Human-Computer Interaction with an Intelligence Virtual Analyst

Denis Gouin, Valérie Lavigne, Alexandre Bergeron-Guyard

Defence R&D Canada, 2459 Pie XI Blvd North, Quebec, Qc, G3J 1X5, Canada
Contact: denis.gouin@drdc-rddc.gc.ca

ABSTRACT
Defence R&D Canada has undertaken an R&D initiative to investigate and develop Intelligent Software Assistant (ISA) technologies to support intelligence analysts in sense making tasks. One aspect of the project is related to the Human-Computer Interaction between the virtual assistant and the users. Two key requirements are to support a dialogue between the ISA and the analyst(s), and to optimize the presentation of the results. In order to address these requirements, a number of technologies need to be exploited, including smart room environments, multimodal interaction, adaptive interfaces, augmented cognition, avatars and storytelling.

Author Keywords
Intelligent Software Assistant, Virtual Assistant, Human-Computer Interaction, multimodal interaction, adaptive interfaces, augmented cognition, avatars and storytelling.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Military intelligence analysts have a mandate to collect, process and analyze information, and disseminate required intelligence products. In the context of modern dynamic military operations, analysts are faced with overload problems (information, task, cognition) and knowledge system technologies become important. In order to better address these problems, it is relevant to go beyond traditional approaches and make use of emerging sense making tools.

Recently, a novel, very promising paradigm in artificial intelligence (AI) has emerged: the Intelligent Software Assistant (ISA), identified by MIT's Technology Review [1] as one of 2009's most promising emerging technologies. The idea behind the ISA is to synthesize the current state of AI research and to develop a personalized assistant that organizes information, learns processes, adapts to changing situations, and interactively supports individuals in their tasks in a seamless, intuitive fashion.

Defence R&D Canada (DRDC) has undertaken an R&D project to investigate and develop ISA technologies towards the development of an Intelligence Virtual Analyst Capability (iVAC). The iVAC is meant as a computerized software assistant supporting the intelligence analysts in sense making tasks, while being ultimately capable of taking on autonomous analytical tasks in concert with other analysts (virtual or human). One aspect of the project is related to the Human-Computer Interaction of the iVAC, which is how users and the ISA communicate with each others.

The initiative is a three-year project started in April 2011. The first part of the project has consisted of a requirement elicitation and a literature survey. This included the identification of technologies that would support intelligence analysts in sense making tasks, such as natural language processing, knowledge engineering and machine learning. One key aspect is the Human-Computer Interaction between the virtual assistant and the users, which includes the dialogue between the ISA and the analyst(s), and the presentation of the results. In order to address these requirements, a number of technologies need to be exploited, such as smart room environments, multimodal interaction, adaptive interfaces, augmented cognition, avatars and storytelling.

RELATED WORK
During the period 2003 to 2008, the US Defense Advanced Research Projects Agency (DARPA), through the PAL (Personalized Assistant that Learns) program, investigated and integrated a number of AI technologies into CALO (Cognitive Assistant that Learns and Organizes). A good vision of the use of an ISA in a military context is provided in the PAL video [2].

Figure 1 shows the main CALO functions. Figure 2 provides examples of CALO's learning capabilities.

KSCO 2012, February 15th – 17th, 2012, Pensacola, FL, USA.

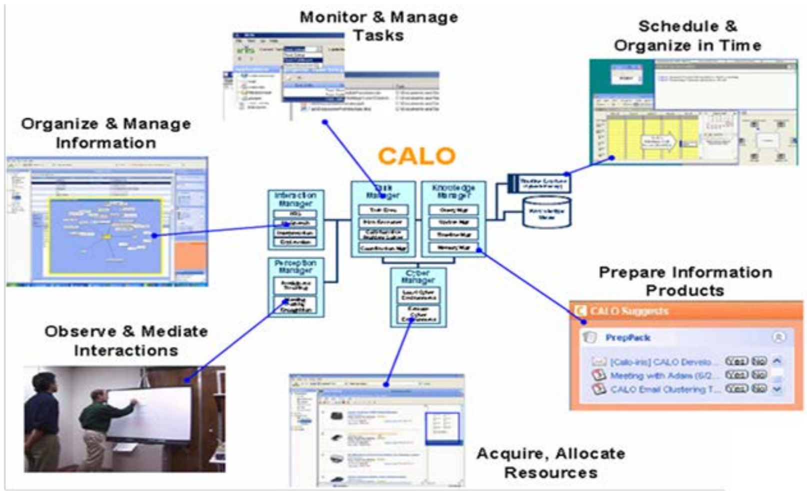

Figure 1 - CALO Functions

CALO Function	Learning In the Wild
Organize Information	Learn associations among email, people, files, web pages, appointments and tasks Learn filing preferences Learn people's expertise Learn email urgency
Prepare Information Products	Learn information priorities for different kind of meetings Learn to summarize documents and email
Observe & Manage Interactions	Learn to track topics Learn to recognize participants Learn to recognize action items Learn vocabulary Learn social network
Schedule & Organize in Time	Learn scheduling preferences (personal) Learn scheduling strategies (by meeting type)
Monitor & Manage Tasks	Learn to recognize tasks Learn to modify tasks
Allocate Resources	Discover new resources Learn resource quality Learn resource acquisition strategies

Figure 2 - Examples of the CALO Learning Capabilities

In the early 2000's, Defence Science and Technology Organisation (DSTO), in Australia, as part of their Future Operations Centre Analysis Laboratory (FOCAL) initiative, has explored the use of virtual advisers. "Virtual advisers are real-time, photo-realistic, animated characters that dialogue with users through spoken-language understanding and speech synthesis… Virtual advisers can brief the command team on a developing situation using text, images, video, and other multimedia; point out significant events for further attention; and suggest alternative courses of action. By tuning the characters' appearance and combining facial gestures and emotional cues, the virtual advisers can also provide context and convey appropriate levels of trust" [3].

Researchers at the Learning Agents Center from George Mason University have developed a "personal cognitive assistant, called Disciple-LTA, that captures analytic expertise, provides effective analytic assistance, and train new analysts and facilitate collaboration with complementary experts and their agents" [4].

More recently, IBM has promoted Watson, a question answering (QA) computing system based on their DeepQA technology. It is "an application of advanced Natural Language Processing, Information Retrieval, Knowledge Representation and Reasoning, and Machine Learning technologies to the field of open domain question answering" [5]. Watson doesn't have a tangible HCI but has created quite a breakthrough in natural language processing and exploitation of massive unstructured data.

A spin-off of the PAL project, SIRI has developed ISA technology onto the iPhone 4s. Siri understand a variety of questions from the user, in natural language, related to his day-to-day activities, such as contacting relatives, holding meetings, looking for restaurants, and it can take appropriate actions. So the user can tell Siri things like

‘Text Ryan I’m on my way’ or ‘Will it be sunny in Miami this weekend?’ [6].

ISA AS A KNOWLEDGE-SYSTEM COMPONENT

ISA can be viewed as a Knowledge System technology that help solve the human cognitive overload by providing assistance to the user by conducting a wide variety of tasks. This includes searching and organizing information, tracking people, managing schedules, assigning tasks, summarizing documents, mediating interactions, guiding and reminding the user, learning procedures and preferences.

IVA HCI REQUIREMENTS

Two high-level HCI requirements have been identified in support of the iVAC. The first requirement is to support a dialogue between the ISA and the analyst(s). This would address the following general questions: How does a user interact with the iVAC in a natural manner? How does a user task the ISA? How does the ISA communicate the results of a question to the user? How does a user input knowledge into the system?

The second HCI requirement deals with the optimization of the presentation of the results. This would address the following general question: How does the ISA adapts the interface to the users’ role, tasks and preferences?

In order to address these requirements, a number of technologies need to be exploited, including smart room environments, multimodal interaction, adaptive interfaces, augmented cognition, avatars and storytelling.

Smart Room Environments

Although an ISA can run on a single computer and be associated to the user of that computer, it is anticipated that in the future, an ISA or ISAs would be available to multiple users in a room. In this case, there would be a need to identify the people in the room or suite of rooms, determining their location and tracking their activities.

Ideally, the users would operate in a smart room environment which can be defined as: a physical space that is instrumented with various networked sensors and devices (e.g. cameras, microphones, biometry, identification tags, computer vision) and support ubiquitous computing permitting to sense, interpret and react to human activity in order to enable better collaboration, increased productivity, creative thinking and decision making. Collaboration can take place within and across meeting rooms.

Multimodal Interaction

An ISA should allow the users to interact with the system(s) using multimodal interaction (e.g. voice, pointing, writing, drawing, gesture, eye/gaze, neural/brain interfaces, emotion detection). Moreover, “full understanding requires identification of speakers and addressees, along with resolution of reference to other participants and objects, and integration of both verbal and non-verbal communication” [7]. In a military context, as illustrated in the PAL video [2], the communication with an ISA will evolve around topics such as tasks, planning activities, standard operating procedures (SOPs), briefing material, documents.

In many cases, as the user will be using natural language interaction combined with movement, there will be some ambiguities. For example, in [2] the user says: ‘These are my priorities… I’m attending this meeting… I need you to setup my briefing package’, while pointing on the screen. The deictic references to ‘my priorities and this meeting’ cannot be solved by speech alone, and in some cases the location being pointed to cannot be resolved to a high-enough precision by vision alone.

Within the CALO initiative, in order to support the understanding of the interaction taking place and resolve ambiguities, a unified multimodal discourse ontology and knowledge base was designed. This ontology is coupled with a dialogue-understanding framework which maintains and shares multiple hypotheses between discourse-understanding components [7].

Information Presentation and Adaptive Interfaces

From an HCI perspective, the ISA should be able to interact with the users by conducting several tasks: Present tools and information; answer questions; ask for clarification; remind the user of some tasks or procedures, and propose alternative possibilities. The ISA may even give a briefing (see Storytelling subsection). This interaction is provided through voice output and/or information display. As illustrated in the PAL video [2], the ISA should be able to interact directly with the user’s screen, in: highlighting information elements already displayed; display information in a new window, organized as the user wants to see it, gather a set of documents; filter information based on user requests.

The ISA could also provide tools to capture information about a meeting. For instance, the CALO Meeting Assistant provides for distributed meeting capture, annotation, automatic transcription and semantic analysis of multiparty meetings [8].

The interface should adapt to the role and current tasks of the user, as well as his/her preferences. The ISA should be proactive, observing what the user is doing and his mental state (see Augmented Cognition subsection), and suggest sequences of events. If the user moves to a large screen display, the user should be recognized using biometry and the interface should adapt to the distance of the user to the display, in terms of font sizes, granularity and quantity of information presented.

Augmented Cognition

ISAs are there to support the analysts. They must intervene when asked to or step in at appropriate moments. The later requires more attention. There might be instances when they should not disturb the users. On the other hand, there are moments where the ISA will be very welcomed, as it could provide augmented cognition.

Using neural interfaces and biometry, augmentation cognition tracks the sensory and cognitive overload of the users and employs computational strategies to restore operational effectiveness. This includes: Intelligent interruption to improve limited working memory; attention management to improve focus during complex tasks; cued memory retrieval to improve situational awareness and context recovery; modality switching (i.e., audio, visual) to increase information throughput [9].

ISA Representation / Avatars

The ISA is represented in some way to interact with the user. It could be in the form of a speech output (voice), a simple icon or a two- or three-dimensional representation of a character (avatar). All of these can be personified, in particular from a gender perspective (male, female, neutral). The avatars can also exhibit other characteristics (age, ethnic group, profession - civil vs military) based on the users' social-cultural context and preferences. Facial and voice features of the avatar (serious / smiling, tone of the voice) could reflect the importance / urgency, or certainty of a message. Different avatars could be used to support different ISA tasks. For example, the avatar for the weather analyst might be different, in terms of gender, age and profession, than the one that recommends the course of action.

In most cases, the avatar would be transmitting information by talking. The voice and facial components are then the important elements. However, in some cases, such as during 'on-the-map planning or video conferencing activities', it may be useful to have a full-size avatar embedded in the display as if it were another user. An example of this is in Figure 3 from PAL.

Figure 3 - ISA Avatar Embedded in the Display

"Experiments conducted by DSTO and the University of Adelaide have shown that different facial characteristics displayed by a speaker will influence a user's affective trust, implicitly imparting uncertainty to the user if required, which at times can be a vital form of qualification for any information delivered" [10].

We must remain careful, however. Wark and Lambert [11] discuss at length the spectre of the Uncanny Valley (discovered by Masahiro Mori [12]) where it is postulated that "as you make a simulacrum look more human, people will identify with it more strongly until a point is reached as the simulacrum approaches 'human looking' where their affinity will suddenly drop steeply, as the differences become more important than the similarities". What level of realism should we implement in the avatar? Will Mori's principle still hold as a new wave of people, grown up with 3D / virtual reality gaming experience, become military analysts? More experimentation with avatars will be needed.

Storytelling

ISAs could be used to present information in the form of stories. Because of its richness, "storytelling is recognized as an effective mechanism for establishing shared context and transferring tacit knowledge throughout an organization" [11]. Analysts ultimately tell stories in their presentations, with the stories providing a way to organize evidence by events and by source documents [13].

[14] discusses the formalism of stories. Stories should be organized around the actions or events, identifying the actors, action type, modality of action, context, rationale, but most importantly identify the relationships between these.

DSTO has conducted interesting work in the form of multimedia narratives, similar to television news services. In this work, the virtual advisers act as automated story tellers and combine narrative with multimedia presentation, to convey situation awareness about complex events [11].

IVA IN COALITION OPERATIONS

The recent significant advances made by IBM in its Watson [5] and the inclusion of the SIRI features within the IPhone 4s [6] will provide a great stimulus in moving ISAs into the military. Military staff will be able to use ISAs onto their Personal Device Assistants (PDAs) and within military command centers. It expected that the ISAs will be able to interact not only with users but with one another; in the longer run, one can expect that ISAs will also become available to support coalition operations.

In coalition settings, ISAs could provide the following benefits: Improved interoperability between coalition forces in terms of disseminating information; translating information between languages; sharing differences in tactics, techniques and procedures (TTPs); synchronizing coalition activities; improving cognitive assistance by providing shared coalition awareness and task support; enabling coalition organizations to learn by managing a knowledge base; and providing better and faster decision making through access to comprehensive knowledge developed through time.

CONCLUSION

ISA can be viewed as a Knowledge System technology that helps solve the human cognitive overload by providing assistance to the user by conducting a wide variety of tasks, including searching and organizing information, tracking people, managing schedules, assigning tasks, summarizing documents, mediating interactions, guiding and reminding the user, learning procedures and preferences. One key aspect is how users and ISAs communicate with each others. This paper has presented a number of the enabling technologies and associated requirements, in particular, smart room environments, multimodal interaction, adaptive interfaces, augmented cognition, avatars and storytelling.

REFERENCES

1. Naone, E.(2009), TR10:Intelligent Software Assistant, Technology Review, MIT, March/April 2009, http://www.technologyreview.com/read_article.aspx?ch=specialsections&sc=tr10&id=22117
2. DARPA (2011), Personalized Assistant that Learns video, http://www.youtube.com/watch?v=BF-KNFlOocQ&feature=youtu.be
3. Wark, S., Broughton, M., Nowina-Krowicki, M., Zschorn, A., Coleman, M., Taplin, P. and Estival, D. (2005), The FOCAL Point - Multimodal Dialogue with Virtual Geospatial Displays, in Proc. SimTecT 2005, Sydney, AU, 9-12 May 2005
4. Tecuci, G., Boicu, M., Ayers, C. and Cammons, D. (2005), Personal Cognitive Assistants for Military Intelligence Analysis: Mixed-Initiative Learning, Tutoring, and Problem Solving, in Proc ICIA, McLean, VA, 2-6 May, 2005
5. IBM Corporation (2011), pQA Project: FAQ, http://www.research.ibm.com/deepqa/faq.shtml
6. Apple (2011), Your Wish is its Command, http://www.apple.com/iphone/features/siri.html
7. Purver, M., Niekrasz, J. and Peters, S. (2005), Ontology-Based Multi-Party Meeting Understanding, in Proc. CHI'05, Portland, OR, April 2005
8. Tur, G., Stolcke, A., Voss, L., Dowding, J., Favre, B., Fernandez, R., Frampton, M., Frandsen, M., Frederickson, C., Gracairena, M., Hakkani-Tür, D., Kintzing, D., Leveque, K., Mason, S., Niekrasz, J., Peters, S., Purver, M., Riedhammer, K., Shriberg, E., Tien, J., Vergyri, D., and Yang, F. (2008), The CALO Meeting Speech Recognition and Understanding System
9. Dorneich, M.C., Mathan, S., Creaser, J.I., Whitlow, S.D., and Ververs (2005), Enabling Improved Performance though a Close-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State, in Proc. HCI (Augmented Cognition), Las Vegas, NV, 22-27 July 2005
10. Australian DefenceScience (2008), Clones that Counsel - Technical Sheet, Australian Defence Science, Volume 16, Number 2 2008
11. Wark S. and Lambert, D. - Presenting the Story Behind the Data - Enhancing Situational Awareness, in Proc. MILCOM 2007, Orlando, Fl, 29-31 Oct 2007
12. Mori, M. (1970), Bukimi no tami (The Uncanny Valley), Energy, vol. 7, 1970, pp. 33-35
13. Bier, E.A., Card, S.K. and Bodnar, J.W. (2008) Entitybased Collaboration Tools for Intelligence Analysis, in Proc. IEEE Symposium on Visual Analytics Science and Technology, 19-24 Oct 2008, Columbus, OH, 99-106
14. Baber, C., Andrews, D., Duffy, T. and McMaster, R., Sensemaking as Narrative: Visualization for Collaboration, in Proc. VAW2011, University London College, 7-8 Sept 2011, London, UK