Integrative Use of Information Extraction, Semantic Matchmaking and Adaptive Coupling Techniques in Support of Distributed Information Processing and Decision-Making

Paul R. Smart*††, Dave Braines†, Alun Preece‡, Anne Kao§, Stephen R. Poteet§, and Ping Xue§

*School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK
†Emerging Technology Services, IBM United Kingdom Ltd, Hursley Park, Winchester, Hampshire, SO21 2JN, UK
‡School of Computer Science and Informatics, Cardiff University, Cardiff, CF24 3AA, UK
§Boeing Research & Technology, P.O. Box 3707, M/C 7L-43, Seattle, Washington, 98124-2207, USA

††Corresponding author: ps02v@ecs.soton.ac.uk

Abstract—**In order to press maximal cognitive benefit from their social, technological and informational environments, military coalitions need to understand how best to exploit available information assets as well as how best to organize their socially-distributed information processing activities. The International Technology Alliance (ITA) program is beginning to address the challenges associated with enhanced cognition in military coalition environments by integrating a variety of research and development efforts. In particular, research in one component of the ITA ('Project 4: Shared Understanding and Information Exploitation') is seeking to develop capabilities that enable military coalitions to better exploit and distribute networked information assets in the service of collective cognitive outcomes (e.g. improved decision-making). In this paper, we provide an overview of the various research activities in Project 4. We also show how these research activities complement one another in terms of supporting coalition-based collective cognition.**

I. Introduction

In order to press maximal cognitive benefit from their social, technological and informational environments, military coalitions need to understand how best to exploit available information assets as well as how best to organize their socially-distributed information processing activities. The International Technology Alliance program is beginning to address some of the challenges associated with improved cognition in military coalitions by integrating a variety of research and development efforts. In particular, research in one component of the ITA ('Project 4: Shared Understanding and Information Exploitation') is seeking to develop capabilities that enable military coalitions to better exploit and distribute networked information assets in the service of collective cognitive outcomes (e.g. improved decision-making). The research activities in Project 4 are spread across three research tasks:

- **Task 1: Collective Cognition in Military Coalition Environments.** Task 1 aims to explore the effect that features of the military coalition communication environment have on the ability of human and machine agents to engage in collective cognitive processing. The task seeks to identify and model a number of features associated with military coalition communication environments, and it then aims to assess the impact of these features on collective cognition in a multi-agent simulation environment. The primary outcome of this task is an improved understanding of how network features contribute to the cognitive abilities of human-agent collectives.
- **Task 2: Improving Information Extraction through Controlled Language.** Task 2 aims to improve fact extraction capabilities by exploiting controlled natural language (CNL) representations. CNLs will be used in the task to model both the inputs and outputs of the information extraction process, as well as to support human end-users in configuring information extraction tools via these CNL-based input formats. The primary outcome of this task is an improved understanding of the value of CNL representations in supporting information extraction processes.
- **Task 3: Matching Missions to Assets using Natural Language.** Task 3 aims to improve the exploitation of coalition information assets by developing advanced asset matching capabilities. These matching capabilities will exploit knowledge about (e.g.) decision intent, the value of information, and features of the coalition communication network in order to support the efficient exploitation of coalition information assets. The primary outcome of this task will be an advanced information asset exploitation capability that will be sensitive to a range of contextual factors (e.g. network load).

The goals of these tasks contribute to coalition-based collective cognition in the following ways:

1) **Information Extraction:** The information extraction work in task 2 enriches the range of information as-

sets available to coalition agencies. In particular, by applying information extraction techniques, task 2 is supporting the re-representation of information content (from unstructured natural language resources) in a form that is more amenable to machine-based information processing. This enables machine agents to cooperate with their human counterparts in identifying, accessing and exploiting relevant information. The use of CNLs in this task serves as a further means by which productive modes of human-machine interaction are established. In this case, the use of CNLs enables human agents to adapt, configure and orient the activities of information extraction processes[1].

2) **Semantic Matchmaking:** As mentioned above, one of the outcomes of information extraction is the re-representation of (unstructured) information content in a form that is more amenable to machine-based processing. The use of semantically-enriched representational formalisms to encode the outputs of the information extraction process means that information extraction is effectively enriching the space of information available to military coalitions. This enrichment is important because it supports, among other things, the intelligent identification and retrieval of information assets. In particular, the availability of semantic features provides the basis for semantic matchmaking capabilities that match information assets to the information requirements of particular agencies. This is the focus of work in task 3.
3) **Adaptive Coupling:** The work in task 1 will improve our understanding of how the features of coalition communication environments affect collective cognitive outcomes. This understanding can be used to good effect in terms of optimizing the flow of information within a coalition. In particular, the scientific outcomes of task 1 can be used to control the *effective* access that agents have to particular information resources. In previous work, we proposed the notion of 'adaptive coupling' (see [19, 20]) to express the idea that communication networks should support the time-dependent coupling of distributed resources in ways that best meliorate cognitive processing. This notion can be used in the present context to describe techniques and technologies that optimize the information flow dynamics of coalition-based communication systems.

This paper seeks to provide an overview of the research activities in each of the aforementioned tasks. Sections II, III and IV provide an overview of the research in tasks 1, 2 and 3 respectively. Section V illustrates some of the interdependencies between the three tasks and shows how their research foci complement one another in terms of supporting collective cognition in military coalition organizations.

II. Task 1: Collective Cognition in Military Coalition Environments

Traditionally, the main focus of attention in cognitive scientific research has been the individual human agent, and cognition has typically been equated with the information processing operations of the biological brain. Over the past two decades, however, this 'individualistic' focus has been supplemented with the idea that cognition can be socially distributed (e.g. [9]), and this has focused attention on the information processing operations implemented by groups of interacting individuals [8]. Research in this area goes by a variety of names, such as distributed cognition [9], collective cognition [4, 20], social cognition [2, 3] and group cognition [2, 3][2]. However, what all of these locutions have in common is a commitment to the idea that groups can engage in cognitive processing. In particular, groups are seen as implementing cognitive processes, and the cognitive performance of the group (e.g. the quality of solutions, ideas, opinions, products and decision outcomes) is often linked to group-level variables (e.g. group composition, structure and the dynamics of inter-agent information exchange).

The notion of socially-distributed cognition (i.e. collective cognition) is readily applicable to the case of military coalitions. Military coalitions are complex socio-technical organizations in which many cognitive processing routines (e.g. those associated with planning and decision-making) are inherently distributed. As a result, if our goal is to understand how best to improve cognitive function in military coalition contexts, then it makes sense to adopt a more distributed or collective perspective when it comes to the analysis and engineering of new coalition technologies. In particular, if we see cognitive outcomes (e.g. a particular decision) as resulting from information processing operations that are spread across multiple human agents, then our efforts at cognitive enhancement will need to be directed to those features of the communication environment that implement the information processing operations in question (for example, the physical communication networks that support information exchange).

In previous work, we have shown that collective cognitive outcomes are sensitive to variables affecting the time-variant flow of information around a networked ensemble of (synthetic) problem-solving agents [19]. In particular, we showed that dynamic networks (networks with dynamic as opposed to static topologies) were more effective in terms of enabling a collection of agents to discover an optimal problem solution. This led us to propose the notion of adaptive coupling, which expresses the idea that communication networks should

[1]One way of thinking about this is to see the information extraction components as playing the role of specialist sensor systems. The goal of these 'sensors' is not the usual one of encoding and representing forms of environmental energy; rather, it is the productive re-coding of existing information assets: the information extraction components are effectively encoding information in a way that makes that information more suitably poised to participate in response selection and response execution processes.

[2]Out of these terms, we prefer the term collective cognition because it seems most suited to the nature of our research interests. For example, the notions of team and group cognition tend to be reserved for the case of small human groups, whereas our focus is on groups of synthetic agents with no constraints on group size.

support the time-dependent coupling of distributed resources in ways that best meliorate cognitive processing:

> "**Adaptive Coupling Thesis:** In situations where cognitive outcomes depend on the coordinated activity of multiple resources, cognitive performance will benefit from the ability to dynamically and flexibly couple those resources into transient networks of information flow and influence. Dynamic networks support the realization of multiple time-variant patterns of functional connectivity, and these enable the component resources to adaptively coordinate their activity at critical junctures in a collective problem-solving process."

Unfortunately, understanding the precise ways in which information flow dynamics should be controlled so as to meliorate cognition in any given organizational context (e.g. a military coalition) is not straightforward. Much obviously depends on the technological features of the communication environment – the opportunities for inter-agent communication that are supported by available communication devices and associated networks. In addition to this, however, it is also important to understand the psychological, social and legal factors that constrain and shape the profile of information flow in organizational settings. In military coalitions, for example, we encounter policy restrictions on how information can be distributed, especially in joint, multi-national military operations, and these may help or hinder collective cognitive functions in particular ways.

In order to improve our understanding of how best to support collective cognition within coalition organizations, research in task 1 of Project 4 within the ITA program is concerned with the development of a simulation capability that provides insights into how specific features of the military coalition communication environment affects one particular aspect of cognition: the development of shared interpretations of ambiguous environmental information. The simulation capability will be built on top of a model of collective cognition, which has been detailed in previous work [21]. The model features a network-of-networks approach to cognitive modeling, in which the cognitive dynamics of individual agents is influenced by forces and factors at a variety of levels. For example, at the individual level, an agent's beliefs are determined (in part) by the need to establish 'consistency' between all the beliefs that an agent holds. Inconsistencies between belief states provide a rough analogue to the notion of cognitive dissonance (see [5]), which provides the basis for cognitive change at individual (psychological) levels of analysis. The model also incorporates influences at both the social and cultural levels, thereby enabling factors such as social network topology and group affiliation to influence agent cognition.

In the context of our ITA work, the aforementioned model is being used to investigate how a number of factors affect cognitive convergence (the convergence of all agents on a common set of belief states). The target factors in question are drawn from a review of studies exploring the features of the military coalition communication environment. They include:

- Variable inter-agent trust relationships (particularly trust relationships that change as a result of previous interaction or experience).
- Variable certainty in information received from external sources (e.g. variable certainty assigned to information from particular sensors).
- Group-specific differences in communication network structure.
- Partial and restricted views of relevant environmental information (e.g. different agents have different levels of access to particular sensor feeds or information sources).
- Group-specific differences in information sharing policies.
- Differences in background knowledge and beliefs (cultural differences).

The systematic manipulation of these factors in a simulated context will provide insight into how features of the military coalition communication environment affect one particular aspect of collective cognition. Given that we have drawn a parallel between the notion of cognitive convergence in the aforementioned model and the notion of shared situation awareness [13] in the human factors literature, our work will hopefully begin to shed some light into how to improve shared situation awareness by influencing the dynamics of information flow within a coalition organization.

III. Task 2: Improving Information Extraction through Controlled Language

The extraction of task-relevant information from large volumes of unstructured or semi-structured textual content is an open research topic with significant interest from many institutions and individuals worldwide [15]. Given the rise of information production in a military context, and the proliferation of unstructured textual content from both formal and informal (e.g. open source) sources, the ability to rapidly and accurately identify and extract such information is a key capability, especially with the need to decrease the "Data to Decision"[3] timeframe. Any such extracted information can then be used directly by the decision maker or commander to support rapid decision-making and can help to build shared understanding amongst human and machine agents in a coalition context. Other significant capabilities are also enabled through this finer-grained representation of extracted information [17], such as structured query support, compatibility with model-aware policy systems, integration with other structured information sources, and so on.

In order to support the development of an environment in which human and machine agents work together to perform information extraction we propose the use of a CNL representation format, specifically that of ITA Controlled English (CE) [10]. A CNL is a subset of natural language with a

[3] See https://agora.cs.illinois.edu/download/attachments/37359137/ARL_SEDD_SIP_Overview.pdf

restricted grammar to reduce complexity and avoid ambiguity. Earlier research in the ITA led to the creation of CE as an evolution of the earlier specification of Common Logic Controlled English (CLCE)[4] by John Sowa. ITA CE [10] (hereafter referred to simply as "CE") has a formal syntax and semantics and is based on First Order Predicate Logic, but it can be extended to support other logics as required. In the context of task 2 research, CE is used to support the configuration of information extraction components, as well as represent the outputs of information extraction processes. This simplifies the interaction of human agents with the information extraction components, whilst simultaneously preserving the benefits of underlying machine-readable representations.

The basic approach to information extraction in task 2 is a supervised (or semi-supervised) approach, within which the underlying natural language information is analysed and where specific entities, attributes, relationships or other information are identified (these are output in the form of CE sentences). Some agents may be concerned with the structure of the documents, looking for any semi-structured (and therefore easier to process) information, such as information stored in tables, whereas other agents may be more traditional Natural Language Processing (NLP) agents, relying on linguistic rules, computational or statistical methods. Additional agents may be tasked with carrying out logic processing based on the underlying semantics of the conceptual model and the corresponding entities that are being extracted from the corpus. Any agent in the system can access the CE for the conceptual model of the domain (also known as the ontology), enabling these agents to take advantage of this structural knowledge of the domain to aid information extraction capabilities [23] in a CE based form of Ontology Based Information Extraction (OBIE) [1].

In this environment, agents can be configured to account for particular linguistic or structural styles that the natural language authors may adopt. And in the domain of intelligence reports, which are created by specifically trained users, such styles are likely to be more common (compared to less formal open source blogs or newsfeed articles, for example). Such idiosyncrasies of individual datasets can be very valuable to the accuracy of the NLP agents, but only if they can be configured in such a way that it is easy to communicate such information. Also, and very importantly, some agents will be human and can therefore configure and augment the information in ways that machine agents are currently unable to do.

In all cases, the agents are able to consume CE sentences in order to determine their context, and they may produce CE sentences to document their results. The output of one agent may be the input for another, and agents can create CE sentences which affect any aspect of the CE-based domain; e.g. "Facts" (about known entities or relationships), statements which extend the conceptual model of the domain, and logical inference rules which define the semantic relations between the components of the domain. In addition to this, we aim to encapsulate relevant contextual aspects of linguistic communication and interpretation, such as presupposition and conversational implicature [7], into the CE environment so that this information can also be made available to the agents.

We propose that the flexible and extensible nature of this environment, coupled with the human-friendly CE representation format, could provide a more agile and capable environment that is suited for use by intelligence analysts or task-focused specialist users rather than requiring intervention by technical specialists or knowledge engineers.

Our research is motivated by a number of observed shortcomings of current information extraction capabilities. These include:

- **Fidelity:** Specifically, we are aiming to improve the accuracy (precision and recall) of the entities, attributes, relationships and contextual information that can be automatically extracted from unstructured sources in a rapid timeframe, based on more contextual information being provided to the extraction agents, and an improved ability for the analyst user to configure and guide the extraction tools.
- **Configuration complexity:** Through the use of the pervasive CE format for agent input and output we aim to reduce the configuration complexity for the human user, at least in terms of unified language and simplified syntax, if not in terms of actual conceptual complexity since the core NLP problem remains a very challenging problem.
- **Domain specificity:** By directly linking the agents to the analyst's conceptual model of the domain we propose that the domain specificity of the information extraction agents could be increased. It is possible that this may decrease recall in favour of increased precision, but with the ability to also perform rapid updates to the model, such as additional linguistic information, this accuracy and domain specificity could be improved.
- **Disambiguation:** The proposed multi-agent, multi-focus approach may help to improve the issue of disambiguation in certain cases. For example, agents that can classify certain paragraphs of text according to their type, style or position within a document may provide this information for consumption by specializations of more general agents that are tuned to deal with these specific cases (e.g. differentiating between a heading, a caption, a report paragraph and a comment).

Taking the multi-agent CE-based approach outlined earlier, we aim to focus our research in the following areas:

- Pervasive human-friendly representations, specifically CE.
- Rich expressivity (as provided by CE in terms of creating the model, logical inference rules, statements, truth values, uncertainty, assumptions and rationale).
- Iterative feedback mechanisms (allowing all agents (human and machine) to interact via CE both in terms of input and output).

[4]http://www.jfsowa.com/clce/specs.htm

- Investigation of “Simplified Technical English”[5] as a potential mechanism for improving the “processability” of future unstructured information created by trained human users.
- A complementary suite of processing capabilities reflecting the fusion of existing NLP algorithms, inferential processing, human capabilities and other relevant capabilities or linguistic resources.

This research will build upon the earlier ITA research into CE, ontology development, shared understanding [18] and situation awareness, and we hope to demonstrate strong collaboration with other related research areas through the use of the CE representation format. We also aim to evaluate the potential for more rapid integration of information from structured open sources such as linked data, organizational data in Excel spreadsheets or databases. Furthermore, there is good opportunity for use of other ITA-related capabilities and collaborations here.

IV. Task 3: Matching Missions to Assets using Natural Language

An agile ISTAR/ISR information infrastructure requires information to flow in two directions:

1) “Forwards”, from data to decision: a commander needs to take decisions based on actionable information, processed from data collected by sensors or from other sources (e.g. HUMINT, OSINT).
2) “Backwards”, from information requirements to assets: a commander needs to determine what kinds of information will help them achieve their intent, and thereby identify suitable assets.

Ideally, the two flows are connected: working backwards from intent, determine information requirements and hence assets, then work forward from the collected data to make a decision to achieve the intent. Supporting this two-way process requires high-level matching between intent, information needs, and potential decisions (normally expressed in natural language), and information assets (normally expressed in machine-readable formalisms such as queries and service ontologies). Agility requires that this be done as an iterative process, allowing a user to pose problem-space queries (in terms of intent, need, or decision), receive solution-space responses (in terms of potential sources and how they can be accessed), then refine their query accordingly. Suitability of solutions needs to consider what information they can provide, the potential value of that information (VoI), and how they provide it (e.g. a user may need to pull updates asynchronously rather than having a stream of data pushed at them). They should also consider the manner in which the user interacts with the system and the additional knowledge required to do so.

Our previous work in this area took as its starting point the Military Missions and Means Framework (MMF) [16] which was originally defined as a conceptual model to support reasoning about the tasks that military units at various echelons need to perform and the assets available to carry out those tasks. MMF was presented as a structured but not mathematically formal model. Subsequent work formalized aspects of MMF in terms of set theory [22] and as a description logic ontology [6], allowing the matching of tasks to assets to be automated to a certain extent, provided that tasks and assets were expressed in the relevant formal representation. A subset of the MMF description logic ontology is shown in Figure 1. On the left hand side, we have concepts related to the mission: a Mission comprises one or more Operations to be carried out, and each operation breaks down into a number of Tasks that must be accomplished. The important feature of a Task is that it is defined as requiring some capabilities.

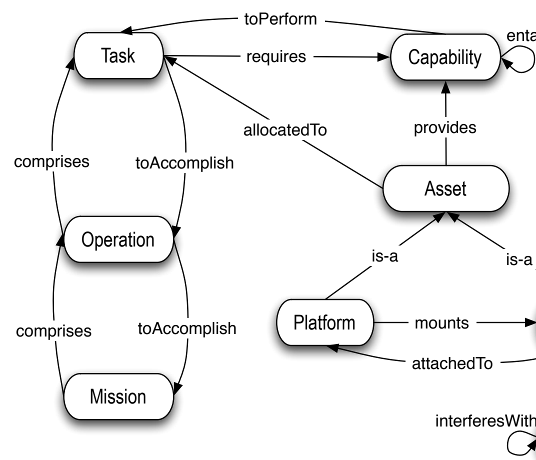

Fig. 1. Main concepts and relations in the MMF ontology

On the right hand side of Figure 1 we have concepts related to capability-provision (“means” in MMF terms). The ontology was developed in the ISTAR/ISR context, with a focus on matching sensors to intelligence tasks. Platform and Sensor are two sub-classes of Asset; a Sensor is a kind of System that can be attached to a Platform; inversely, a Platform can mount one or more Systems. Assets provide capabilities, which link to Tasks as discussed above. Moreover, a Capability can entail a number of more elementary capabilities[6]. Using semantic matchmaking techniques [6] we were able to provide an automated solution to the “backwards” process of matching tasks to assets.

More specifically, we define a Capability as a quadruple, $\langle NC, DS, A, T\rangle$, where:

- NC is one of three primitive capabilty types: detect (find or discover the presence or existence of an entity), distinguish-between (determine that two detected entities are of different types), and identify (name an object by type);

[5]See http://www.asd-ste100.org/

[6]Primitive capabilities are called *functions* in MMF, but we have simplified this.

- DS is a set of "detectable" types of entity (for example, people, vehicles, or installations at various levels of specificity);
- A defines an area-of-interest;
- T defines a period of time.

The primitive capabilities are derived from the National Imagery Interpretability Rating Scale (NIIRS) for various kinds of imagery intelligence [12].

There are two issues with this representation as it stands: Firstly, it is somewhat "unnatural" to military users. To ease usabilty, we have created "form-filling" interfaces, but these still require a user (for example an ISTAR/ISR analyst) to express their information requirements in a somewhat cumbersome manner. Secondly, these information requirements are very low-level. While an analyst may well be interested in specific objects such as vehicles or buildings, it would be desirable for them to state higher-level requirements such as "detect threats to friendly units" in an area of interest. We have previously shown how higher-level information requirements can be manually broken down into our lower-level forms [14]; however, it would be valuable if the analyst could state their requirement *at a high level, using a more natural language*.

We propose CNL as a candidate for an improved representational format for information requirements. Natural language allows a military analyst to express their needs as directly as possible, and it provides a flexible and extensible capability which can react quickly to changing domains or focus areas. CNL – specifically CE – has previously been used for representing mission plans [11]. As this dialect of CE was founded on first-order logic, there exists a mapping from the CE representation via FOL to the description logic MMF ontology, since description logics are less expressive than FOL.

Moreover, mapping the entire MMF ontology to CE — including the "means" part — offers a potentially effective way for a matching system to convey solution options to the user. This could lead to a more "conversational" style of interaction between system and user: the user may refine their requirements (queries) iteratively, adding new information each time, and along the way provide some additional useful information of their own, augmenting the information already available within the system. Rather than seeing this interaction as a disconnected series of separate commands, we are interested in exploring the benefits from seeing this as a "conversation" in which a context is established, where convenient short hand techniques can be used to refer to things in previous query phrases (results) during the conversation. As a simple example, a user may specify a detection task which the system informs them could be achieved via imagery or acoustic intelligence. The user may then refine their requirement to specify imagery. If many users over time specify imagery for that task, the system may infer this as a general preference when making allocation recommendations. Furthermore, by taking into account network constraints (for example, restricted bandwidth) the system can suggest alternatives such as a still image if the user asks for video. Such exchanges are expected to be more convenient when conducted in natural language.

Extending this argument further, having a CNL representation of assets opens up the possibility of users being able to extend the matching knowledge base in a relatively direct way, by describing new means in natural language. This is particularly applicable to the addition of "local" knowledge at the edges of the network. Consider an example where a particular asset has proven highly effective at detecting a certain kind of object in a local context. This could be as simple as recommending a local database of situation reports, or setting the tasking parameters on a sensing system a particular way. If such experience and local learning can be readily incorporated into a local version of the matching knowledge base, the overall system has the potential to become more agile, flexible, and efficient in terms of how ISTAR/ISR resources are used and shared.

In these ways, we turn the knowledge base from a largely static entity (as it was in our previous work) to an evolving repository that captures user's preferences and learning.

Use of CNL also provides us with a more "natural" way of addressing issues of terminology and structural mismatch between different representations across coalition partners, as this is a common problem in a coalition context where no central authority can mandate a universal catalogue of concepts, terms and processes. The proposed use of ontologies, models and CE representations provides a means for aligning alternative model representations from different partners.

V. Task Linkages and Cross-Cutting Themes

Although the research associated with the three tasks in Project 4 has been presented separately, it should be clear that there are a number of research themes that are common to multiple tasks. Some of these cross-cutting themes are described in Table I.

In addition to the cross-cutting themes, the research foci of each task complement one another in respect of the overarching aim of supporting coalition-based collective cognition. For example, the research in task 2 supports the process by which the content of unstructured information resources is made more amenable to machine-based automated processing. This complements the work in task 3, which aims to match human agents with information assets. Importantly, the matchmaking capability relies on the existence of semantically-enriched representations of the kind that are delivered by the information extraction capabilities in task 2. Therefore, task 2 can be seen as enriching the analytic substrate against which task 3 capabilities execute. By re-representing the content of unstructured information resources in a semantically-enriched, machine-readable form, information extraction processes effectively improve the conditions in which matchmaking capabilities operate.

The output of matchmaking processes (task 3) also has an impact on the kind of capabilities targeted by the information extraction processes (task 2). In particular, the matchmaking capability has the potential to reveal information gaps – situations in which there is a shortfall in the information assets

TABLE I
CROSS-CUTTING RESEARCH THEMES.

Research Theme	Task 1	Task 2	Task 3
Shared Understanding	Understand how network-mediated information flows affect the shared interpretation of ambiguous, uncertain or conflicting environmental information.	Support a shared understanding of disparate information via extraction and re-representation in a common format.	Improve inter-agent understanding via ontologies and controlled languages.
Human-Machine Interaction		Develop techniques to enable human end-users to configure information extraction components.	Use conversational dialogue techniques to support human-machine communication.
Semantic Enrichment	Develop an ontology of the coalition communication environment.	Use a shared ontology to represent extracted information content.	Use semantic features to improve asset matching and vocabulary alignment processes.
Controlled Natural Languages		Use controlled languages to support information extraction processes.	Use controlled languages to represent background knowledge and support human-machine communication.
Adaptive Coupling	Understand how to distribute information in a way that meliorates collective cognitive processing.	Provide an environment for rapid configuration and response to re-tasking of IE agents based on emerging contexts.	Support the exploitation of information assets in a way that meliorates coalition decision-making.
Network Modeling	Develop an ontology to support the representation of coalition communication networks.		Develop models to represent network features, particularly those that influence information flow and information retrieval processes.
Value of Information	Understand the ’cognitive value’ of information. For example, understand how the same body of information may yield different cognitive outcomes in different communication contexts.	Provide human/machine friendly representational formats for encapsulating VoI and other related metrics from various (hard and soft) information sources.	Compute and represent the value of information assets in relation to decision-making processes.
Decision Making	Understand how to configure networks so as to best support coalition decision-making.		Understand how to exploit the technological environment in order to improve coalition decision-making.
Information Exploitation	Understand how information assets should be distributed to a multi-agent team in order to support optimal decision outcomes.	Facilitate the exploitation of unstructured information assets by re-representing information content in controlled languages.	Facilitate the exploitation of information assets by using state-of-the-art, context-sensitive asset matching techniques.

required to meet a particular processing objective. This can be fed back to the information extraction components, so that they can be configured, adapted and oriented in ways that address the information shortfall. The linkages with task 2 research should be evident here, since it is part of the remit of task 2 to investigate the ways in which information extraction technologies can be optimally configured so as to support distributed information processing requirements.

Finally, the work in task 1 seeks to provide a better understanding of the factors that affect collective cognition. This is important because it ties in with one of the research aims of task 3; in particular, the notion of delivering information in ways that respect a variety of dynamic constraints (for example, network load). Of course, task 3 is primarily concerned with constraints that are imposed by the physical structure of the coalition communication environment. These include, among other things, bandwidth, power and connectivity constraints. Nevertheless, an additional constraint relates to the need to optimize the flow of information to agencies in a way that meliorates collective cognition. This is precisely the kind of information that will be provided by the task 1 research. Task 1 research can thus be seen as providing additional constraints on information distribution that task 3 can factor into its efforts to support advanced information exploitation capabilities.

VI. CONCLUSION

This paper has described the research activities in one particular component of the ITA, namely Project 4. The overarching goal of the ITA program is to develop capabilities that enable military coalitions to coordinate their efforts in the realization of particular goals. When those goals relate to (e.g.) the specification of plans, the joint understanding of situations and the sanctioning of particular decision outcomes, it sometimes makes sense to see the coalition organization (or aspects thereof) as engaged in a form of collective cognitive processing. Research in Project 4 is concerned with the development of capabilities that 1) enable coalitions to maximally exploit the information resources available to them and 2) adaptively regulate the flow of information resources in ways that meliorate cognition. Together, these activities take us a step further in our efforts to support military coalitions in pressing maximal cognitive benefit from their social, technological and informational environments.

Acknowledgment

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

[1] *Proceedings of the 1st International and KI-08 Workshop on Ontology-based Information Extraction Systems*, volume 400, 2008. DFKI. URL http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-400/proceedings.pdf.

[2] E. H. Chi. Augmented social cognition: using social web technology to enhance the ability of groups to remember, think, and reason. In *35th SIGMOD International Conference on Management of Data*, Providence, Rhode Island, USA, 2009.

[3] E. H. Chi, P. Pirolli, B. Suh, A. Kittur, B. Pendleton, and T. Mytkowicz. Augmented social cognition. In *AAAI Spring Symposium on Social Information Processing*, California, USA, 2008.

[4] I. Couzin. Collective cognition in animal groups. *Trends In Cognitive Sciences*, 13(1):36–43, 2009.

[5] L. Festinger. *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, USA, 1957.

[6] M. Gomez, A. Preece, M. Johnson, G. de Mel, W. Vasconcelos, C. Gibson, A. Bar-Noy, K. Borowiecki, T. L. Porta, D. Pizzocaro, H. Rowaihy, G. Pearson, and T. Pham. An ontology-centric approach to sensor-mission assignment. In *Proc 16th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2008)*, Acitrezza, Catania, Italy, 2008.

[7] H. P. Grice. Presupposition and conversational implicature. *Radical Pragmatics*, pages 183–198, 1981.

[8] V. Hinsz, R. Tindale, and D. Vollrath. The emerging conceptualization of groups as information processors. *Psychological Bulletin*, 121:43–64, 1997.

[9] E. Hutchins. *Cognition in the Wild*. MIT Press, Cambridge, Massachusetts, USA, 1995.

[10] D. Mott. Summary of controlled english. Technical report, Emerging Technology Services, IBM, 2010.

[11] D. Mott, D. Braines, C. Giammanco, and M. C. Dorneich. Hybrid rationale for shared understanding. In *4th Annual Conference of the International Technology Alliance (ACITA'10)*, London, UK, 2010.

[12] National Imagery Interpretability Rating Scale (NIIRS). http://www.fas.org/irp/imint/niirs.htm.

[13] A. A. Nofi. Defining and measuring shared situational awareness. Technical Report CRM D0002895.A1/Final, Center for Naval Analyses, 2000.

[14] A. Preece, M. Jackson, G. Pearson, and T. Pham. Sensor-mission assignment: A scenario-driven walkthrough. In *2nd Annual Conference of the International Technology Alliance (ACITA 2008)*, London, UK, 2008.

[15] S. Sarawagi. Information extraction. *Foundations and Trends in Databases*, 1(3):261–377, 2008.

[16] J. H. Sheehan, P. H. Deitz, B. E. Bray, B. A. Harris, and A. B. H. Wong. The military missions and means framework. In *Proceedings of the Interservice/Industry Training and Simulation and Education Conference*, pages 655–663, 2003.

[17] P. R. Smart and N. R. Shadbolt. The semantic battlespace infosphere: A knowledge infrastructure for improved coalition inter-operability. In *4th International Conference on Knowledge Systems for Coalition Operations (KSCO)*, Massachusetts, USA, 2007. URL http://eprints.ecs.soton.ac.uk/14114/.

[18] P. R. Smart, D. Mott, K. Sycara, D. Braines, M. Strub, and N. R. Shadbolt. Shared understanding within military coalitions: A definition and review of research challenges. In *Knowledge Systems for Coalition Operations (KSCO'09)*, Southampton, UK, 2009. URL http://eprints.ecs.soton.ac.uk/17173/.

[19] P. R. Smart, T. D. Huynh, D. Braines, and N. R. Shadbolt. Dynamic networks and distributed problem-solving. In *Knowledge Systems for Coalition Operations (KSCO'10)*, Vancouver, British Columbia, Canada, 2010. URL http://eprints.ecs.soton.ac.uk/21508/.

[20] P. R. Smart, T. D. Huynh, D. Braines, K. Sycara, and N. R. Shadbolt. Collective cognition: Exploring the dynamics of belief propagation and collective problem solving in multi-agent systems. In *1st ITA Workshop on Network-Enabled Cognition: The Contribution of Social and Technological Networks to Human Cognition*, Maryland, USA, 2010. URL http://eprints.ecs.soton.ac.uk/17743/.

[21] P. R. Smart, W. R. Sieck, D. Braines, T. D. Huynh, K. Sycara, and N. R. Shadbolt. Modelling the dynamics of collective cognition: A network-based approach to socially-mediated cognitive change. In *4th Annual Conference of the International Technology Alliance (ACITA'10)*, London, UK, 2010. URL http://eprints.ecs.soton.ac.uk/20920/.

[22] P. J. Tanenbaum and W. P. Yeakel. A framework linking military missions and means. In *Mathematics for Industry: Challenges and Frontiers*. SIAM, 2005.

[23] D. C. Wimalasuriya and D. Dou. Ontology-based information extraction: An introduction and a survey of current approaches. *Journal of Information Science*, 36(3):306–323, 2010.