HCI dilemmas for Context-Aware Support in Intelligence Analysis

Daniel Lafond[1], René Proulx[1], Alexis Morris[2], William Ross[2], Alexandre Bergeron-Guyard[3], Mihaela Ulieru[4]
[1] Defence and Security Unit, Thales Research and Technology Canada, Quebec, Canada
[2] Faculty of Computer Science, University of New Brunswick, Fredericton, Canada
[3] Defence Research and Development Canada – Valcartier, Quebec, Canada
[4] School of Information Technology, Carleton University, Ottawa, Canada
e-mail: daniel.lafond@ca.thalesgroup.com

Abstract— **The RECON (REcommending cases based on CONtext) system is a prototype adaptive technology designed to support intelligence analysis using dynamic load balancing and advanced human-machine synergy. RECON combines a brain-computer interface, machine learning, and simulation in order to create an innovative case-based recommendation capability. Several dilemmas emerge when designing joint cognitive systems endowed with an adaptive capacity. Herein, we critically discuss these dilemmas related to human modeling methodology and human-computer interaction.**

Keywords: adaptive system; human computer interaction; context awareness case-based recommendation; brain-computer interface; information relevance; trust; modeling; virtual assistant; complex systems

I. Introduction

Human-machine systems involve the often-complex interplay of human and technological components as interconnected actors sharing a common goal. These systems, while found in many domains, are particularly relevant in the case of defence and security, where intelligence analysts must make effective use of relevant information, communication and logistic systems and technologies to improve situational awareness. Information overload is a critical area of concern for intelligence analysts who must sift through large volumes of data to uncover trends and make sense of unfolding situations [1].

The day-to-day activities of the intelligence analyst are driven by the intelligence cycle, illustrated in Fig. 1.

Figure 1. The intelligence cycle (adapted from [2])

The intelligence cycle is defined as "*the process of developing raw information into finished intelligence for policymakers to use in decision-making and action*" [3]. The intelligence cycle encompasses many sense-making tasks that the intelligence analyst must accomplish in an iterative fashion. Such tasks include: gather relevant information, represent and organize the information in a schematic way that will ease the analysis process, develop an understanding of the situation by subjecting the information to various hypotheses, produce intelligence packages and recommendations for courses of actions, etc.

As described in [4], the overall process is organized into two major loops of activities: (1) a *foraging loop* ([5]) that involves processes aimed at seeking information, searching and filtering it, and reading and extracting information, possibly into some schema, and (2) a *sense-making loop* ([6]) that involves iterative development of a mental model (a conceptualization) from the schema that best fits the evidence. This process is illustrated in Fig. 2.

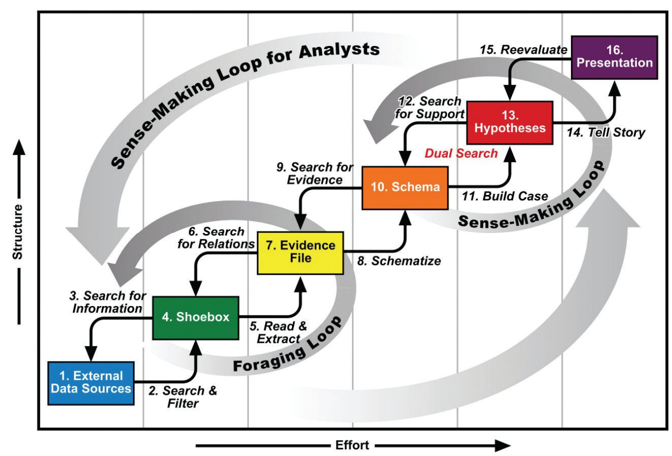

Figure 2. Notional model of sense-making (from [4])

The analyst's activities within the intelligence cycle are subjected to a number of contextual factors (e.g., psycho-physiological or environmental), that can severely impede intelligence analysis due to excessive workload, time pressure, or uncertainty. Below, we present a prototype adaptive technology designed to support intelligence analysis using dynamic load balancing and advanced human-machine synergy and discuss important dilemmas in the design of joint cognitive systems endowed with an adaptive capacity.

II. RECON: CONTEXT-AWARE CASE-BASED REASONING FOR THE INTELLIGENCE VIRTUAL ANALYST CAPABILITY (IVAC)

The Intelligence Virtual Analyst Capability (iVAC) ([7]) is a recent Defence Research and Development Canada initiative that forms an intricate part of a future intelligence analysis capability (FIAC) ([8]). iVAC is a knowledge system with an important human computer interface component that aims to alleviate the problem of cognitive overload by conducting a wide-variety of tasks. This initiative envisions a computerized software assistant supporting the intelligence analysts in sense-making, while ultimately being capable of taking on autonomous analytical tasks in concert with other analysts (virtual or human).

As part of this research, an identification of iVAC sub-capability requirements was performed, based on literature reviews [9] and workshops held with experts from the military, the industry, and academia. The capabilities of the iVAC system were classified into seven broad categories:

- Context management;
- Acquisition of data, information, and knowledge;
- Monitoring, scheduling, management, and evaluation of activities;
- Learning of user and task models;
- Supporting complex intelligence tasks;
- Interaction with humans and other systems.

RECON (REcommending cases based on CONtext) is a context-aware system being developed for integration with the iVAC. The central objective of RECON is to assist the intelligence analysts during the collection, processing, and analysis phases of the intelligence cycle (see Fig. 1), by focusing on context management, monitoring of activities, and interaction with humans. RECON will further the goal of alleviating human-cognitive overload in two ways: firstly, by providing a system capable of sensing the user's contextual state using a brain-computer interface; and, secondly, by adapting to the user's context, identifying other similar contexts, and recommending relevant information to the user based on the system's level of awareness.

The RECON architecture includes the following integrated layer components:

- *Brain-computer interface (BCI) layer.* Classification of user state, and assessment of attention and interest to the displayed information;
- *Human-computer interaction (HCI) layer.* Interface elements and interaction procedures;
- *Data layer.* Information from multiple sources;
- *Context layer.* Transforms information from explicit and implicit sources into contextual knowledge;
- *Case-based recommendation (CBR) layer.* Provides case recommendations based on analyst's context.

The architecture components are conceptually organized according to the relations illustrated in Fig. 3. A more thorough description of the RECON architecture can be found in [10]. The user state classification component of RECON constitutes the heart of the adaptive system capability, performing a real-time analysis of EEG signals, HCI logs, and data to achieve contextual classification.

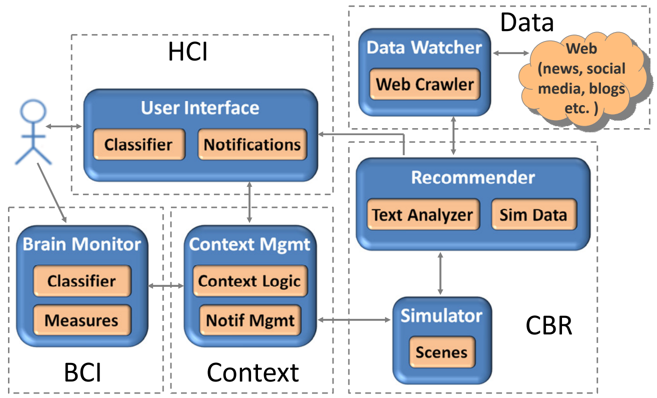

Figure 3. RECON components and relations

The system monitors in real-time the information being reviewed by the user and assesses the degree of interest of the user in regard to that information. This assessment aims to provide critical feedback to the case-based recommendation component, helping it provide more relevant recommendations to the user. Furthermore, EEG signal monitoring allows an assessment of the user's state in regard to the current pressures he/she is facing, which serves to modulate system behavior in accordance with this context (e.g., off-loading the user through adaptive automation, postponing non-critical notifications).

State classification is based on the Contextual Control Model [11], which posits that human decision makers can operate in one of four control modes:

- *Scrambled.* Trial-and-error approach, reactive or random with no forward thinking;
- *Opportunistic.* Based on salient situation characteristics. Lack of anticipation or planning;
- *Tactical.* Procedural or rule-based decision making. Some anticipation/planning;
- *Strategic.* Broad planning horizon with significant anticipation, consideration of high-level goals and global context.

The selection of a control mode is a function of the subjective estimation of the time required to perform a task and of the time available [12]. While the human analyst can dynamically adapt his/her control mode to cope with situational constraints, RECON aims to recognize these changes in control mode and adjust its behavior accordingly. One major technical and scientific challenge is to derive effective and reliable classification models using EEG signals as inputs in the applied context of intelligence analysis. A two-stage process is employed to achieve this. First, an experiment will provide the critical human data necessary for initial model comparison and selection. Second, individual user feedback will allow validation and fine-tuning of the classification rules for each analyst.

Five key dilemmas, relevant to the design of adaptive systems at large are critically discussed below. These generic dilemmas essentially touch human modeling methodology (model selection and calibration) and human-computer interaction (model transparency, user feedback, explicit vs. implicit contextual inputs)

III. Dilemmas

A. Model Selection: Statistics vs. Machine Learning

A first dilemma for modeling user state is whether to opt for statistical analyses based on the general linear model (GLM) or for a machine learning (ML) algorithm to appropriately capture the underlying pattern of cerebral activity associated with a given state. The GLM approach traditionally taught to neuroscientists has a proven track record and comes with robust analysis software, yet the linearity constraint means that complex non-linear relations cannot be 'discovered' using this method (i.e., the underfitting problem). On the other hand, the linearity contraint makes the GLM very robust to noise (i.e., measurement error or intrusions from confounding factors), thus minimising the overfitting problem. Underfitting means that the model lacks functional flexibility to capture a phenonemon, while overfitting means that the model's flexibility allows it to 'fit' both the true regularities in the data but also false patterns that are actually noise (leading to an overestimation of a model's real accuracy)[13]. ML algorithms (or "data mining" algorithms) provide highly flexible models capable of discovering highly complex patterns in datasets. However, the flexibility of ML algorithms makes them highly vulnerable to overfitting.

To resolve this dilemma, the approach proposed here is to concurrently consider models that differ in their functional flexibility and compare their predictive accuracy [14][15]. Indeed, the gold standard in model selection is to assess a model's predictive accuracy by using one (or several) "training samples" for model calibration (i.e., to learn the pattern in the data) and one (or more) "test samples" for model validation. Models that tend to overfit to noise in the data will thus tend to perform worse on the test sample than on the training sample (i.e., a phenomenon called shrinkage)[16]. Alternatively, models that start simple and "grow" to accommodate more complex patterns in the data (e.g., decision trees, cascade correlation) can include stopping rules that check when the prediction error stops improving (i.e., finding the "sweet spot" between underfitting and overfitting).

B. Individual Calibration vs Collective Calibration

A second dilemma relevant to user state modeling is whether to perform model calibration at the group level (i.e., resulting in a single model for all potential users), or at the level of the individual. Clearly, individual modeling has the disadvantage of requiring a new data collection for each user in order to extract an individualized model. Nonetheless, this individualized approach may be necessary in order to reach high levels of model accuracy, particularly when the average is the result of idiosyncratic patterns [17][18]. The alternative is to treat individual differences as noise (leading to a potential underfitting of the user state). The solution proposed herein is to focus on discriminating between broad state categories (as opposed to continuous scales of the concept of interest), which may not require individual user modeling to achieve a satisfactory accuracy.

C. Model Transparency to the User

A third dilemma, related to human-computer interaction, is whether or not to display to the user the model's inputs, its logic, and its resulting assessment. A transparent model offers the possibility to increase user trust, but there is also the risk of a backlash if the user disagrees with the model or simply does not understand it. Conversely a "black box" model may foster doubt and mistrust in the system. This issue also refers to the classic invisibility dilemma which is about choosing between minimizing distractions from the primary task versus providing an added value through explicit interaction [19]. The proposed solution to this dilemma is to make only the model output (e.g., the inferred state) transparent to the user, thus reducing risks and distractions yet allowing the user to develop a sense of trust over time as a function of the tool's classification accuracy.

D. Learning Model Based on User Feedback

The fourth dilemma involves whether or not to collect user feedback in order to sample the correct state at different moments in time, at least for an initial model calibration phase. The alternative is to resort to indirect indicators of user state such as observer judgments or behavior patterns associated with each state (note that unsupervised learning methods are not considered here) [20].

The proposed solution to this dilemma is to combine the two approaches in order to combine self-ratings and observer ratings into a more reliable metric, with observers being supported by access to behavioral markers to help discriminate between the different user states considered.

E. Explicit vs Implicit Contextual Inputs

Knowledge about the context is of primary interest to the RECON system, because we want a system with the capability to adapt to the context. From the point of view of the system the context is what describes the environment, situation, state, surroundings, task, user social settings and roles, and so on. The context evolves according to events and changes occurring during system operation. These events and changes can be introduced by direct explicit interactions from the user (e.g. user overtly indicates current context parameters such as time pressure, psycho-physiological state, availability, critical constraints, current interest in certain types of information, etc.) or implicit interactions based on the situational context (e.g. sensor-based perception, database monitoring, HCI monitoring). Accordingly, the RECON architecture supports both an explicit or implicit classification capability within the HCI layer (see Fig. 3).

The behaviour of the system relies on a combination of the explicit and implicit context which is represented and managed in the context layer of the RECON architecture. The more the system is allowed to respond to implicit interactions, the more it will display agency and autonomy in off-loading the user and selecting what cases to recommend to the user and when. Implicit induction of the context allows the user to concentrate on his or her primary tasks. Explicit user-based contextual inputs can help provide a sense of control over the system and provide contextual data that may not be available otherwise.

The balancing between explicit and implicit interaction support in designing man machine interfaces raises a dilemma because it will affect the usability and effectiveness of the system. A system that relies too much on explicit context will put a heavier workload on the user as he or she must provide a larger amount of information to the system. This will require a more complex GUI and a larger number of manipulations that may interfere with the user's ability to focus on the task at hand. Conversely, a system that emphasizes implicit context frees the user from tedious data input operations but requires the system to monitor data and perform reasoning to infer contextual information. This requires a significant a priori effort to develop effective user-state and contextual classification models.

The proposed solution to this dilemma is expected to come from the results of the upcoming data collection, where some contextual information will be readily extracted from implicit sources (BCI, HCI, and Data layers), while other contextual information may only be obtained through explicit user prompts.

IV. Conclusion

The RECON system currently in development aims to provide an innovative context-aware case-based reasoning framework for the intelligence virtual analyst capability (iVAC). This ongoing research aims to advance the state of the art in the study of context-based systems, case-based reasoning, brain-computer interfaces, and human-computer interaction, through an upcoming proof-of-concept experiment. The research hypothesis is that in situations involving information overload, uncertainty and time pressure, analyst effectiveness can be significantly improved through adaptive off-loading and high-relevance system recommendations derived from the innovative exploitation of BCI data and available contextual information.

Acknowledgment

Thanks are due to Prof. Amedeo D'Angiulli and Prof. Michael Fleming for their insights and institutional support. This work was funded by Defence R&D Canada, by Thales Research and Technology Canada, and by a research partnership grant from the Department of National Defence of Canada and the Natural Sciences and Engineering Research Council of Canada to Prof. D'Angiulli.

References

[1] E. S. Patterson, D. D. Woods, D. Tinapple, E. M. Roth, J. M. Finley, and G. G. Kuperman, "Aiding the intelligence analyst in situations of data overload: From problem definition to design concept exploration," Institute for Ergonomics/Cognitive Systems Engineering, ERGO-CSEL 01-TR-01, 2001.

[2] M. Chesbro, "Intel-Cyclopedia: A Guide to Sources of Information for the Intelligence Community,"Homeland Security Digital Library, 2012. Retrieved from http://www.hsdl.org/

[3] Central Intelligence Agengy, "The work of a Nation,"Library of Congress, 2009.

[4] P. Pirolli, and S. Card, "The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis," Proc. IEEE Symp. Computational Intelligence Analysis, 2005.

[5] P. Pirolli, and S. K. Card, "Information foraging," Psychological Review, 106, pp. 643-675, 1999.

[6] D. M. Russell, M. J. Stefik, P. Pirolli, and S. K. Card, "The cost structure of sensemaking, " Paper presented at the INTERCHI '93 Conference on Human Factors in Computing Systems, Amsterdam, 1993.

[7] D. Gouin, V. Lavigne, and A. Bergeron-Guyard, "Human-computer interaction with an intelligence virtual analyst," in Proc. Knowledge Systems for Coalition Operations, Pensacola, FL, 2012.

[8] D. Poussart, "Future intelligence analysis capability—towards a cohesive R&D program definition," DRDC Valcartier, TM 2012-9999, 2012.

[9] J. Hong, E. Suh, and S. J. Kim, "Context-aware systems: A literature review and classification," Expert Systems with Applications, vol. 36, no. 4, pp. 8509-8522, 2009.

[10] W. Ross, A. Morris, M. Ulieru, and A. Bergeron-Guyard, "RECON: An Adaptive Human-Machine System for Supporting Intelligence Analysis," IEEE International Conference on Systems, Man, and Cybernetics, 2013. pp. 782-787, doi: 10.1109/SMC.2013.138.

[11] E. Hollnagel, "Context, cognition, and control," in Co-operative Process Management, Y. Waern, Ed. London: Taylor & Francis, 1998, pp. 27-52.

[12] M.-E. Jobidon, R. Rousseau, and R. Breton, "Time in the Control of a Dynamic Environment," Proc. of the Human Factors and Ergonomics Society 48th Annual Meeting (pp. 557-561), 2004, doi: DOI: 10.1177/154193120404800360.

[13] S. Roberts, and H. Pashler, "How persuasive is a good fit? A comment on theory testing. Psychological Review, vol 107, April 2000, pp. 358–367, doi: 10.1037/0033-295X.107.2.358

[14] M. Browne, "Cross-validation methods," Journal of Mathematical Psychology, vol 44, March 2000, pp. 108–132.

[15] J. R. Busemeyer and Y. Wang, "Model comparisons and model selections based on the generalization criterion methodology," Journal of Mathematical Psychology, vol 44, March 2000, pp. 171–189.

[16] B. S. Everitt. Cambridge Dictionary of Statistics (2nd Edition), CUP, 2002.

[17] W. K., Estes, and W. T. Maddox, "Risks of drawing inferences about cognitive processes from model fits to individual versus average performance," Psychonomic Bulletin & Review, vol 12, June 2005, pp. 403–408.

[18] P. N. Mohr, and I. E. Nagel, "Variability in brain activity as an individual difference measure in neuroscience?" Journal of Neuroscience, vol 30, June 2010, pp. 7755-7757; doi: 10.1523/JNEUROSCI.1560-10.2010

[19] A. Schmidt, M. Kranz, and Paul Holleis. "Interacting with the ubiquitous computer: towards embedding interaction," In Proceedings of the joint conference on Smart objects and ambient intelligence, October 2005, pp. 147–152.

[20] S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias. "Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment." Multimedia Tools and Applications, vol 41, Feb. 2009, pp. 469-493.