The 2016 AAAI Fall Symposium Series:
Cognitive Assistance in Government and Public Sector Applications
Technical Report FS-16-02

Persuasive AI Technologies for Healthcare Systems

Daniel Sonntag
German Research Center for Artificial Intelligence
Stuhlsatzenhausweg 3
66123 Saarbruecken, Germany

Abstract

Cognitive assistance may be valuable in applications that reduce costs and improve quality in healthcare systems. Use cases and scenarios include persuasion, i.e., the design, development and evaluation of interactive technologies aimed at changing users' attitudes or behaviours through persuasion, but not through coercion or deception. We motivate persuasion for healthcare systems and propose solutions from an artificial intelligence (AI) perspective for conceptual design and system implementation. The goal is to develop an IoT (Internet-Of-Things) toolbox towards AI-based persuasive technologies for healthcare systems.

Motivation

Smartphones and watches have the potential not only to help collect user data (wearable devices can detect sensor data in user contexts), but to help the user stay on track for achieving a personal health-related goal. They can for example communicate health-related messages to the user (see, e.g., mobile notifications). *Persuasion technologies* can hence be contextualised: they can be issued in contexts when the chance is higher that the notification is being attended to and is not ignored. This is a step towards widely available medical decision support systems, delivering health interventions by multimodal interfaces.

The approach we attempt to make by persuasion is moving people in certain directions: to pay full attention and possess complete information for self-control by, e.g., smartphone notifications, augmented reality glasses, or robot companions. The approach assumes that user choices are not blocked, fenced off, or significantly burdened. Two different approaches of choice architecture are relevant: first, the one promoted by (Thaler and Sunstein 2009) with the key claim that real people make mistakes systematically, and that people are often subject to making mistakes that are the result of widely occurring biases, heuristics, and fallacies. These include for example when people predict the frequency of an event based on how easily an example can be brought to mind; or when people are very likely to continue a course of action since it has been traditionally the one pursued, even though this course of action may clearly not be in their best interest. Second, (Jameson et al. 2014) promote a choice architecture for human-computer interaction with the key goal of providing a solid understanding of the psychology of choice and decision making, i.e., how do people go about making choices in their everyday lives, with or without computing technology and strategies and multimodal technologies for supporting everyday choices. Here, persuasion means to "increase the likelihood that the chooser will choose a particular option (e.g., fruit salad instead of cake); or will choose an option from some particular class (e.g., fruits and vegetables); or adopt a particular goal (e.g., eat in a more health-conscious way)." (Fogg 2002) introduces the useful distinction of macro- and micro-level persuasion; macro-level persuasion is designed to encourage people to reduce the risk of diabetes for example, especially if they are overweight or have a family history of the disease. Micro-level elements can be reminders and visualizations towards the goal of changing health-related habits, e.g., to get more physical activity and improve blood sugar control by eating more bread, pasta products and cereals. In the context of multimodal dialogue, micro-level persuasion could be implemented by clever system initiative or turn-taking behaviour related to personal health records (Saparova 2012) and *quantified self*, to reason about target health behaviours.

We can identify special target groups in the multimodal-multisensory interface context as public sector applications: cognitive assistance in terms of persuasion can improve quality care in healthcare systems. For example, young children in educational contexts (e.g., traffic, nutrition), adults with obesity, or dementia. Persuasion complements reality orientation and validation dialogue (Sonntag 2015). *Quantified self* is often viewed as a strategy to enable persuasion as pure information can be an effective nudge. Fundamental science and technology for cognitive assistance requires conceptual design and prototypical system implementations to evaluate societal aspects of cognitive assistant-people interactions. Therefore, IoT frameworks become more important for *quantified self* solutions as they collect the required multisensory data streams. In addition, personalised persuasive systems and persuasion through gamification and serious games have great potential. We develop a prototypical IoT toolbox towards AI-based persuasive technologies for healthcare systems.

Conceptual Design

This research is situated within a larger project (Sonntag 2015) with the ultimate goal of developing a companion robot (Biundo et al. 2016) that gets user information from the IoT framework and exhibits comprehensible multimodal dialogue for persuasion and is entertaining to interact with. The design of interactions, experiences, and persuasion strategies is the point at which the AI experts will combine technology for understanding people, their processes, their needs, their contexts, in order to create scenarios in which AI technology can be integrated: to predict intentions or anticipate actions; then to change users' attitudes or behaviours (e.g., eat in a more health-conscious way).

Context Modelling

Context modelling takes place at the different levels of persuasion, i.e., increase the likelihood that the chooser will choose a particular option, choose an option from some particular class or adopt a particular goal. Context modelling subsumes the physical activity context and the task context (Bettini et al. 2010). Context attributes are, e.g., context objects, location, and user activity. Therefore, activity recognition has a crucial role to play in our scenarios. AI-based models of context information include to capture ontology-based classes and relations, as well as axiomatic models (e.g., constraints on conflicting location information) and software architecture for observing and modelling human activity, to achieve context awareness (Crowley et al. 2002).

Scenarios

Interpreting, using, and developing behavioral theory (Hekler et al. 2013) tries to include behavioral theory in persuasion-related presentation design. In this way, important conversation rules of persuasive messages for given scenarios (figure 1) and situations of brief motivational interventions (Lisetti, Amini, and Yasavur 2015; Gelissen and Sonntag 2015) can be formulated. (1) interprets the eye tracking signal and offers nutrition facts label for unpackaged foods (orange rectangles illustrates persuasive information); (2) illustrates the use of wearable IoT frameworks to allow quantified self solutions to generate context data for persuasive information; (3) depicts the use of mixed reality interfaces (head-worn displays) to present personalised persuasive information that is registered into the environment and spatial context (Kiyokawa 2016); (4) illustrates to self-report food and medication intake data via mobile devices for personalising persuasive information; (5) illustrates the medical cyber-physical system character, to give persuasive feedback messages in real-time. In order to include information from multiple biosensors and other low-level context sensors (e.g., 2D and 3D vision sensors), this information must be turned into higher level context information, a semantic interpretation. In order to recognise activities and detect event and situations, low-level contextual cues must be integrated into a (human-understandable) ontological representation. User defined context labels of situations, events or even emotions are a prerequisite for supervised machine learning experiments in this sector—dominated by activity recognition.

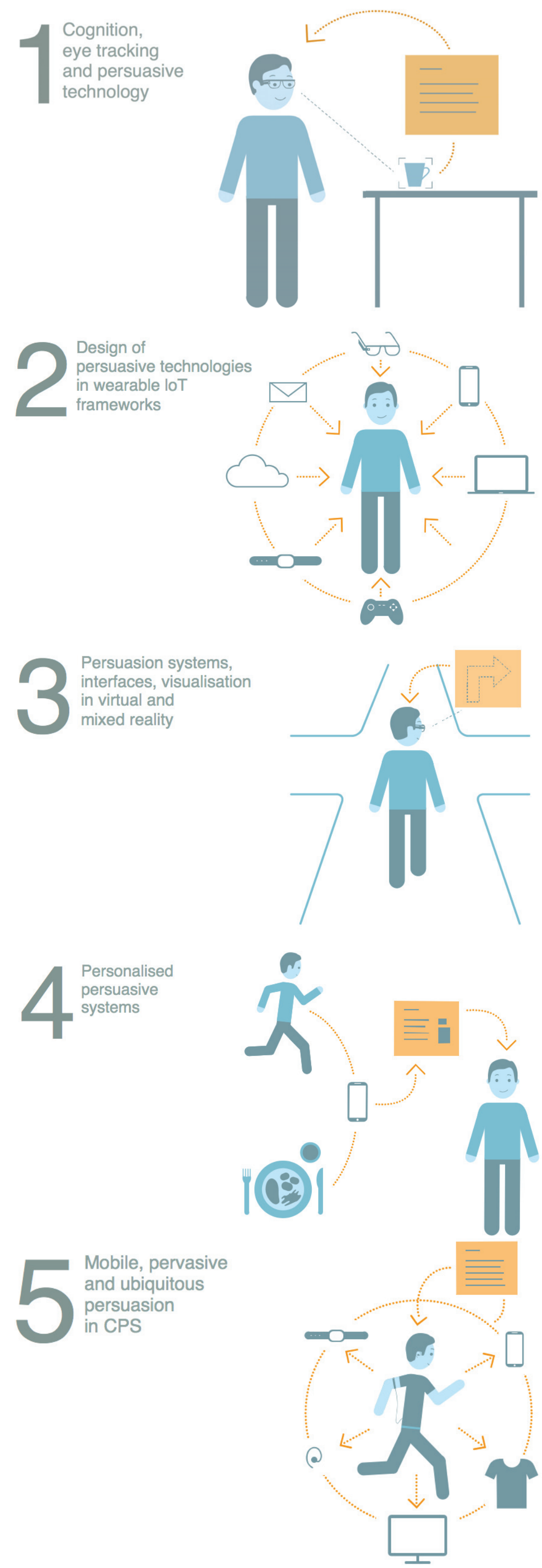

Figure 1: Persuasion Scenarios

Cameras

Device	Output	Platform
Webcam (UVC)	RGB	Windows, Linux
Intel RealSense F200	RGB, depth, point cloud	Windows
Creative Senz3D	RGB, depth	Windows, Linux
PMD Nano	RGB, depth, amplitude	Windows, Linux

Interaction Devices

Device	Output	Platform
Leap Motion	hand tracking data	Windows, Linux
Myo	IMU, EMG data	Windows

Figure 2: 2D and 3D Spatio-Temporal Sensor Devices

System Implementation and AI Toolbox

Along with each of these scenarios and mobile situations and contexts, new AI tools are being developed. Although they are all in early states of development, we can expect that either they will become part of an AI ”toolbox” for persuasion over the next few years.

We implemented a sensor network architecture to observe ”states” of the physical world and provide real-time access to the data for interpretation. In addition, context-aware applications may need access to a timeline of past events (and world states) in terms of context histories for reasoning purposes. In addition, the prediction of future states (for goal-oriented systems such as persuasive multimodal dialogue systems) can make substantial contributions to a context management system. Context attributes are, e.g., context objects, location, and user activity. Activity recognition has a crucial role to play in our scenarios:

- For semantic labelling, we developed a multi-channel annotation tool that works on several input devices (figure 2) and produces events on a timeline (episodic memory).
- For automatic object detection (figure 3), we (1) incorporated the gaze signal and the egocentric camera of the eye tracker to identify the objects the user focusses at; (2) employed object classification based on deep learning and common objects in context of non-iconic images based on (Lin et al. 2014). Our architecture uniquely combines image patch detection with object detection of non-iconic images, and image descriptions (Karpathy and Li 2015).
- For activity recognition, we use a case-based reasoning system (Toyama and Sonntag 2015) that is combined with conceptNet[1].

The persuasion architecture with the multimodal-multisensory dialogue interface is shown in figure 4. The sensor network architecture is displayed in orange, the input interpreter and dialogue framework is the second tier (Neßelrath 2016); the databases are in the backend. Uncertainty of context information, and the need to reason about uncertainty for inferring context models, brings us to differentiate between several representations of context models in the dialogue architecture. The backend databases as context models consist of an episodic timeline and a Cloudant DB

[1]conceptnet5.media.mit.edu

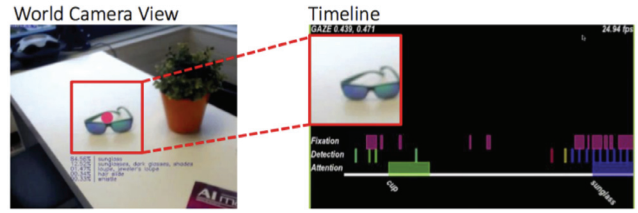

Figure 3: Gaze-Guided Object Classification

which covers higher-level event abstractions. Of particular interest is the face tracker (figure 5), which allows for emotion recognition and gaze direction estimation: with the help of this tool that works with NAO’s camera, a webcam or a smartphone camera on a wheelphone robot, we can avoid head-worn eye trackers and increase the usability and utility in educational contexts with children or dementia patients. Facetracker will be combined with model-based hand-pose estimation (Bellon et al. 2016). See the complete conceptual design videos[2] and the current system implementation of the AI toolbox video[3].

Future Research

The goal is to apply the persuasion ”toolbox” to treat patients as intelligent healthcare consumers and to bring into balance their understanding of their own situation and persuasive decision support technology (nudges). We will evaluate the significance of expressiveness and attention in human-robot interaction for persuasion (Bruce, Nourbakhsh, and Simmons 2001) and architectures for cognitive computing, such as Watson, that can support persuasion dialogue. The following questions raise some issues: How will the data from the experiment be gathered? Will it be complete? Will it interfere with the anticipated normal use?

References

Bellon, R.; Choi, Y.; Ekker, N.; Lepetit, V.; Olasz, L. M.; Sonntag, D.; Tősér, Z.; Yoo, K.; ; and Lőrincz, A. 2016. Model based augmentation and testing of annotated hand pose dataset. In *KI 2016: Advances in Artificial Intelligence - 39th Annual German Conference on AI*.

Bettini, C.; Brdiczka, O.; Henricksen, K.; Indulska, J.; Nicklas, D.; Ranganathan, A.; and Riboni, D. 2010. A survey of context modelling and reasoning techniques. *Pervasive and Mobile Computing* 6(2):161 – 180. Context Modelling, Reasoning and Management.

Biundo, S.; Höller, D.; Schattenberg, B.; and Bercher, P. 2016. Companion-technology: An overview. *KI - Künstliche Intelligenz* 30(1):11–20.

Bruce, A.; Nourbakhsh, I.; and Simmons, R. 2001. The Role of Expressiveness and Attention in Human-Robot Interaction. In *AAAI Fall Symposium Series*.

[2]http://www.dfki.de/MedicalCPS (PERSUASIVE tab)
[3]http://www.dfki.de/MedicalCPS/Nao.mp4

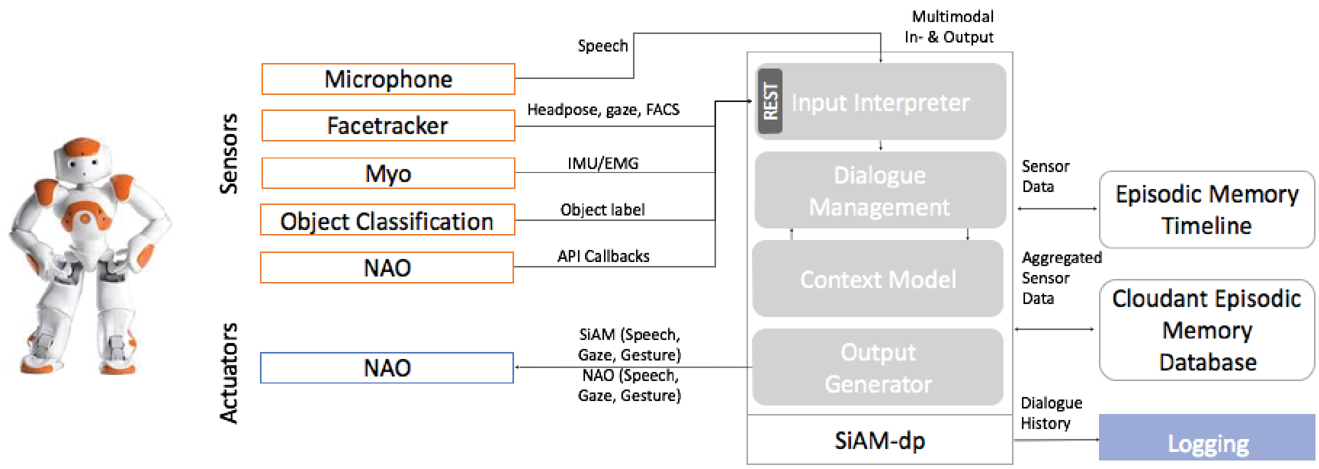

Figure 4: Multimodal-Multisensor Architecture with NAO Robot

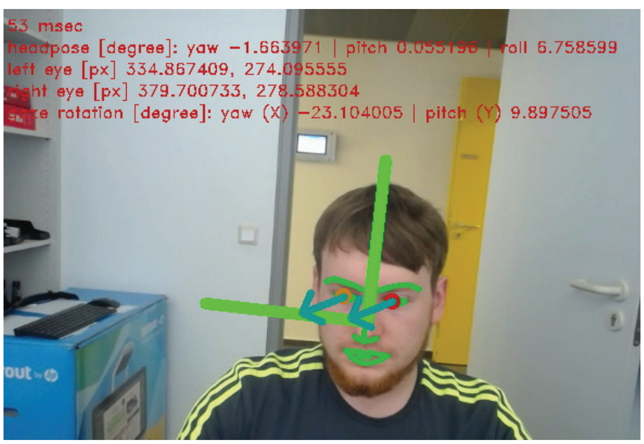

Figure 5: Facetracker (Tősér et al. 2016)

Crowley, J. L.; Coutaz, J.; Rey, G.; and Reignier, P. 2002. Perceptual components for context aware computing. In *Proceedings of the 4th International Conference on Ubiquitous Computing*, UbiComp '02, 117–134. London, UK, UK: Springer-Verlag.

Fogg, B. J. 2002. Persuasive technology: Using computers to change what we think and do. *Ubiquity* 2002(December).

Gelissen, J., and Sonntag, D. 2015. Special issue on health and wellbeing. *KI - Künstliche Intelligenz* 29(2):111–113.

Hekler, E. B.; Klasnja, P.; Froehlich, J. E.; and Buman, M. P. 2013. Mind the theoretical gap: Interpreting, using, and developing behavioral theory in hci research. In *Proceedings of CHI*, 3307–3316. New York, NY, USA: ACM.

Jameson, A.; Berendt, B.; Gabrielli, S.; Cena, F.; Gena, C.; Vernero, F.; and Reinecke, K. 2014. Choice architecture for human-computer interaction. *Found. Trends Hum.-Comput. Interact.* 7(12):1–235.

Karpathy, A., and Li, F. 2015. Deep visual-semantic alignments for generating image descriptions. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, 3128–3137. IEEE Computer Society.

Kiyokawa, K. 2016. Redesigning vision by head worn displays. In *Imaging and Applied Optics 2016*, TM4A.2. Optical Society of America.

Lin, T.; Maire, M.; Belongie, S. J.; Bourdev, L. D.; Girshick, R. B.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; and Zitnick, C. L. 2014. Microsoft COCO: common objects in context. *CoRR* abs/1405.0312.

Lisetti, C.; Amini, R.; and Yasavur, U. 2015. Now all together: Overview of virtual health assistants emulating face-to-face health interview experience. *KI - Künstliche Intelligenz* 29(2):161–172.

Neßelrath, R. 2016. *SiAM-dp: an open development platform for massively multimodal dialogue systems in cyber-physical environments*. Ph.D. Dissertation, Saarland University.

Saparova, D. 2012. Motivating, influencing, and persuading patients through personal health records: A scoping review. *Perspectives in Health Information Management, American Health Information Management Association (AHIMA)*.

Sonntag, D. 2015. Kognit: Intelligent cognitive enhancement technology by cognitive models and mixed reality for dementia patients. In *AAAI Fall Symposium Series*.

Thaler, R. H., and Sunstein, C. R. 2009. *Nudge : improving decisions about health, wealth, and happiness*. New York, Toronto, London: Penguin Books.

Tősér, Z.; Rill, R. A.; Faragó, K.; Jeni, L. A.; and Lőrincz, A. 2016. Personalization of gaze direction estimation with deep learning. In *KI 2016: Advances in Artificial Intelligence - 39th Annual German Conference on AI*.

Toyama, T., and Sonntag, D. 2015. Towards episodic memory support for dementia patients by recognizing objects, faces and text in eye gaze. In *KI 2015: Advances in Artificial Intelligence - 38th Annual German Conference on AI*, 316–323.