HCI Dilemmas for Context-Aware Support in Intelligence Analysis

Daniel Lafond, René Proulx
Thales Research and Technology Canada
Thales Canada Inc.
Quebec City, Canada
e-mail: daniel.lafond@ca.thalesgroup.com
e-mail: rene.proulx@ca.thalesgroup.com

Alexis Morris, William Ross
Faculty of Computer Science
University of New Brunswick
Fredericton, Canada
e-mail: alexis.morris@unb.ca
e-mail: william.ross@unb.ca

Alexandre Bergeron-Guyard
Command, Control and Intelligence (C2I) Section
Defence Research and Development Canada – Valcartier
Quebec City, Canada
e-mail: Alexandre.BergeronGuyard@drdc-rddc.gc.ca

Mihaela Ulieru
School of Information Technology
Carleton University
Ottawa, Canada
e-mail: mihaela@theimpactinstitute.org

***Abstract*—The REcommending Cases based on cONtext (RECON) 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 and human-computer interaction.**

Keywords-adaptive system; human-computer interaction; context awareness; case-based recommendation; brain-computer interface; information relevance; modeling.

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 Figure 1. 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 sensemaking tasks that the intelligence analyst must accomplish in an iterative fashion. Such tasks include: gathering relevant information; representing and organizing the information in a schematic way that will ease the analysis process; developing an understanding of the situation by subjecting the information to various hypotheses; and producing intelligence packages and recommendations for courses of action.

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

As described by Pirolli and Card [4], the overall process is organized into two major loops of activities: (1) a *foraging loop* [5] that involves processes aimed at seeking, searching, filtering, reading and extracting information, possibly into some schema; and (2) a *sensemaking 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 Figure 2.

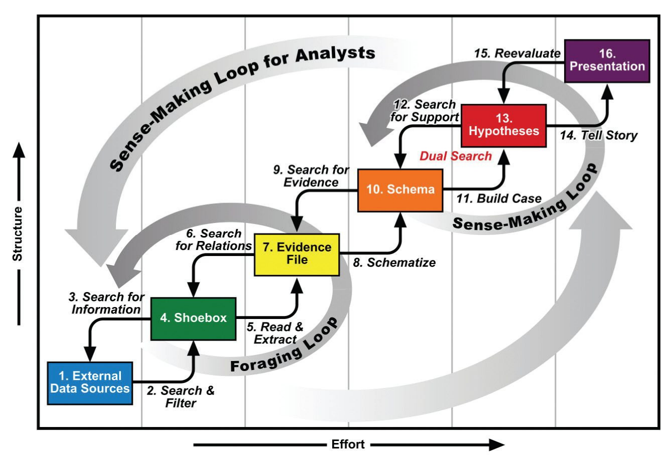

Figure 2. Notional model of sensemaking (from [4])

The analyst's activities within the intelligence cycle are subjected to a number of contextual factors (e.g., psycho-physiological and environmental) that can severely impede intelligence analysis due to excessive workload, time pressure, and uncertainty. The paper is organized as follows. Section II presents a prototype adaptive technology designed to support intelligence analysis using dynamic load balancing and advanced human-machine synergy. Section III discusses important dilemmas in the design of joint cognitive systems endowed with an adaptive capacity. Section IV concludes with a discussion of related work and directions for future work.

II. RECON: Context-Aware Case-Based Recommendation 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 sensemaking, while ultimately being capable of taking on autonomous analytical tasks in concert with other analysts (virtual or human).

As part of the 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;
- Activity monitoring, management, and evaluation;
- Learning of user and task models;
- Supporting complex intelligence tasks;
- Interaction with humans and other systems.

REcommending Cases based on cONtext (RECON) 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 Figure 1), by 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 the system 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:* Classifies user state and assesses user attention and interest to the displayed information;
- *Human-Computer Interaction (HCI) layer:* Presents adaptive interface elements and notifications;
- *Data layer:* Gathers 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 Figure 3. A more thorough description of the RECON architecture can be found in [10]. The context management component of RECON is central to the adaptive system capability, combining HCI logs, data, and user-state classification from real-time analysis of electroencephalogram (EEG) signals to achieve contextual classification.

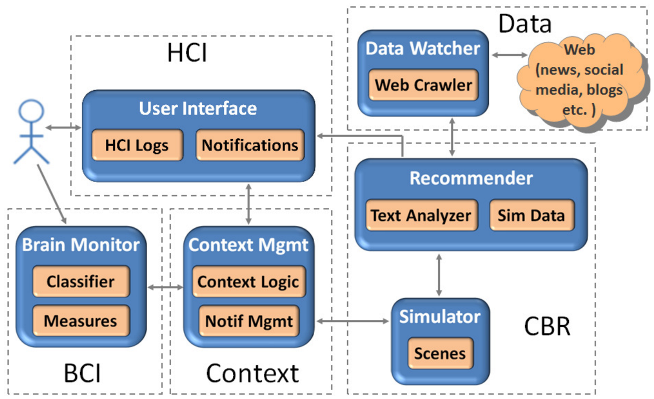

Figure 3. RECON components and relations

The system monitors the information being viewed by the user in real-time and assesses the user's degree of interest 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., reducing user cognitive load through adaptive automation and postponing non-critical notifications).

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

- *Scrambled:* Planning is limited (or non-existent) and actions include trial-and-error, reactive or random approaches with no forward thinking;
- *Opportunistic:* Planning is limited and actions are based on salient situation characteristics;
- *Tactical:* Planning is present but restricted to the current situation and actions are guided by procedural or rule-based decision making;
- *Strategic:* Planning extends beyond the current situation and actions consider 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

Copyright (c) IARIA, 2014. ISBN: 978-1-61208-341-4

dynamically adapt his or 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 [13] in the applied context of intelligence analysis. A two-stage process is employed to achieve this. First, an experimental training set will provide the critical human data necessary for initial model comparison and selection. Secondly, individual user feedback will be incorporated to allow validation and fine-tuning of the classification rules for each analyst. Together with the integrated system components shown in Figure 3, these will allow RECON to achieve its goal of context-aware, cased-based recommendation for iVAC.

III. DILEMMAS

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

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 [14], yet the linearity constraint means that complex non-linear relations cannot be ”discovered” using this method (i.e., the underfitting problem) [15]. 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) [16]. 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 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 [17][18]. 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) [19]. Alternatively, models that start simple and "grow" to accommodate more complex patterns in the data (e.g., decision trees and 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 [20][21]. 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. For example, RECON could use a classification model, such as low, medium, and high, to discriminate among different categories of "interest toward a type of information," instead of using a continuous equal-interval scale.

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 relates to the classic invisibility dilemma which involves choosing between minimizing distractions from the primary task and providing added value through explicit interaction [22].

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. For example, RECON could show the analyst the currently estimated control mode (e.g., scrambled, opportunistic, tactical, or strategic) without displaying the current input values and the classification model.

D. Learning Model Based on User Feedback

A 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) [23].

The proposed solution to this dilemma is to use both approaches in order to combine self-ratings and observer

Copyright (c) IARIA, 2014. ISBN: 978-1-61208-341-4

ratings into a more reliable metric, with observers being supported by access to behavioral markers to help discriminate between the different user states considered. For example, the classification of the control mode in RECON could be calibrated based on feedback in a training phase, using self reports (after the fact) from the intelligence analysts' perceived control mode at differents moments in time, combined with judgments from an expert observer.

E. *Explicit vs Implicit Contextual Inputs*

A fifth dilemma involves knowledge about user context, which is central to system adaptation. Context is what describes the environment, situation, state, surroundings, tasks, social settings, and roles, among other things [10]. This context evolves according to events and changes occurring during system operation either by direct explicit interactions from the user (e.g., a user manually indicates current context parameters such as time pressure, psycho-physiological state, availability, and current interest in certain types of information) or indirect implicit interactions based on the situational context (e.g., automatic data monitoring, HCI monitoring, and sensor-based perception).

Explicitly specifying context affords the user a sense of control over the system and provides contextual data that may not be otherwise available. However, 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, requiring a more complex graphical-user interface and a larger number of manipulations which 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 to combine both explicit and implicit context within RECON. Implicitly, context will be derived from the BCI, HCI, and Data layers (Figure 3), while other contextual information such as a user's current task will be obtained through explicit user input.

IV. CONCLUSION

The RECON system, currently in development, aims at providing an innovative context-aware case-based recommendation framework for the intelligence virtual analyst capability (iVAC). This work builds on previous research in intelligence analysis [7][8], context-aware systems [9], BCI [13], human factors [11][12] and classification modeling [15][20]. It is expected that in situations involving information overload, uncertainty, and time pressure, the effectiveness of intelligence analysts can be significantly improved through context-aware adaptive systems. The approach described in this paper relies heavily on psycho-physiological measurement to infer the user's cognitive state in order to implicitly coordinate the system and the user. An alternative approach is to focus on explicit coordination through human-machine teamwork, enabled through interaction with a virtual assistant [7]. The iVAC initiative seeks to combine these two complementary approaches.

This paper presents five HCI dilemmas for context-aware support in intelligence analysis related to model selection, calibration, model transparency, user feedback, and contextual inputs. Moreover, how these are addressed in RECON is also presented, along with the architecture and core motivations. While the five HCI dilemmas delimit a solution space for designing adaptive joint cognitive systems, the existence of a general optimal configuration is unlikely. The solutions proposed in the context of RECON may not provide an ideal cost-benefit tradeoff in other contexts (and this may also depend on the user). A future design methodology that could parse various combinations and determine the optimal configuration for a given context/user would be very useful. There are also interdependencies between these dilemmas that need to be better understood. Finally, it should be noted that this non-exhaustive list of dilemmas relevant to adaptive systems could be complemented by additional HCI dilemmas such as those identified for supervisory control tasks [24].

With its focus on adaptive off-loading and high-relevance system recommendations, RECON aims to advance the state of the art in the study of context-management systems, case-based recommendation, brain-computer interfaces, and human-computer interaction, through an upcoming proof-of-concept experiment.

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 et al., "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. http://www.hsdl.org/ [retrieved: April 2014].

[3] Central Intelligence Agency, "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, May 2005, pp. 1-6.

[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, Apr. 1993, pp. 1-9.

[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, Feb. 2012, pp. 1-5.

[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, Oct. 2013. pp. 782-787, doi: 10.1109/SMC.2013.138.

[11] E. Hollnagel and D. D. Woods, “Joint cognitive systems: Foundations of cognitive systems engineering.” Boca Raton, FL: Taylor and Francis, 2005.

[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), Sept. 2004, doi: 10.1177/154193120404800360.

[13] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for EEG-based brain–computer interfaces.” Journal of neural engineering, vol. 4, March 2007, pp. R1-R13.

[14] G. D. Hutcheson and N. Sofroniou, “The multivariate social scientist: Introductory statistics using generalized linear models,” Sage, 1999.

[15] M. A. Pitt, W. Kim, and I. J. Myung, “Flexibility versus generalizability in model selection,” Psychonomic Bulletin & Review, 10, pp. 29-44, March 2003.

[16] 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.

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

[18] 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.

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] T. B. Sheridan, “HCI in supervisory control: Twelve dilemmas,” in Human Error and System Design and Management, Elzer, P., Kluwe, R. & Boussoffara, B. (Eds.), Springer-Verlag: London, pp. 1-12, 2000.