

Figure 2. Mean Frequency (in %) for identified image schemas underlying speech interaction with a heating device

In accordance with the second hypothesis, gender, age, and technical experience were not correlated with both the frequency of early and late image schemas. This suggests that image schemas as basic building blocks of human knowledge are available to users of all ages and technical experience independently from gender.

Intuitively usable technical devices should enable users to effectively and easily interact with a system without the need for long controlled cognitive processes (Mohs et al., 2006; Turner, 2008). Interaction should, thus, be based on automatically retrievable prior knowledge available to all potential users such as knowledge based on image schemas. Our investigation, however, extents previous research by demonstrating that not all image schemas are equally intuitive in Human-Computer-Interaction (HCI): The developmental occurrence of image schemas impacts upon the frequency of applied image schemas when interacting with technical devices. Early image schemas should, thus, be given preference over late image schemas in interface designs.

Technical Aspects of Semantic Processing

Cognitive systems (including speech dialogue systems) mainly consist of three parts allowing them to interact with their environments: perceptor, behaviour control, and actuator (Haykin, 2012). Of these, perceptor and actuator each comprise a hierarchy of transforming units and bidirectional communication between the particular levels (Römer & Wirsching, 2013).

The responsibility of the perceptor hierarchy is translating input signals into semantic representations where the relevant parts from the input are related to semantic categories relevant to the system. Semantic representations abstract from concrete syntax and must, therefore, be something substantial different. Since image schemas contribute to meanings of interaction turns, they need to be included in the representation as well. Our approach of assigning suitable values to distinct features already is capable of integrating different sources and aspects of meaning.

Feature-Values-Relations

We use the concept of *feature-values-relations* (FVRs) as semantic representations. These were first defined by Huber, Kölbl, Lorenz, and Wirsching (2008) and further refined by Wirsching, Huber, Kölbl, Lorenz, and Römer (2012). In principle, these are hierarchical structures consisting of semantic categories and assigned values as in Figure 3. Whereas the structure is not restricted to tree-like forms, i.e. any node can have more than one parent, there can be more than one root node and even so-called N-forms (Huber & Römer, 2018). As was argued by Wirsching and Wolff (2014), Huber, Römer, and Wolff (2017) and Huber and Wolff (2017) semantics in general non-sequential but repeated parts need to be in some sequence. This can best be seen on the example in **Error! Reference source not found.** where the semantic structure for the name Wolfgang Amadeus Mozart is depicted. The order relation between Wolfgang and Amadeus is retained by the order of their semantic categories whereas there is no order relation between given names and surname. On the semantic level that ordering is of no relevance. To represent uncertainty, nodes of FVRs can also be equipped with weights. Huber and Wolff (2017) give a brief overview of related work also using non-sequential semantic representations. Wirsching and Wolff (2014) and Huber and Wolff (2017) showed that in general these structures can be arbitrarily nested, so there is no way of knowing in advance how much repetition, e.g. how many given names, there will be.

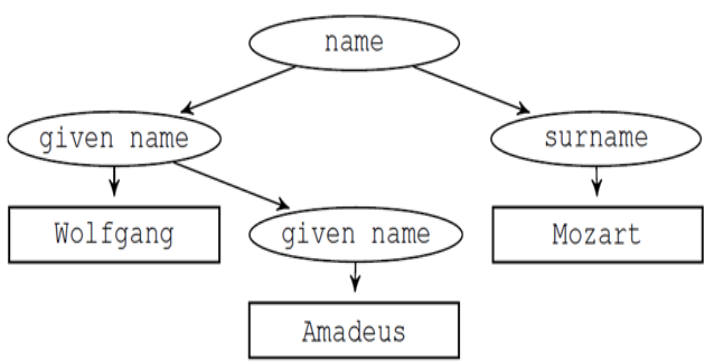

Figure 3. Semantic representation of "Wolfgang Amadeus Mozart" (cf. Huber & Wolff, 2017)

Huber and Römer (2018) described an approach allowing cognitive systems to reveal semantic categories in data consisting of features and assigned values. They follow an approach from Römer, Huber, and Wirsching (2016) as well as Schmitt, Römer, Wirsching, and Wolff (2017) where such data is represented in Hilbert spaces – vector spaces with some additional properties. Using techniques from quantum theory several of these vector spaces can be combined (cf. Widdows, 2004, for an application to *Information Retrieval*). Now searching for mappings inside the data can be interpreted geometrically as looking for orthogonal subspaces. Huber and Römer (2018) concluded that this results in structured Hilbert spaces of different dimensions corresponding to semantic categories. Thus semantic structures can be learned by cognitive systems. Note that the very basic example studied by Huber and Römer (2018) includes an N-form in the resulting structure. A first step towards not only representing the values as vectors but map whole structures into a vector space is done by beim Graben, Huber, Römer, and Wolff (2018). They use the Fock space from quantum field theory where different sectors of the space correspond to semantic categories. The mentioned approach also allows focusing on different situations. This means that solving problems need not result in processing a global state in a large vector space but keeping it as local as possible and therefore reducing complexity.

A more complex example of an FVR is depicted in Figure 4 that originates from our example application. It represents the semantics of an utterance consisting of two statements. A node labelled STEP with all necessary data below of it represents each statement. The ordering of the statements is given by the directed edge between the two-cascaded STEP-nodes.

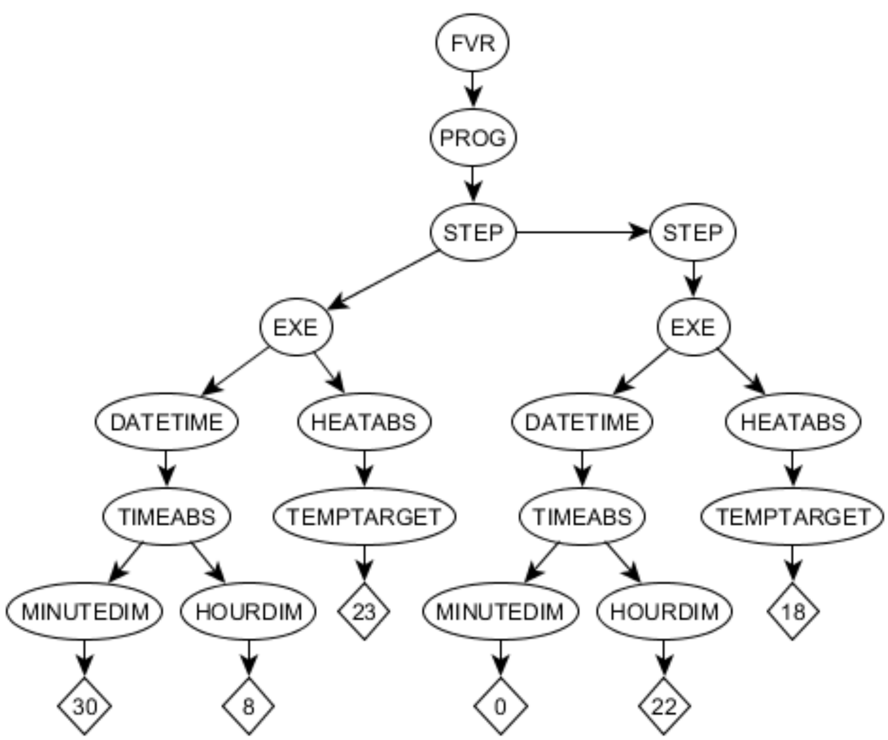

Figure 4. Cascaded semantic representation of "Heat 8:30 am 23 degree and 10 pm 18 degree!"

This way the formalism of FVRs is capable of representing arbitrarily often iterated semantic concepts while retaining their relative order. Although this order is of no relevance for the example, consider a route guidance system where you request a trip from Cottbus to Pisa via Munich and Trento. Would you like to visit Trento first and go on to Munich before finally arriving at Pisa?

Partial Orders and Petri Net Transducers

To represent FVRs mathematical objects called *labelled partial orders* (LPOs) can be used. These are well-studied generalisations of strings (Lorenz, Huber, & Wirsching, 2014). There exist several classes of underlying partial orders where the most general ones are those which include N-forms (Gischer, 1988). Operations for LPOs are described for example by Pratt (1986) or Gischer (1988). Mathematically speaking, an LPO corresponding to an FVR is its transitive closure. Some of the operations on LPOs correspond directly to operations on FVRs described by Geßler (2017). Lorenz et al. (2014) also defined *weighted LPOs* (wLPOs) where every node carries an additional weight from an algebraic structure. All operations on LPOs were lifted to wLPOs and an overall weight for a wLPO was defined.

Finally, *Petri net transducers* (PNTs) are processing machines for (w)LPOs. PNTs were first formally defined by Lorenz et al. (2014) as generalisation of finite automata. Compositions of PNTs and finite automata are able to translate between sequences and non-sequential structures (Huber & Römer, 2015). In connection with *subsymbol-symbol transducers* (SSTs) as defined by Wolff, Tschöpe, Römer, and Wirsching (2013) and finite automata, PNTs allow for a seamless, bidirectional translation between signal and semantics and complete the perceptor and actuator hierarchies.

PNTs where extended by Huber, Römer, and Wolff (2017) to handle recursive structures. This extension allows the processing of arbitrarily many iterations of semantic concepts where the order of repeated parts in a sequence is retained to the non-sequential structure. Huber and Wolff (2017) further extended PNTs to propagate segments from output to input. This way semantic concepts can induce segments on the input signal that makes aligning input signals to semantic concepts possible. Furthermore, weights get propagated from input signals to the output on the semantic level.

The semantic level of our example system is fully realised with FVRs. To create FVRs from speech input we use *utterance-meaning-transducers* (UMTs) (Lindemann, 2016), which can emulate PNTs for tree-like FVRs of known maximal width. This is possible in our example domain since it is very limited. A UMT is a special kind of an automaton. In Figure 5, an excerpt of an example UMT is represented. An automaton has system states depicted as circles connected by transitions depicted as directed edges. The labels above a transition include an input symbol, an output symbol, a weight and optionally a stack token. A colon in the named order separates all parameters. Input symbols represent the words needed to follow the corresponding transition. The output symbols of taken transitions are collected by the system. An epsilon symbol means that there is no input needed or no output given. In the figure the weight is only shown when different from zero or a stack token is present.

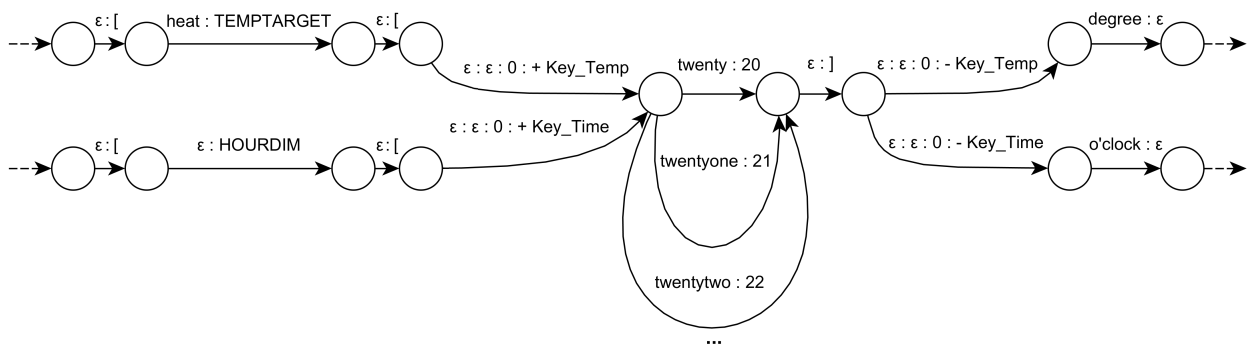

Figure 5. Excerpt of an Utterance-Meaning-Transducer

The depicted example from Figure 5 allows three different utterances regarding temperature or time (note that in German the days have 24 hours instead of two times 12 hours). For the input "…heat twenty-one degree…" the output "…[TEMPTARGET[21]…" would be generated and for the input "…twenty-two o'clock…" the output "…[HOURDIM[22]…". Note that an utterance like "…heat twenty-two o'clock…" is not possible since the stack tokens prohibit such a sequence of transitions. A stack token beginning with a plus sign restricts possible upcoming transitions to those with stack tokens beginning with a minus sign. This technique allows for more compact automata where common parts are defined only once and are reused, like the numbers in our example.

The concatenation of all output symbols of a complete path results in a string encoding a FVR where the squared brackets define the order of features and values. Inverse UMTs, where input and output symbols are interchanged, are used to generate speech output. A pre-processing step creates the set of all possible interleavings of paths from the FVR tree that is then used for synthesizing the systems answer.

Instinctive and Adaptive Behaviour Control

In the UCUI project we create a cognitive system that can handle different modalities of user input (speech, gesture, touch) to manage technical devices (Duckhorn et al., 2017). Meyer, Huber, and Wolff (2018) defined the behaviour control for the system in terms of FVRs. All communication from input components to behaviour control and from there to output components exclusively uses FVRs and is thus modality independent. With the operation of unification of FVRs, we can combine input from different sources and therefore base our decisions on all available information in parallel (Geßler, 2017). From behaviour control's point of view there is no difference in combining FVRs originating from different modalities or different interpretations, i.e. a FVR encoding a touch event and a FVR encoding an image schema is equal in that sense.

To represent possible actions of the system we use a world model as a collection of schemata. Each schema is a FVR containing some placeholders representing parameters of an action. For our example application the world model contains 14 schemata. Missing values are resolved by queries until the FVR is complete and the action can be executed.

For the queries there are two different paradigms. First, there is an instinctive behaviour control that compares user input to the world model and chooses a matching schema. Then it queries missing information from the user. Second, there is the adaptive behaviour control that learns from dialogues. It stores information in a dynamically constructed state space that can also be queried. Thus, it can adapt to repeatedly occurring situations using former dialogues to resolve missing values.

Our adaptive behaviour controller is based on a Markov decision process (Ertel, 2009). The state space is constructed by the instinctive behaviour control during interaction with the user. Weights assigned to the states represent their frequency of occurrence.

Consider the user said *"Heat!"*, then the system has no clue to which temperature. It asks the user and stores the answer. Later, in the same situation, e.g. *"Set the temperature higher!"*, the system may decide to use the stored value to complete the task. To meet the restrictions of limited space, we implemented a rule-based state dropping mechanism (Meyer, Huber, & Wolff, 2018). Computation of missing information is based on the work of Geßler (2017).

Using FVRs and Image Schemas for Touch Interaction

As outlined in the previous section, our approach of behaviour control is modality independent. Therefore, a user interface through a touch screen should appear to the system like any other input or output. This implies that the behaviour control should receive FVRs representing the meaning of touch events. Inversely, the screen must be able to display FVRs sent by the behaviour control.

There is a difference that makes screen-based user interfaces special. The user first chooses a task and then provides the needed data. This means the schema is selected by the user prior to any other information and there is no need for the system to deduce the schema from the given (possibly incomplete) input. Often it is even possible to provide a completely filled-in schema with a single click or touch. A major consequence is that screen-based user interfaces need access to the world model since they need to know a list of available schemata. Moreover, what is sent by the behaviour control should always include the selected/deduced schema. Expanding this method to any input and output components, allows them to be in sync enabling switching of modality during interaction with the system. On the technical side the difference between generating speech output for a query and displaying graphical elements for the very same query was almost annihilated since the concepts and processing machines are the same.

For the UCUI device we implemented a minimalist graphical user interface applying some image schemas. Besides a few buttons it mainly consists of a list (a container) and a slider (a scale in up-down direction). The list contains either the available schemata or the needed properties for an already selected schema. The slider is used to manipulate the values of the displayed properties, depicted on the right of Figure 6. The left side shows the display for a prominent one-property schema that is available by one of the four shortcut-buttons. The original interface on the device has a black background.

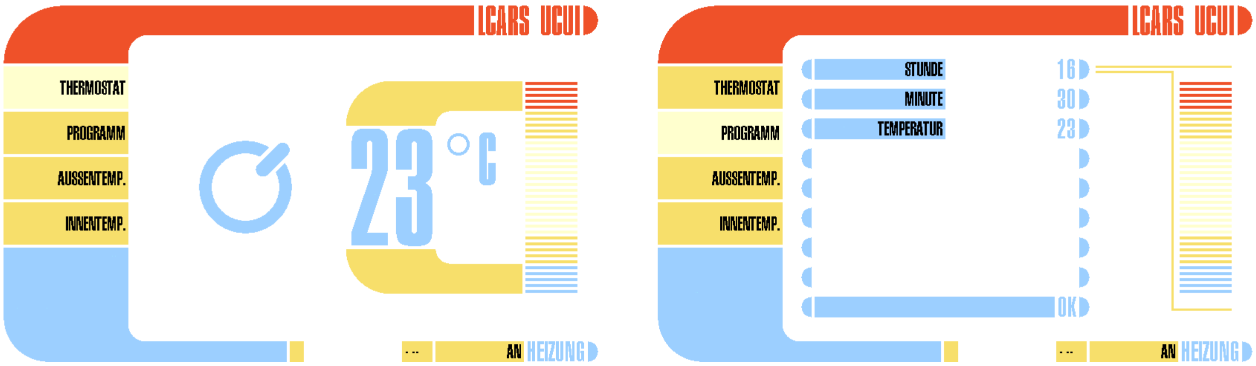

Figure 6. Screenshots from the UCUI device with German interface language

Conclusion

Our results show an impact of developmental occurrence of image schemas on speech interaction with technical devices. Our findings suggest that knowledge based on image schemas developed early at pre-verbal stages should be given preference in interface design. However, this was a first investigation of the influence of image schemas on human interaction behaviour depending on developmental occurrence. Further research into touch and gesture interaction behaviour is mandatory and underway. We also presented theoretical approaches for the representation and processing of semantics within cognitive systems. Here, image schemas cannot only fit into the representation part. Developing truly human-centred technical devices requires an implementation of basic models of human thought. Here, image schemas as basic building blocks of human knowledge may be a promising starting point for providing users with intuitive human-computer interaction environments that allow technology to work for people by anticipating human knowledge, decisions and needs. Utilising image schemas for the processing part, however, is subject to further research while the mechanisms to do so are already in place. We mentioned first promising results regarding the structuring of available data. These techniques should also be helpful in the context of self-organising systems (Żytniewski, 2017). We presented some results from our implementation of a hardware module where users can interact with a heat control. The system allows switching of modality during interaction and adapts to the users' habits. It learns and forgets.

Acknowledgement

This work has been developed in the project Universal Cognitive User Interface (UCUI), which is partly funded by the German Federal Ministry of Education and Research (BMBF) within the research program IKT2020 (grant #16ES0297).

References

Duckhorn, F., Huber, M., Meyer, W., Jokisch, O., Tschöpe, C., & Wolff, M. (2017). Towards an autarkic embedded cognitive user interface. *Proceedings of "Interspeech,"* 3435–3436.

Ertel, W. (2009). *Grundkurs künstliche intelligenz* (2nd ed.). Wiesbaden, Germany: Vieweg+Teubner Verlag.

Geßler, P. (2017). *Kognitive gerätesteuerung* (Master's thesis, Brandenburg University of Technology Cottbus-Senftenberg, Germany).

Gischer, J. L. (1988). The equational theory of pomsets. *Theoretical Computer Science*, *61*, 199–224.

beim Graben, P., Römer, R., Huber, M., & Wolff, M. (2018). Der fockraum als labyrinth: wissensrepräsentation und sprachverarbeitung am beispiel des mouse-maze-problems. *Proceedings of Elektronische Sprachsignalverarbeitung,* 1-8.

Haykin, S. (2012): *Cognitive dynamic systems: Perception-action cycle, radar and radio*. New York, NY: Cambridge University Press.

Huber, M., & Jokisch, O. (2017). Cognitive data retrieval using a wizard-of-oz framework. In *Proceedings of the 2017 Knowledge Management Conference,* 31–40.

Huber, M., Kölbl, C., Lorenz, R., & Wirsching, G. (2008). Ein petrinetz-modell zur informationsübertragung per dialog. *Proceedings of the 15th German Workshop on Algorithms and Tools for Petri Nets, Algorithmen und Werkzeuge für Petrinetze*, 15–24. Retrieved from http://ceur-ws.org/Vol-380/paper03.pdf

Huber, M., Meyer, W., Nowack, K., & Geßler, P. (2016). WoOF: Ein framework für wizard of oz experimente. *Proceedings of Elektronische Sprachsignalverarbeitung,* 127-134.

Huber, M., & Römer, R. (2015). Modellierung des semantik-syntax grenzübergangs kognitiver systeme am beispiel des "mouse-maze"-problems. *Proceedings of Elektronische Sprachsignalverarbeitung,* 232–239.

Huber, M., & Römer, R. (2018). Agenten in höhren sphären. Quantenmechanische situationsmodellierung am kognitiver system am beispiel des mouse-maze-problems. *Proceedings of Elektronische Sprachsignalverarbeitung,* 127-134.

Huber, M., Römer, R., & Wolff, M. (2017). Little drop of mulligatawny soup, miss sophie? Automatic speech understanding provided by petri nets. *Proceedings of Elektronische Sprachsignalverarbeitung,* 122–129.

Huber, M., & Wolff, M. (2017). Segmenting sequences semantically. Using petri net transducers for the translation from sequential data to non-sequential models. *Proceedings of the International Workshop on Algorithms & Theories for the Analysis of Event Data 2017,* 139–157. Retrieved from http://ceur-ws.org/Vol-1847/paper11.pdf

Hurtienne, J., Stößel, C., Sturm, C., Maus, A., Rötting, M., Langdon, P., & Clarkson, P.J. (2010) Physical gestures for abstract concepts. Inclusive design with primary metaphors. *Interacting with Computers, 22(6),* 475–484.

Karolić, B. (2013). Increasing the availability of information using modern technologies of the open Web to build user interfaces for mobile devices. *Online Journal of Applied Knowledge Management, 1(2)*, 143-155.

Lakoff, G., & Johnson, M. (2011). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic Books.

Lindemann, J. (2016). Semantische interpretation und artikulation mit ausserungs-bedeutungs-transduktoren. *Proceedings of Elektronische Sprachsignalverarbeitung,* 119–126.

Lorenz, R., Huber, M., & Wirsching, G. (2014). On weighted petri net transducers. *Proceedings of the 35th International Conference of Application and Theory of Petri Nets and Concurrency,* 8489, 233–252.

Mandler, J. & Pagán Cánovas, C. (2014). On defining image schemas. *Language and Cognition*, *6(4),* 510-532.

Meyer, W., Huber, M., & Wolff, M. (2018). Verhaltenssteuerung einer universellen kognitiven Benutzerschnittstelle anhand einer heizungssteuerung. *Proceedings of Elektronische Sprachsignalverarbeitung,* 90, 301-308.

Mohs, C., Hurtienne, J. , Kindsmüller, M. C. , Israel, J. H. , Meyer, H. A., & the IUUI Research Group (2006). *IUUI – Intuitive use of user interfaces: Auf dem weg zu einer wissen-schaftlichen basis für das schlagwort "Intuitivität"*. MMI-Interaktiv, 11, 75-84.

Pratt, V. (1986). Modelling concurrency with partial orders. *International Journal of Parallel Programming*, *15*, 33–71.

Römer, R., Huber, M., & Wirsching, G. (2016). Ein beitrag zur gedankengeometrie kognitiver systeme. In O. Jokisch (Ed.), *Proceedings of Elektronische Sprachsignalverarbeitung (ESSV),* (pp. 101–110). Dresden, Germany: TUDpress.

Römer, R., & Wirsching, G. (2013): Ein beitrag zu den natur- und geisteswissenschaftlichen grundlagen kognitiver systeme. *Proceedings of Elektronische Sprachsignalverarbeitung,* 93–102.

Römer, R., & Wolff, M. (2015). Konzeption eines kognitiven systems für den experimentellen einsatz in forschung und lehre. In G. Wirsching (Ed.), *Proceedings of Elektronische Sprachsignalverarbeitung,* 212–213.

Schmitt, I., Römer, R., Wirsching, G., & Wolff, M. (2017). Denormalized quantum density operators for encoding semantic uncertainty in cognitive agents. *Proceedings of the 8th IEEE International Conference on Cognitive Infocommunications*, 165–170.

Silber-Varod, V. & Geri, N. (2014). *Can automatic speech recognition be satisficing for audio/video search? Keyword-focused analysis of Hebrew automatic and manual transcription*. *Online Journal of Applied Knowledge Management, 2(1)*, 104-121.

Turner, P. (2008). Towards an account of intuitiveness. *Behaviour & Information Technology, 27(6)*, 475–482.

Widdows, D. (2004). *Geometry and meaning*. Stanford, CA: CSLI Publications, Lecture Notes.

Wirsching, G., Huber, M., Kölbl, C., Lorenz, R., & Römer, R. (2012). Semantic dialogue modeling. *Proceedings of the Cognitive Behavioural Systems –2102 International*

Training School, Revised Selected Papers, Lecture Notes in Computer Science (pp. 104–113). Berlin, Heidelberg: Springer.

Wirsching, G., & Wolff, M. (2014). Semantische dekodierung von sprachsignalen am beispiel einer mikrofonfeldsteuerung. *Proceedings of Elektronische Sprachsignalverarbeitung,* 104–109.

Wolff, M., Tschöpe, C., Römer, R., & Wirsching, G. (2013). Subsymbol-symbol-transduktoren. *Proceedings of Elektronische Sprachsignalverarbeitung,* 197-204.

Żytniewski, M. (2017). Autopoiesis of knowledge management systems supported by software agent societies. *Proceedings of the 2017 Knowledge Management Conference,* 41–52.

Authors' Biographies

Markus Huber is a researcher at Brandenburg University of Technology Cottbus-Senftenberg and Innotec21 GmbH Leipzig. He studied educational sciences and mathematics at the Catholic University of Eichstätt-Ingolstadt and worked as a computer scientist at the University of Augsburg. Markus holds a graduate degree each in education and mathematics. His current research interests include Petri nets based transductions, formal languages, abstract algebra, quantum theory, semantic processing and cognitive systems.

From 1992 to 1997, **Matthias Wolff** studied Electrical Engineering at Technische Universität Dresden. He received the Dr.-Ing. degree and his Habilitation in 2004 and 2011, respectively, also from TU Dresden. Since 2011, he has been a full professor of Communications Engineering at Brandenburgische Technische Universität Cottbus-Senftenberg (BTU). He is a co-founder of the Cognitive Systems Research Cluster at BTU and has been the speaker of the cluster since 2015. Matthias Wolff is a university teacher and researcher on the fields of acoustic pattern recognition, speech technology, and cognitive technical systems. He also lectures on systems theory, communications engineering, and audio and signal processing. Matthias Wolff is member of IEEE and VDE. For more information, see https://www.b-tu.de/fg-kommunikationstechnik.

Werner Meyer is a researcher at Brandenburg University of Technology Cottbus-Senftenberg. He studied electrical engineering at DHBW-Mannheim and Brandenburg University of Technology Cottbus-Senftenberg and holds a Master of Science. Werner's current research interests include behaviour control of cognitive systems and their algorithmic formulation, including decision processes, semantic processing and abstract algebra.

Oliver Jokisch is teaching as a professor for signal and system theory at the Leipzig University of Telecommunications (HfTL), Germany. He studied information technology at TU Dresden in Germany as well as at the Loughborough University in United Kingdom. Oliver graduated as a diploma engineer and holds a PhD degree in information technology from TU Dresden. His research is dedicated to different areas in audio and speech communication such as audio coding, speech prosody, speech synthesis and language learning systems. He served as local chair of the 6th Workshop on Speech and Language Technology in Education (SLaTE'15) and of the 27th Electronic Speech Signal Processing Conference (ESSV'16). Beyond he acted as publicity chair of the 16th Annual ISCA Conference (INTERSPEECH'15). Oliver co-founded and mentored the

IT-oriented companies voiceINTERconnect GmbH, Linguwerk GmbH, COSEDA Technologies GmbH and ambisone GmbH as well as the education and knowledge management firm IBWM GmbH. Personal site: http://www1.hft-leipzig.de/ice/jokisch_en.php

Kati Nowack at the chair of Applied Media Sciences, Brandenburg Technical University Cottbus-Senftenberg holds a Ph.D. from Humboldt University at Berlin in psychology. Her fields of expertise cover both cognitive strategies in human-computer-interaction as well as the human perception of time (e.g. time succession, time perspective, circadian preferences). She applies empirical research methods derived from cognitive psychology and psychophysiology (e.g., reaction times, pupillary response, eye movements).