Available online at www.sciencedirect.com
ScienceDirect
Procedia Manufacturing 3 (2015) 1052 – 1059
Procedia MANUFACTURING

6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

A Context-Based Approach to Robot-Human Interaction

Wayne Zachary*, PhD[a], Matthew Johnson, PhD[b], R. Hoffman,PhD[b], Travis Thomas[a], Andrew Rosoff[a], & Thomas Santarelli[a]

[a]*CHI Systems Inc., 2250 Hickory Road, Suite 150, Plymouth Meeting, PA 19462*
[b]*Instsitute for Human-Machine Cognition, 40 South Alcaniz St., Pensacola, FL 32502-6008*

Abstract

CARIL (Context-Augmented Robotic Interaction Layer) is a human-robot interaction system that leverages cognitive representations of shared context as a basis for a fundamentally new approach to human-robotic interaction. CARIL gives a robot a human-like representation of context and an ability to reason about context in order to adapt its behavior to that of the humans around it. This capability is "action compliance." A prototype CARIL implementation focuses on a fundamental form of action compliance called non-interference -- "not being underfoot or in a human's way". Non-interference is key for the safety of human-co-workers, and is also foundational to more complex interactive and teamwork skills. CARIL is tested via simulation in a space-exploration use-case. The live CARIL prototype directs a single simulated robot in a simulated space station where four simulated astronauts are engaging in a variety of tightly-scheduled work activities. The robot is scheduled to perform background tasks away from the astronauts, but must quickly adapt and not be underfoot as astronaut activities diverge from plan and encroach on the robot's space. The robot, driven by CARIL, demonstrates non-interference action compliance in three benchmarks situations, demonstrating the viability of the CARIL technology and concept.

Peer-review under responsibility of AHFE Conference

Keywords: human-robot interaction; context model; action compliance

* Corresponding author. Tel.: (215) 542-1400; fax: (215) 542-1400
E-mail address: wzachary@chisystems.com

Peer-review under responsibility of AHFE Conference
doi:10.1016/j.promfg.2015.07.171

1. Introduction

Context is an important part of human-human interaction. Ideally, human-robot interaction should be no different. Unfortunately, context is often overlooked or ignored when designing robotic systems. Traditional engineering approaches focus on controls and autonomy – trying to make the robots as independent as possible. More recently, the view of robots as teammates has grown as the field of robotics has matured (e.g., [1-4]). Intuitively, effective teamwork implies coordination of activity, cooperation among participants and collaboration (see [5]). However, such concepts are generally too abstract to give direct guidance to human-machine system designers and developers. The challenge is to translate high-level concepts such as teamwork and collaboration into specific requirements that can be implemented within control algorithms, interface elements, and behaviors.

We believe that a basis for a solution to this challenge lies in applying the growing understanding of human-human collaboration to re-think the human-robot interaction problem. There are two aspects to this approach. The first is understanding that the main basis for human-human interaction and coordination is the common underlying cognitive architecture that all people share. As Suchman [6] noted in her ground breaking studies of human-machine interaction, human cognition is situated in the social and physical context, and we (people) use our cognitive representation of the situation in acting and *inter*-acting. This common understanding can simplify the communication that is needed for effective interaction, and establishes (usually implicitly) frames for expectations and predictions of each other's behavior. As the scientific understanding of the functional aspects of cognitive architectures have grown (see, for example, [7]), substantial progress has been made in creating computational models that allow context to be represented and even embedded into tools (such as decision support systems and intelligent tutors).

The second aspect of this approach is understanding that automated systems do *not* share this common architecture, and are unable to implicitly participate in interactions that presume it. As Hoffman and Woods [7] put it, "*machines do not know that they are in the world that they have within themselves as a model of that world*". Thus, if machines are to be given any chance of being intelligent subordinates (much less collaborators or cooperators), we would argue that they need to have access to some representation of the situational context *as the humans involved would understand it.* We believe that computational models of context can create this heretofore missing layer needed to enable more efficient, robust, and adaptive human-robot interactions.

A new framework for human-robotic interaction is presented here, based on cognitively-inspired computational representations of context. Called CARIL (the Context-Augmented Robotic Interaction Layer), the framework is intended to give a robot a representation of the context in which it is operating, and an ability to reason about that context to adapt its behavior to the actions of the human astronauts around it. We call this capability "action compliance" – a behavioral/task analog of the "force compliance" concept that has driven much prior research in robot physical and environmental interactions. Force compliance uses information about force to adjust joint control, where as CARIL uses information about context to adjust behavior.

2. CARIL Organization and Architecture

One way of thinking about CARIL is as a cognitive processing system for a robot or as a shared cognitive processing system for multiple robots. That is, CARIL is not physically co-located inside a particular robot, but rather executes separate from the robot (or robots) involved; it can in fact be distributed across multiple computational components. There are many advantages for this, including physical/logistical one: A distributed cognitive system can rely on separate hardware, power-supply, environmental controls, even service access. Moreover, if CARIL is relatively nearby to the robot, the communication delays are minimal. Figure 1 shows how CARIL enables such a distributed cognitive system. CARIL consists of three interacting components. One builds and maintains a model of the situational context relevant to the robot (including the robot's location, plans and intentions). A second component uses the dynamic content of context model to identify action compliance behaviors for the robot. These first two together comprise the actual CARIL. The third component interfaces with the robot to implement these behaviors, and/or to communicate with human co-workers to disambiguate or refine the underlying context understanding.

The research reported here focused on a fundamental form of action compliance – that of "not being underfoot or in the astronauts' way", which is termed non-interference. Specifically, it uses an example case dealing with space exploration, and involves interactions between human astronauts and a robot worker on a space station. Although the entertainment media frequently presents anthropomorphic robots that are both highly intelligent and generally capable of interacting with people in very human-like ways, the humanoid robots of today are much less capable (see [8-10]). While a medium range goal for robots in space exploration and space-based work is to have them perform repair/construction tasks under human supervision or to perform tasks cooperatively with human astronauts, there is little or no use for robots in space environments today. A main reason is safety – the robots can't be trusted not to collide with astronauts or even to avoid unintentionally disrupting astronauts' work. Given this concern, it is problematic to have a robot perform even solitary tasks in the presence of humans. This is not a trivial problem, as evidenced by the fact that intelligent domestic animals (e.g., dogs and cats) cannot reliably achieve this goal, and that children require years of time to develop the cognitive capability to reliably achieve this goal. If a robot can't stay out of a human co-worker's way, it can pose an unacceptable hazard for the humans and to itself. In addition, this capability or skill is fundamental to more complex cooperative skills, such as taking direction from a human astronaut, or cooperating to perform a shared task. In general, for a robot to participate in teamwork, it needs to know not just how to carry out a directed task, but how do so in a way that does not interfere with, bother, or potentially harm the human or other robots.

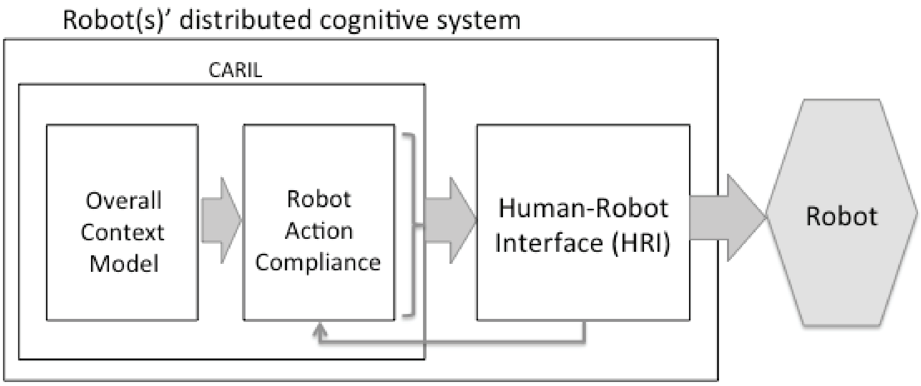

Figure 1. CARIL as A Distributed Robot Cognitive System

The key feature of CARIL is that it uses a cognitively-based computational model of context. A CARIL context model is tailored to a specific work-environment, and includes the physical environment, the work environment (i.e., tasks involved), and the kinds of action compliance that the robot will need in that environment. In the work reported here, the CARIL context model was developed using the Narratively-Integrated Multilevel (NIM) context representation (see [11,12]).

The NIM-based CARIL context representation is a structured, declarative representation, inspired by human situation awareness and narrative reasoning structures. The CARIL context-representation is a hierarchical, declarative, knowledge structure onto which pieces of declarative information are posted. (Note: we say 'on' rather than 'in' because it is easier to conceptualize the representation as a whiteboard onto which pieces of information are constantly being written, (re-)arranged, and erased). The top level of this structure is a *panel*, which consists of a number of *levels* of information, with each higher level representing more abstract information than the lower levels. There are core basic levels on: Perception (what is perceived in the environment), Significance (what it means), Projection (what might happen in the future), and Adaptations (changes to own actions that may required in the future).

Context understanding is built and maintained through a self-organizing process that includes two classes of procedures plus a *conflict-management process.* That is, there are many self- contained reasoning elements (described below), that are activated by external events or that activate themselves based on the presence of specific patterns of declarative information on the context representation.

There are three types of reasoning elements:

- Event-driven procedures that each extract a specific kind of information from the external environment and import it to the context representation, called *perceptual extractor* or *PE* knowledge elements;
- Self-activating procedures that manipulate parts of the context representation, adding details about significance of perceptual information, projections of future possible events of actions, or even how information on the context model can be fit into narratives about the situation. These are called *representation building* or *RB* knowledge elements; and
- Self-activating procedures that use the context information to identify potential situations requiring action compliance and to develop behaviors (including communicative behaviors) to either comply with the situation or add context information that could enable compliance. These are called *action building* or *AB* knowledge elements.

Together all of these reasoning elements are termed the procedural swarm. We note that the action-builder knowledge elements actually constitute the Robot Action Compliance component in Figure 1.

Figure 2 shows the structural components and processing relationships in the CARIL context model. The CARIL system receives a stream of data from various sensory and data communications (both within the robot(s) and within the work environment), and constantly passes these to a cognitive component, labeled the NIM virtual machine (VM). This is a cognitively-inspired, opportunistic computation process, based on the NIM Principles of Operation [12]. Sensory information stimulates the PEs that can process those data to wake up, demand attention, and (eventually) be executed by the NIM VM. Once executed, they post new context concepts on the context representation "whiteboard"; alternatively they can transform information already on it and/or delete other information. As patterns of context information arise on the context representation, various RPs or ABs recognize those patterns as instances of information to that they can apply specific reasoning knowledge. Those RBs and ABs then also wake up, demand attention, and (eventually) come to be executed by the NIM VM. Throughout all this, the NIM VM uses its Principles of Operation (POPs) to police the various context elements demanding attention and ensure that the highest priority and most relevant reasoning element is executed at each quantum of time.

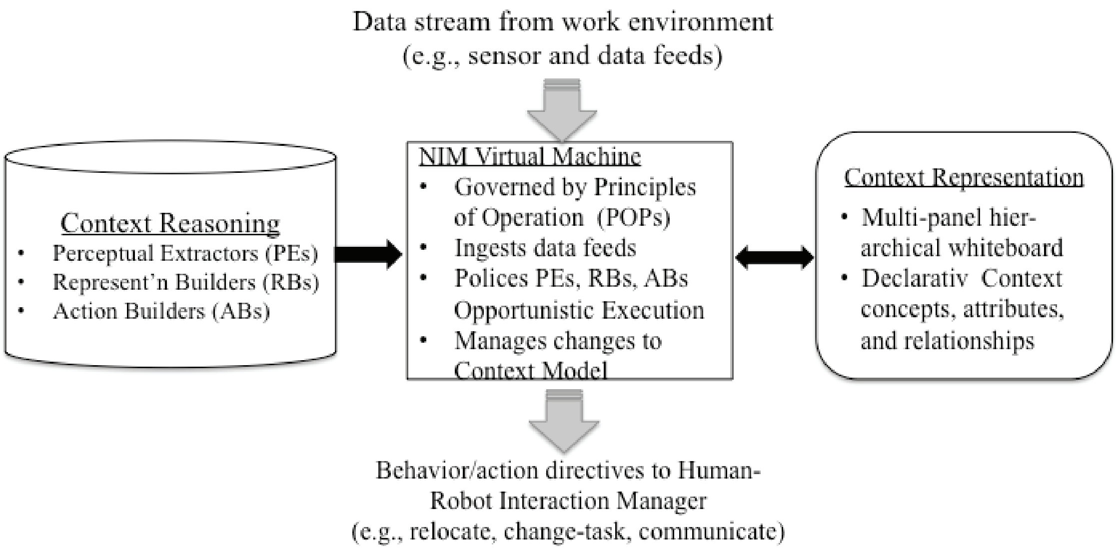

Figure 2. Structure and Organization of CARIL Context Model Processing

As ABs are executed, they create action directions to the Human-Robot Interface manager, providing behaviors that should be executed by the robot to achieve specific action compliance goals.

Figure 3 lists the high level structure of the Context Reasoning elements and the Context Representation. For the Perceptual Elements, each PE name refers to the sensory event to which it can respond. For the Representation Builders, each RB name reflects a new information pattern to which it can apply knowledge. And for the Action Builders, each AB name reflects a possible situational pattern that is associated with a specific action compliance challenge, and about which it can reason and recommend a possible action compliance behavior.

Context Reasoning

Perceptual Elements:
Astronaut update
Robot Update

Representation Builders:
Astronaut Stops Moving
Astronaut Is Moving
Astronaut Arrived at Activity-location?
Unexpected Astronaut Movement
Astronaut Staying Late
Time to Move Robot
Robot Start New Work Segment

Action Builders:
Conflict if Astronaut Staying Late?
Conflict if Astronaut Arrives Early
Conflict if Astronaut Moving Off Plan?
Robot Needs Master Plan Update

Context Representation

Perception
- Astronaut (name, location-vector, new/old flag)
- Robot (name, location-vector, new/old flag)

Significance
- Astronaut (name, status, meta-status, Task)
- Module (name, occupied?, {occupied-by})
- Robot (name, status, meta-status, intention, Task)

Projections/Assumptions
- Astronaut fixed-location (name, exp. location, exp. time-to-end)
- Astronaut path-projection (name, exp. path, exp. time-to-end)
- Robot next-expected-movement (name, exp. time-to-start, destination)
- Robot next-expected-work-activity (name, exp. time-to-start, loc., activity)

Conflict/Adaptations
- Next Movement Blocked
- Unexpected Movement Toward Me

Context-Representation Background Knowledge
- SSS Structure Panel
- Daily Plans Panel

Figure 3. CARIL Context Model for Non-interference Action Compliance

The Perception, Significance, and Projection/Assumption levels of the Context Representation are fixed by the NIM structure (a fourth level, Narrative, was not required in this initial model). The conflict/adaptations level contains context information about specific potential or actual current situations that can require robot action compliance. The Background Knowledge panel provides static knowledge about the physical work environment and the current (in this case, daily) work plan for both human and robot workers. At each level, the bulleted items define specific kinds of concept elements (and their potential attributes), instances of which can be posted at that level of the context representation.

With the context representation and various knowledge elements described above, the robot can develop enough context to figure out:

- where the astronauts and the robot (itself) are located;
- what the astronauts and the robot are doing, and why (i.e., the activity goals);
- how the astronaut activities compare with the plan; and
- what assumptions or expectations can be made about astronaut activities for the immediate future.

It is important that the robot is itself part of this context model. This gives the robot information about itself and allows it to incorporate that information into its reasoning about possible conflict and reasoning to generate action compliant behavior. This is a key foundation for (future) reasoning about communicative interactions with astronauts as part of action compliant behavior. It is also worth noting that in this version of the non-interference problem, there turns out to be no need to make use of narratives. This is because the perturbations in the astronaut schedules are independent and local in time, making them all variation of a common story (normal deviations from schedule). If more systematic variations occur, as would be the case with a general safety alarm, the context model would need to recognize that as an instance of a different story requiring a different adaptation strategy.

3. Simulation-Based Implementation and Test of the Non-interference CARIL Model

The non-interference CARIL Model was implemented and tested in a simulation environment in which a (single) robot works with a group of human workers (i.e., astronauts) in a contained work setting. The test environment consisted of a simple space station in which four simulated astronauts are engaging in a full day of activities, some of which are performed alone and some of which are performed in pairs or teams. A single simulated robot is engaged in background maintenance activities, planned around the astronauts' daily plan such that the robot is always working in unoccupied compartments. As the human activities unfold and sometimes diverge from plan, the robot needs to understand the evolving context of *who-is-where-doing-what-and-why* so that it can adapt its activities and not be in the way of the changing astronaut work activities.

Figure 4 shows the organization of the simulation testbed. CARIL, including both the context model and the action compliance reasoning, was fully implemented and run live. Everything else – the simple space station (SSS), the astronaut activities, and robot's physical presence and activities – was simulated. The locations of each astronaut (and the robot) were sent to CARIL to simulate a motion-sensor feed. Each astronaut and the robot were simulated as separate agents within this physical environment. In the simulation, the activities of each astronaut were directed by an action plan, which was constructed to keep the robot working in different modules than any of the astronauts. The astronauts were given names of Alpha, Beta, Gamma and Delta, and the modules of the SSS were taken from the first six modules of the International Space Station (Destiny, Harmony, Unity, Tranquility, Columbus, and Kibo). Visualization software was created to allow the activity within the simulated SSS and within CARIL to be visually examined as the simulation was running.

To test the adaptation capabilities of the context model, however, multiple changes from the baseline plan were created to emulate various kinds of events that could occur in any normal work day. However, only the baseline version of the plan was made available to the NIM context model. Thus, the robot could not cheat and inspect the actual activities of the astronauts, but could only interpret them according to a context constructed from the daily work plan. It is also noted that this is the situation the other astronauts would face – as human beings, they would adapt to the minor changes as they occurred, using their context understanding and almost without notice. To date, however, this is not the case for robots. However, using commands from CARIL's Action Compliance component, the robot was able to deviate from its plan temporarily and adaptively, to avoid interfering with an astronaut, based on its context understanding and its non-interference goals.

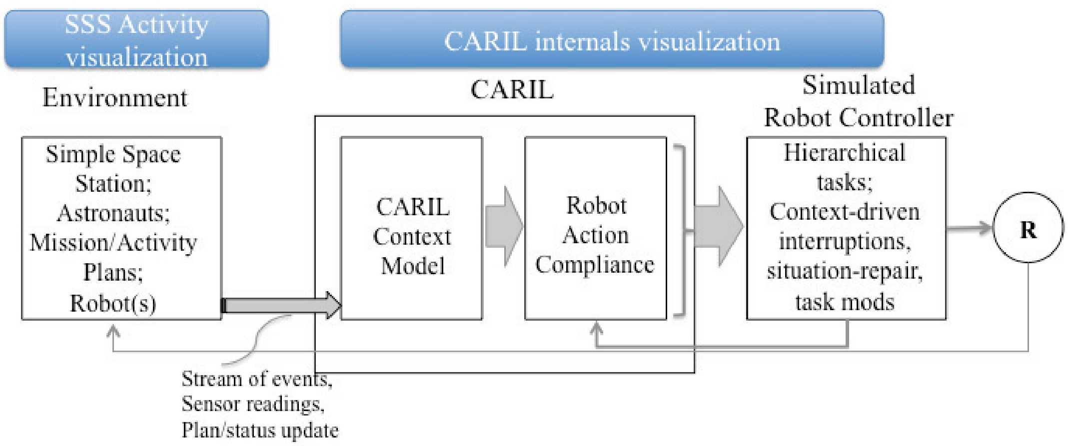

Figure 4. Simulation Testbed to Test and Assess the CARIL Non-interference Context Model

A run of the simulation generates a single workday. When the day begins, the astronauts are all on schedule (i.e., where the schedule indicates they should be at the current time), and the robot's NIM-based context reasoning lets it develop expectations about where each astronaut will be in the immediate future (e.g., how long the astronaut might remain at the current location and where the astronaut should go next). When an astronaut is not where that astronaut should be at (or very near to) the current item, the context model infers that the astronaut is *not* on schedule. All of these inferences and expectations then themselves become part of the context.

The robot goes about its scheduled activities using the baseline daily plan. The context model uses the daily plan, and a constant stream of sensor feeds from space station itself on where each entity is located, at each time quantum (nominally in 10 second intervals), to build and maintain a model of the work context inside the simple space station. The astronauts, however, are guided by a different and perturbed schedule, as discussed above. Each deviation from the plan presents a different type of potential interference which the robot will need to identify and adapt to, using its context model. Three such vignettes are described below. In all three, the context model and its adaptive reasoning capabilities allowed the robot to detect the possible interference situation in advance, and to adaptively avoid any interference with any astronaut.

Vignette 1: At 8:45 Astronaut Gamma is working in Harmony but stays later than scheduled. Without some adaptation by the robot, there could be a potential collision with the robot as it would make its next scheduled move at 9:05 going into Harmony to work. The context model allows the robot to assume that Gamma would leave Harmony at 8:50 and that its own next scheduled movement, from its current location in Columbus, would be at 9:05 to begin work in Harmony. When the astronaut is still in Harmony past 8:50, the context model detects that this violates the assumption about Gamma's next move, and examines whether Gamma's staying there would conflict with the robot's next scheduled move and work activity, which it would. The context model then makes an assumption that Gamma might be working late in Harmony but also assumes that Gamma will eventually go to the next task. From the schedule, this next task is in Tranquility. With this assumption, the context-based adapted reasoning directs the robot to wait in Columbus until Gamma moves. When this move occurs and, as expected, Gamma moves to Tranquility, the context-based adaptive reasoning directs the robot to move to Tranquility and begin the scheduled work activity there.

Vignette 2: At 10:30 in the day, the robot knows via the context model that it is working in Destiny and that Gamma is exercising in Tranquility, where the context model assumes Gamma will stay until about 11:00. However, at 10:45 the context model notes that Gamma has left Tranquility early, and assumes that this is an early departure for Destiny, triggering a potential conflict, as Gamma would be moving toward the robot. The possible conflict triggers the context model to use the context information to let the robot adapt to this early departure. Assuming that the path from Tranquility to Destiny (the assumed destination) will not take Gamma through the adjacent module Harmony, the adaptive reasoning directs the robot to stop working and move into Harmony and wait until Gamma passes. The robot does this, avoiding Gamma. It waits in Harmony until Gamma has arrived in Destiny, around 10:50. Since this is when the robot was scheduled to move to Kibo and began a new task, CARIL directs the robot to move directly to Kibo.

Vignette 3: At 15:45 in the day, the context model shows the robot working in module Harmony. The context model notices that Delta has left the hatch-cleaning task in Tranquility early, and assumes that Delta is proceeding to the next work task in Columbus – which would be towards the robot, and thus creating a new conflict. As it did previously, the adaptive reasoning component uses context model contents to identify an option to move out of Delta's way by going into Kibo. The robot stops its work task and moves into Kibo and stands by waiting. However, it turns out – though the context model would have no way of knowing this – that Delta left the work task in Tranquility to look for a tool, which Delta believed to be in Kibo. So when Delta arrives in Harmony and turns toward Kibo (not Columbus), the context model detects another conflict with Delta. As Delta enters Kibo, the adaptive reasoning component directs the robot to make the only possible adaptation, by freezing-in-place, to make as much room as possible for Delta. After the context model detects that Delta has left and returned to Tranquility, the adaptive reasoning component directs the robot to return to the scheduled task.

4. Conclusions

Using simulation, we examined the robot's CARIL-directed behavior under different types of deviations by astronauts from baseline plan:

- staying late at an activity, blocking robot's movement
- leaving work activity early for next activity
- leaving work in middle of activity to different location (to retrieve a forgotten tool)

CARIL was able to respond appropriately to each, by using its context understanding to recognize a deviation from the plan, recognize deviations that could put it at risk of interfering with astronaut activities, reason about possible adaptive behaviors, and execute the best available option.

An important aspect of this experiment was that explicit communication with astronauts was not an option, yet the action compliance was successful in each case. However, the context reasoning clearly points out places in the reasoning process where communicative strategies could be efficiently used to disambiguate context interpretations and to support adaptive actions. These extensions can be explored in future research.

Acknowledgements

This research was supported in part by Contract N00014-14-P-1186 from the Office of Naval Research for development of the NIM framework and computational implementation, and in part by Contract NNX14CJ38P from the NASA Johnson Space Center, for the application to the human-robot interaction problem. We gratefully acknowledge their support. This paper reflects only the ideas and conclusions of the authors and not the sponsoring agencies; any errors or omissions are ours alone. Special thanks are due to Dr. Kimberly Hambuchen from NASA and Dr. Jeffrey Morrison from ONR for thoughtful comments and insights on our research.

References

[1] Barnes, M., & Jentsch, F. (Eds.). (2010). Human-robot-interactions in future military operations. Farnham: Ashgate.

[2] Kaupp, T., & Makarenko, A. (2008). Measuring human-robot team effectiveness to determine an appropriate autonomy level. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on* (pp. 2146-2151). IEEE.

[3] Hoffman, G., & Breazeal, C. (2007, March). Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction* (pp. 1-8). ACM.

[4] Wang, J., & Lewis, M. (2007, March). Human control for cooperating robot teams. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on* (pp. 9-16). IEEE.

[5] Robertson, S., Zachary, W., and Black, J. , Editors (1990) *Cognition, Computation and Cooperation: An Interdisciplinary Study of Multi-Agent Systems*. Norwood: Ablex Press,.

[6] Suchman, L. A. (1987). *Plans and situated action: The problem of human-machine communication.* New York, NY: Cambridge University Press.

[7] Hoffman, R.R. and Woods, D.D. (2011, November/December). Beyond Simon's slice: Five fundamental tradeoffs that bound the performance of macrocognitive work systems. *IEEE: Intelligent Systems*, pp. 67-71.

[8] Stone, R. & Lavine R., Eds., (2014) The Social Life of Robots. *Science, Vol. 346* 179-203

[9] Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C., van Riemsdijk, B., & Sierhuis, M. (2012). Autonomy and interdependence in human-agent-robot teams. *Intelligent Systems, IEEE, 27(2),* 43–51. doi:10.1109/MIS.2012.1

[10] Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, M. B., & Sierhuis, M. (2014). *Coactive design: Designing support for interdependence in joint activity.* Human Robot-Interaction.

[11] Zachary, W., Rosoff, A., Miller, L., and Read, S. (2013). Context as Cognitive Process. In Laskey, K., Emmons, I., & Costa, P. (Eds). *Proc. of 2013 Semantic Technologies in Intelligence, Defense and Security (STIDS).* CEUR Conf. Proc Vol-1097. http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol 1097/STIDS2013_T07_ZacharyEtAl.pdf

[12] Zachary, W., Rosoff, A., Thomas, T., Iordanov, I, Read, S., and Miller, L. (2015) Toward Computational, Embeddable, Cognitive Models of Context. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015.* NY: Elsevier