CE-SAM: a conversational interface for ISR mission support

Diego Pizzocaro[a], Christos Parizas[a], Alun Preece[a],
Dave Braines[b], David Mott[b], Jonathan Z. Bakdash[c]

[a]School of Computer Science and Informatics, Cardiff University, UK;

[b] Emerging Technology Services, IBM United Kingdom Ltd, Hursley Park, Winchester, UK;

[c] US Army Research Lab, Aberdeen Proving Ground, US

ABSTRACT

There is considerable interest in natural language conversational interfaces. These allow for complex user interactions with systems, such as fulfilling information requirements in dynamic environments, without requiring extensive training or a technical background (e.g. in formal query languages or schemas). To leverage the advantages of conversational interactions we propose CE-SAM (Controlled English Sensor Assignment to Missions), a system that guides users through refining and satisfying their information needs in the context of Intelligence, Surveillance, and Reconnaissance (ISR) operations. The rapidly-increasing availability of sensing assets and other information sources poses substantial challenges to effective ISR resource management. In a coalition context, the problem is even more complex, because assets may be "owned" by different partners. We show how CE-SAM allows a user to refine and relate their ISR information needs to pre-existing concepts in an ISR knowledge base, via conversational interaction implemented on a tablet device. The knowledge base is represented using Controlled English (CE) - a form of controlled natural language that is both human-readable and machine processable (i.e. can be used to implement automated reasoning). Users interact with the CE-SAM conversational interface using natural language, which the system converts to CE for feeding-back to the user for confirmation (e.g. to reduce misunderstanding). We show that this process not only allows users to access the assets that can support their mission needs, but also assists them in extending the CE knowledge base with new concepts.

Keywords: Controlled English, ISR, decision support, analyst, intelligence, conversational interaction, conversational interface, mobile app, CE-SAM, sensor mission assignment, resource allocation

1. INTRODUCTION

A first responder or soldier on the streets of a disaster scenario or a conflict area typically needs to tap into the best-available information about the environment and the situation around them. Sources of information may include cameras, eyewitnesses, patrols, and textual data. Data might, in fact, be produced by ISR sensing assets, such as Unmanned Aerial Vehicles (UAVs), provided with different sensing capabilities based on the mounted sensing devices (e.g. IR Cameras). Other information might instead be already collected e.g. in the form of textual data or, in some cases, it might be generated in real-time from eyewitnesses or patrols on the field. Recently, mobile "apps" have become the common way of accessing information on the go, allowing mobile users to take the best decisions based on facts which were previously ignored.

In the case of soldiers in the field or in an emergency response scenario, where lives are usually at stake, the problem of selecting the right apps - i.e. selecting the right data and information sources - is typically much more complex than a simple search on the "app-stores". We call this the app-selection problem or more simply the *Selection Problem*. One of the important issues which make the selection problem in the ISR domain usually human-intractable is the wide variety of data available and the heterogeneity of source capabilities. Consider for example the latter, in a particular scenario two different ISR sensing assets (e.g. UAVs) might provide infrared imagery or standard visual data (in the visible range). Given a particular information need, like "detecting tracked vehicles", we need to decide which ISR sensing asset (or groups of assets) might be fit to satisfy with a high success-rate our information request. This is a hard problem to solve because the choice of allocating different ISR platforms to that ISR need depends also on how well they work together, e.g. triangulation is achievable with both visual and IR cameras but if only one visual or IR camera is provided then triangulation might be difficult.[1] In addition, an emergency response or military environment is usually highly dynamic, therefore sources might come and go (e.g. ISR assets be destroyed). Sources might be busy supporting other ISR information requests – thus we also need to decide if those sources can be tasked to support different requests considering their priorities. Finally, both military and humanitarian missions are often performed by a coalition composed of multi-national parties and therefore we often have to deal with different access rights to data and diverse sources availability across the coalition.

In such situations, there is clearly a need for an automated mechanism to aid users in the field in choosing the best available "apps" as fast as possible, in order to allow for an improved decision on the course of action. Another important issue is related to the fact that information needs might be expressed via high level requests, e.g. "I am looking for an intruder" rather than "I require a Predator UAV with an IR sensor". By allowing for high level requests as statements of requirement rather than capabilities, we allow non-technical personnel to take advantage of the best available information on the field even if they do not have enough knowledge about sensing asset capabilities or data availability. In parallel, the enormous advantage of apps – which has also determined their success – is that they typically do not require training to be used. In general they pose low cognitive overhead to users in that they are easy to install, configure and use. Therefore, we propose that by both allowing high level requests and wrapping this experience into a zero-overhead/training app is key to delivering problem-focused information to the teams on the field which are often comprised of mainly non-technical people.

To further push the paradigm/success of low training overhead, there is considerable interest in natural language conversational interfaces. These allow for complex user interactions with systems, such as fulfilling information requirements in dynamic environments, without requiring extensive training or a technical background (e.g. in formal query languages or database schemas). Apple Siri[*] and Google Now[†] – currently, two of the most popular natural language interfaces – basically allow users to navigate through apps intended as both data and sources of data using natural language voice-based queries.

To leverage the advantages of such conversational interactions we propose CE-SAM (Controlled English Sensor Assignment to Missions), a system that guides users through refining and satisfying their information needs in the context of Intelligence, Surveillance, and Reconnaissance (ISR) operations. As explained above, the rapidly-increasing availability of sensing assets and other information sources poses substantial challenges to effective ISR resource management. In a coalition context, the problem is even more complex, because assets may be "owned" by different partners. In this paper, we show how CE-SAM allows a user to refine and relate their ISR information needs to pre-existing concepts in an ISR knowledge base, via conversational interaction implemented on a tablet device as a mobile app. The knowledge base is represented using Controlled English (CE) - a form of controlled natural language that is both human-readable and machine processable (i.e. can be used to implement automated reasoning). Users interact with the CE-SAM conversational interface using natural language, which the system converts to CE for feeding-back to the user for confirmation (e.g. to reduce misunderstanding). We show that this process not only allows users to access the assets that can support their mission needs, but also assists them in extending the CE knowledge base with new concepts.

The rest of the paper is organised as follows. In Section 2 we briefly formalize the selection problem which CE-SAM tackles by subdividing it into two subproblems. In Section 3 we focus on the first subproblem which is the information request refinement, describing the conversational approach based on Controlled English which is used by CE-SAM. In Section 4, we describe a concrete use case of CE-SAM, via screenshots of a prototype tablet-interface. Finally, in Section 5 we summarise our contributions and discuss future work.

2. THE SELECTION PROBLEM

In this section we formalize further the problem that CE-SAM is tackling, illustrated as a diagram in Figure 1. As presented in the introduction choosing the right "app" (to access wrapped data and sources) for each of the information needs of the users on the field is a complex task. We call this the *Selection Problem*, which basically

[*] `http://www.apple.com/ios/siri/`

[†] `http://www.google.com/landing/now/`

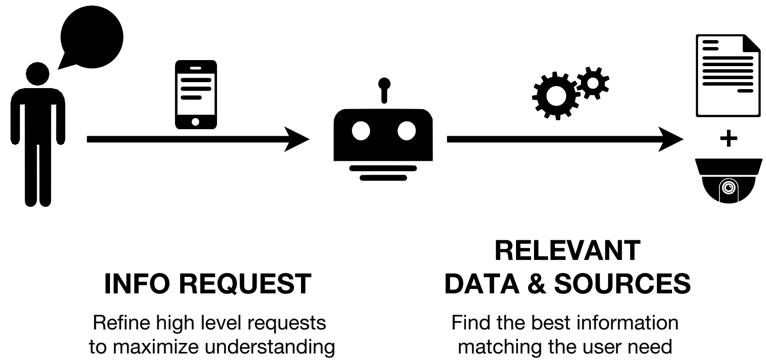

Figure 1. Selection problem: "High level" requests to "low level" data & sources.

consists of parsing, refining and understanding the user's information requests and then finding the relevant data and sources on the field via automatic source/data selection and allocation mechanisms.

In particular, the decision maker on the field may not have fully-articulated the semantic features of their problem (in terms of the kinds of entities and relationships), and they are unlikely to be aware of all candidate information resources in a coalition context, or all the factors involved in selecting the best resources. It is likely a user will need to go through a process of iteratively refining their problem, and matching its features to available ISR resources. The relevance of a resource to some aspect of the users problem is hard to automatically assess, particularly when the problem is stated at a much higher level of abstraction than the capability the app can deliver. Informed choices and accurate decisions are critical in such a dynamic environment and therefore the need to provide usable and human-friendly solutions is imperative.

CE-SAM approaches the selection problem by splitting it into two parts: (1) Information need refinement (2) Matching the need with the appropriate data and sources. Essentially CE-SAM assists with the translation from a "high-level" information request to "low level" needs for data and sources. This enables automatic selection/allocation mechanisms to find the best available resources to satisfy the "low level" need. The challenge is to ensure that the data retrieved and sources selected are relevant to the high level request which the system needs to break down into a set of low level requirements in order for the selection algorithms to "understand" it. In this paper, we focus on the first subproblem which we approach via a conversational interaction and with the use of a common language, Controlled English (CE) as the ontology and data representation format and basic query language. The second subproblem is covered in a companion paper,[2] where we explain how CE can be used to describe "low level" ISR requirements and ISR resources (for particular sensing assets).

The Selection Problem is in general related to the *Multi-Robot Task Allocation* problem, the *sensor selection* problem and the *multi-agent coalition formation* problems, all addressing intelligent resource allocation in different environments. The Multi-Robot Task Allocation (MRTA) problem[3] aims to allocate robots to sensing and actuating tasks in a generic Multi-Robot System (MRS), here instead we focus on ISR sensing assets with no actuator capabilities. The Multi-Sensor Task Allocation (MSTA) problem[4,5] can be considered a constrained version of MRTA, which considers exclusively ISR sensing assets. The Selection problem is basically an extension of MSTA, which considers both ISR sources and data, but extends further the problem scope by focusing also on refining the user information need. It is also related to the general problem of *sensor selection* in which we usually need to choose a set of "active" sensors to achieve a particular objective. This has received considerable attention in the wireless sensor network research community but always with a focus on the allocation mechanisms rather than also considering the user's information need refinement. Examples of sensor selection problem are represented by algorithms for localization and tracking of targets.[6–8] The Selection Problem differs from these also because it considers multiple ISR information needs possibly competing for exclusive usage of the same sensing resources, while instead in sensor selection the focus is usually on supporting single ISR tasks. If we restrict the Selection Problem to only ISR sensing assets, it can be modeled as a *multi-agent coalition formation problem* which has been extensively studied,[9,10] where autonomous entities ("agents") can form a coalition to achieve one common goal (although in this case the refinement of the user information need is not the focus). Finally, the Selection Problem is related to classic information retrieval problems,[11] although those do not consider allocation of sensing resources contended among different ISR information needs.

3. REFINING INFORMATION REQUESTS

In recent years, there has been considerable attention given to conversational user interfaces (CUIs) which allow for complex user interactions with systems, such as fulfilling information requirements in dynamic environments, without requiring extensive training or a technical background (e.g. in formal query languages or database schemas). One of the main advantages of CUIs over Graphical User Interfaces (GUIs) is that they allow the user to talk about hypothetical objects or future events that have no graphical representation[‡]. CUIs are much more flexible essentially allowing users to interact at a much higher level than GUIs, for this reason they are considered to be the next big step in Human-Computer Interfaces.

Apple Siri and Google Now, two of the currently most popular natural language interfaces, allow users to navigate through apps intended as both data and sources of data using natural language voice-based queries. These are personal mobile assistants which help users navigate through app contents, basically trying to enhance the experience and speed of a search into an app-store or amongst apps already installed, with the aim to satisfy a particular user information need. Low training overhead is key to providing problem-focused information delivery in this way and is often referred to as the zero-overhead principle[§]. To improve retrieval of relevant data and sources of data, conversational-based interfaces such as Apple Siri and Google Now take advantage of two main aspects: (a) Natural language query interaction which aims to greatly reduce the need for training and (b) Context-awareness, using location, time and search history as a form of context. In CE-SAM we focus on point (a) since we want first to help the user to refine their information need via a conversation with the system rather than pushing relevant data before the user requests it. This allows the system to retrieve more relevant ISR data and sources, which in the future might also be contextually delivered, e.g. in a form factor or application that is most relevant to the users current activity. Context-awareness can also be used to better understand the request that the user is expressing (e.g. by using the current user location).

Figure 2. Apple's Siri misinterpretation of information need.

There is an important issue in adopting Natural Language (NL) interaction in order to refine the user requests: the input of the user can often be ambiguous due to the intrinsic ambiguities of NL, e.g. the same syntax and words might have completely different meanings. As an example take the case in which a user might ask Apple

[‡]As recently observed in Wired by Dr. Ron Kaplan from Nuance Communications' NLU R&D Lab – http://www.wired.com/opinion/2013/03/conversational-user-interface/ – checked 27th March 2013.

[§]Building For The Enterprise – The Zero Overhead Principle, Dr. DJ Patil – http://techcrunch.com/2012/10/05/building-for-the-enterprise-the-zero-overhead-principle-2/ – checked 27th March 2013.

Siri "How many people are in the square?", as shown in Figure 2. Siri misinterprets this type of information need and believes instead that we are asking for how many people are currently working in a company called "Square". This type of misinterpretation of the input is due both to (1) lack of enough context about the information need but also to (2) different semantics for each of the words and sentences used.

This second point about the intended semantics of the utterance is actually the crucial issue on which we focus. Note that Siri is to a certain extent aware that "square" might have different meanings and for this reason it presents back to the user its input interpretation[¶]. In certain cases, Siri might also ask for confirmation of different input interpretation, e.g. when asking to be reminded of something "tomorrow" when midnight has just passed, Siri will ask if by tomorrow we actually mean today. We believe that asking for confirmation of the user input is a good place to start for refining the users information need on the field of a humanitarian or military mission in order to retrieve relevant ISR sources and data.

Although, just demanding for user confirmation might not be enough for many "high level" information requests in which the query might be more complex than a simple question. For example, consider the case in which the user is "looking for intruders", then the system would have to interpret what is an intruder according to the user. Clearly in such a scenario, presenting back the "input interpretation" needs to be more flexible than simply showing what the system believes the user meant (e.g. as Siri does with "people" considering them "employees" in Figure 2). We need to provide an interactive way of refining the information request by allowing the user to confirm and perhaps edit the request. We believe that the use of a *common language* for both the user need and the system internal processes can help in resolving this issue of ensuring that the ISR data and sources retrieved are relevant to the user need. In particular, in CE-SAM, we decided to use *Controlled English* (CE) which is a controlled natural language; a natural language with limited syntax and vocabulary. CE is by construction unambiguous, which means it is machine processable, but it still maintains an expressivity which is closer to natural language than traditional programming languages and therefore it can be read by humans without need for special training. To summarise the two features described above which can help avoid misunderstandings and refine the information needs of non-technical personnel on the field are: (1) a shared unambiguous language for human and machine processing (CE) and (2) a conversational interaction.

3.1 Controlled English

In our previous work we showed how information requirements expressed in Controlled English (CE) can be matched against available ISR sources and data.[2] Through the conversational interaction we help the user translating their high level information need into a set of statements (facts) expressed in CE.

Our approach[2] depends on the existence of a set of ontologies representing elements of the domain of interest, such as available sensing assets and their relationships with the possible information need. These ontologies cover not only objects of interest in the world (like vehicles, people, places, time) but also elements of the intelligence cycle (e.g. disparate sensing assets and ISR information needs). Further details on ontologies associated with the ISR domain are given in one of our previous articles.[12] The set of ontologies is extensible and can be substituted with the most appropriate for each domain. In CE-SAM these ontologies are defined in CE and referred to as conceptual models.

The CE sentences that define the conceptual models take the form of concept and relationship definitions (via "conceptualise" sentences) and the definition of logical inference rules.[2] Once the conceptual models are defined, subsequent assertions can be made according to these concepts and relationships, via "there is a..." sentences. Concepts may be specialisations of other concepts, indicated via "is a" declarations. Relationships may be defined between concepts (for example, a relationship "provides" is defined between the concepts "asset" and "ISR capability"). Below we show the definition of the concept "military tracked vehicle", using the taxonomy of detectables defined in our previous work[2] via a series of relationships with other concepts:

```
conceptualise a ~ military tracked vehicle ~ M
  that is a tracked vehicle.
```

[¶]The underlying engine for this type of semantic query is Wolfram Alpha `http://www.wolframalpha.com`

```
conceptualise a ~ tracked vehicle ~ T
  that is a ground platform.

conceptualise a ~ ground platform ~ G
  that is a platform type.
```

As described in our previous work,[2] via semantic matching – using a combination of "there is a" sentences and ISR taxonomies of detectables/assets – CE-SAM determines what kinds of ISR sensing capabilities and NIIRS ratings[‖] are required to identify a military tracked vehicle. We then automate the assignment of sensing assets to tasks using a knowledge base derived from the NIIRS method of rating imagery data.[2] Below, we show how specific NIIRS ratings are associated with ISR assets in order to match those against subsequent ISR requirements. This is achieved via defining "there is a" sentences which define "intelligence clauses":

```
there is an intelligence clause named 'ic010' that
  fulfills the intelligence capability 'detect' and
  is looking for the detectable thing 'tracked vehicle' and
  is capable of the sensing capability 'visible sensing' and
  is rated as the NIIRS rating 'visible NIIRS rating 4'.
```

Following the NIIRS approach, we allow ISR requests (which we call intelligence tasks or simply tasks) to be defined in terms of basic capabilities such as detect, identify, and distinguish, and one or more kinds of object of interest which we call detectables. A task is defined in the ISR model in CE as follows:

```
conceptualise the task T
  ~ requires ~ the intelligence capability IC and
  ~ is looking for ~ the detectable thing DT and
  ~ operates in ~ the spatial area SA and
  ~ operates during ~ the time period TP and
  ~ is ranked with ~ the task priority PR.
```

Finally, a user's information need can be represented by instantiating the above model for a specific requirement and ISR assets are suggested with the appropriate sensing capabilities via semantic matching:

```
there is a task named t327893 that
  requires the intelligence capability detect and
  is looking for military tracked vehicle and
  operates in the spatial area 'current location' and
  operates during 'the next 3 hours' and
  is ranked with the task priority high.
```

Note that according to U.S. and NATO military doctrine,[13] ISR requests are to be made based on "desired effects, operational objectives, and end states" which is consistent with the above CE task description. Moreover, the structure of CE tasks is consistent with doctrine since doctrine can be captured in the underlying CE models.

However, even considering that the above machine-processable statements are human readable – and as a consequence, they are in an appropriate format for being presented back to the user for confirmation – for a non-technical user CE sentences are definitely hard to write. Mainly because they involve learning overhead in particular in terms of understanding the underpinning ISR model (represented in CE), what sentences are allowed, and what are the variable values.

‖The National Imagery Interpretability Rating Scale (NIIRS) associates a value to each pair of detectable thing and intelligence capability required; NIIRS ratings are also associated to a combination of sensing platform types and compatible sensing assets, which then enables the semantic matching process.[2]

3.2 Conversational Interaction

In this section we discuss how the conversational interaction can assist a user to refine their information need from high level natural language to low level CE task sentences, also enabling users to extend the knowledge base of CE-SAM in order to deal with previously unseen ISR information requests when required.

As specified previously CE-SAM approaches the selection problem by first helping to refine the user information need and making sure it understood correctly the request. As seen already the refined input is in the form of CE which then allows the system to retrieve relevant data and sources which are also wrapped in CE (using semantic matching[2,12]). This is summarised in Figure 3. Focusing on the first subproblem, CE-SAM helps to translate the information need from "high level" request to "low level" ISR requirements via the use of two mechanisms: (1) Transformation of natural language to CE and (2) Presenting back Controlled English for the refined information need, to allow user confirmation.

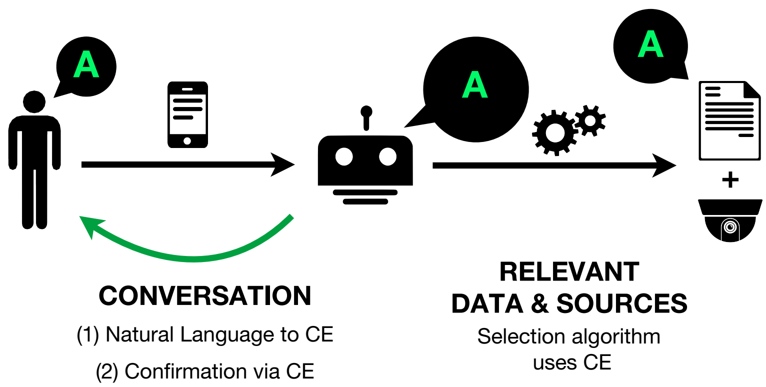

Figure 3. CE-SAM approach to solving the selection problem.

Typical users do not have software development experience, making code writing and in particular stating information requirements derived from NL in CE a challenge. The divide between users and developers has made maintaining and updating knowledge bases in health care too resource intensive in the past.[14] Because Controlled English (CE) is a Domain Specific Language (DSL), it reduces some of the time, cost, and difficulties associated with software development. However, information defined using CE alone is still symbolic, making it still abstract to non-technical users. To close the gap between non-technical users and developers or knowledge engineers, we implement a conversational natural language interface which creates the CE task description representing the user information need. In particular, the natural language interface allows the process to be structured into steps, allowing users to define new information needs without the need for a developer. Each step is comprised of a question with a set of concrete responses. Hence, the interface provides interactive guidance by "walking" users through the process. After the conversation is complete, CE-SAM provides feedback in the form of a human readable representation of the user information requirement in CE, taking advantage of the easy-to-read nature whilst mitigating the harder-to-write issue.

The user verifies the conversation, making modifications to the formal CE if necessary, and then confirms their intent directly in CE, taking advantage of both its human readability and machine processability. System transparency (appropriate feedback, concrete terminology, organization/structure, and direct interaction) is key to human understanding and usability.[15,16] In addition, the simplicity of the conversational approach of CE-SAM is crucial for user acceptance because highly complex systems often lack transparency and may be too cumbersome for integration within typical workflows.[17]

The approach taken to implement the conversational interface is based directly on the use of Controlled English[18] to define the ISR model and associated information as described in our previous work.[2] This CE representation already uses natural English words to define the concepts, properties and relationships within the model and is a good resource on which to base the machine processing of the natural language conversation "fragments" given by the user during the conversation. In CE, these semantic features are defined using normal case (rather than CamelCase as might be found in more formal representation languages) and have no limits on length or form, meaning that the phrases used in the model are already likely to be a good match to some of the natural language used to describe the domain. In addition to this inherent applicability to the task the CE model(s) can be easily augmented with additional lexical information to insert additional phrases and words that are commonly used in natural language to define each of the concepts, properties and relationships in the model. These can be asserted both in the model, and against the instances and can be built up to form a rich inter-linked "bank" of additional terms that are used in natural language. The richer the set of terms the better the accuracy of the natural language processing of the conversation. It is important to note that this approach can easily introduce ambiguity through the use of similar terms in different parts of the model, and any implementation must be able to handle the inherent "fuzziness" of the results, unlike the handling of CE, which is always precise and unambiguous.

The implementation is composed of a set of simple CE agents that take each natural language conversational fragment specified by the user and parse it using a simple word-based analysis mechanism, looking to separate "question words" (and common conversational structures such as interjections) away from other words that are assumed to be part of the language of the domain model in question. The CE model and instances are then scanned to see which of the words in the conversation fragment can be matched to semantic features in the model. The response back to the user is dependent on the interpretation of the fragment and may contain for example the answer to a question, but in all cases the explanation of the interpretation of the fragment is also provided. This explanation takes the form of a mapping from the natural language words to corresponding CE sentences. Since CE is unambiguous and easier to read than it is to write we think that this feedback in CE is a powerful mechanism to capture and communicate the interpreted meaning of the fragment in a convenient and familiar form.

Within this overall pattern there is also the capability to define any number of fairly static "guided conversations" which are defined in CE and specify a multi-stage information collection workflow that can be used by the machine to guide the user through a particular activity. This is the technique that is used in the example given in this paper, enabling the user to specify their complex ISR task requirement via the conversational interface.

The final variation for conversation handling is the most "open" of the three and is useful when handling general statements from a human user about the state of the world, for example when gathering intelligence information. In this situation the fragments specified by the human user are sent to a lexical parsing tool (e.g. the Stanford Parser[19]) for lexical analysis, revealing the linguistic structure of the sentence. This analysis is then passed on to the more traditional information extraction processing that is being researched in the CE environment,[20] thereby treating each conversation fragment as a source of field intelligence in the same way as existing documents or other corpora would be processed. This mechanism offers a very useful alternative mechanism for obtaining meaningful and task-relevant information from a human user in natural language with little or no training overhead.

All three of these techniques can be used to support a conversation with a non-technical human user, with the appropriate technique being used depending on the aim and context of each discussion. This allows questions to be asked or additional information relevant to the task to be specified by the user and can be applied to any CE conceptual models that have been augmented with lexical information as described.

4. CE-SAM ILLUSTRATIVE WALKTHROUGH

In this section we describe, as an illustration, a use case of the current CE-SAM research-grade prototype. In particular we focus on how CE-SAM helps users refine their ISR information need and aids them in the creation of CE-based requirements to capture this. A video of the demo which relates to the detailed description in this section can be found at the following URL: `http://youtu.be/W3ea5DeKAbQ`.

In figure 4 we show the main tablet app interface which is composed of a map displaying the current information requests on the field which we call "sensing tasks" or more simply "tasks"; these might have been requested by the same user or by members of the coalition team which the user belongs to. Tasks are essentially a way of representing on the map ISR information needs that persist over time (since the data gathered by a task might be needed for a certain duration). In order to satisfy each task, we might in fact need to allocate both ISR sensing assets or retrieve ISR collected data. Note that in this walkthrough we limit our attention to ISR sensing assets. It should be noted that our approach is extensible and can also easily include "soft-data[21]", that is data which is not produced directly by ISR "hard" sensing assets but which is instead generated via other means, e.g. by humans, like patrol reports or eyewitness observations in the form of tweets or mobile texts and requires some complex parsing to determine probable intended meaning.

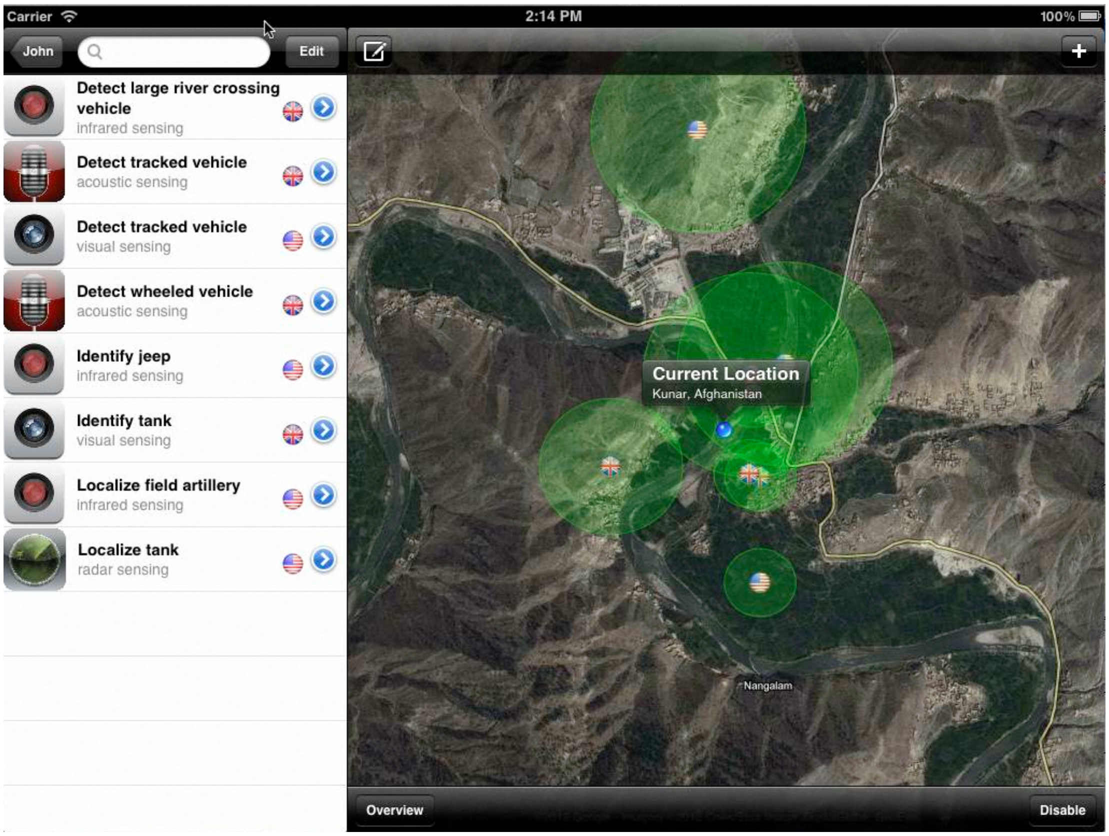

Figure 4. CE-SAM main view: geolocated ISR information needs.

The task list is searchable using the box on the top-left allowing users to filter the displayed list by intelligence capability or type of detectable thing. Tasks are summarised with an icon that provides a quick indication of the kind of sensing capability assigned (for example, by means of acoustic, visual, infrared, seismic, or radar sensing).

We now go into more detail for a particular ongoing task: "Detect wheeled vehicle". This task is expressed at high level as the user specified "what" they wanted to know, rather than which ISR assets are required to satisfy this information need. A technical user could express the same information need by instead stating the type of assets needed: e.g. "mica 7 motes with acoustic sensors". Such a request is equally precise in terms of capabilities required but by stating the capabilities rather than the requirements the system cannot take into account the rest of information needs which might compete for the same resources and lacks of flexibility. In other words, a user might end up getting a negative response from the system only because the system does not know "what" to look for and therefore cannot find alternative ISR sensing assets or data in order to achieve the same goal.

The conversational interface in Fig. 5 could provide great help to assist the user in the refinement of their information need from a higher level than the one provided by the NIIRS taxonomy[2] (i.e. simple pairs of intelligence capability and detectable needed, e.g. "detect wheeled vehicle"). This is achieved by translating the higher level information need to a CE-based task representation which the system can then use in order to match the request with ISR data and assets,[2] as discussed above. Furthermore, the conversational interface allows the provision of new knowledge into the CE-SAM knowledge base during the conversation, allowing a richer and more realistic interaction, enabling the user to know which resources are best fit for a new information need.

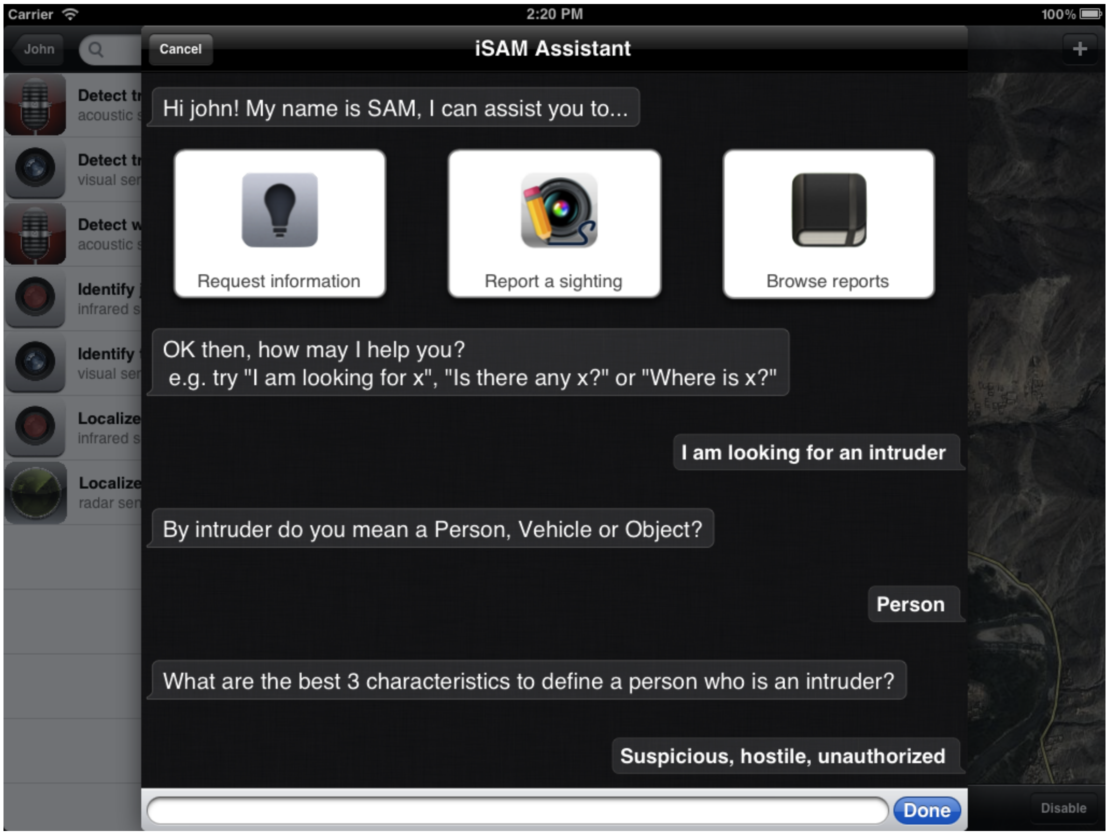

Figure 5. CE-SAM conversational interface: refinement of the information need.

In Figure 5, we consider the example of a user who is looking for intruders: in our walkthrough scenario neither the detectable "intruder", nor the intelligence capability "looking for" are present in the CE-SAM knowledge base so the statement cannot be made via a simple pre-defined form-based request. For this reason a conversational interface is brought up by the user. The conversational interface starts by requesting the information need of the user at high-level and by hinting at possible requests that the system is able to understand (in our example, "I am looking for x"). As explained in Section 3.2, when the user inputs "I am looking for intruders", CE-SAM parses the sentence and it maps "I am looking for" to the intelligence capability "detect". After that it expects a detectable but since "intruder" is not in its knowledge base CE-SAM tries to link that, supposedly, new detectable to pre-existing detectable concepts already known to the system. This is performed in order to match the new detectable with ISR assets that are able to satisfy information needs for known detectables. In this case the system asks if for "intruder" the user meant a person, vehicle or object, trying to link at different levels of its taxonomy (i.e. classes of objects) the new detectable. Finally, once the user has specified that "intruders" belongs to the class "Person", CE-SAM now associates it to that detectable and therefore in the worst case it can match the request with ISR sensing assets able to "detect a person".

Now the problem lies in how CE-SAM can distinguish normal people from intruders and deliver pertinent ISR information or assign appropriate ISR assets. As an example in our scenario this could be achieved – with various levels of accuracy – by asking the user for the best three characteristics which defines the specialisation of the person concept known as "intruder". In this case the adjectives provided by the user could be used to match information coming in from textual reports from both patrols or eyewitnesses. Alternatively we can ask a user to associate in the CE-SAM knowledge base different sequences of actions of a person which determines it to be suspicious. This is not shown in this paper but, for example, the user might state that frequent passages of the same person might classify the person as suspicious. This style of conversation is part of our current ongoing work which has the aim of understanding the degree to which we can deal with a generic information requirement in NL yet still being able to translate it into an understandable CE form for the system and to link it to pre-existent objects in the knowledge base. This brief example is to illustrate how the conversational interface helps the user to specify new knowledge in their NL without needing technical knowledge about CE or the CE model used.

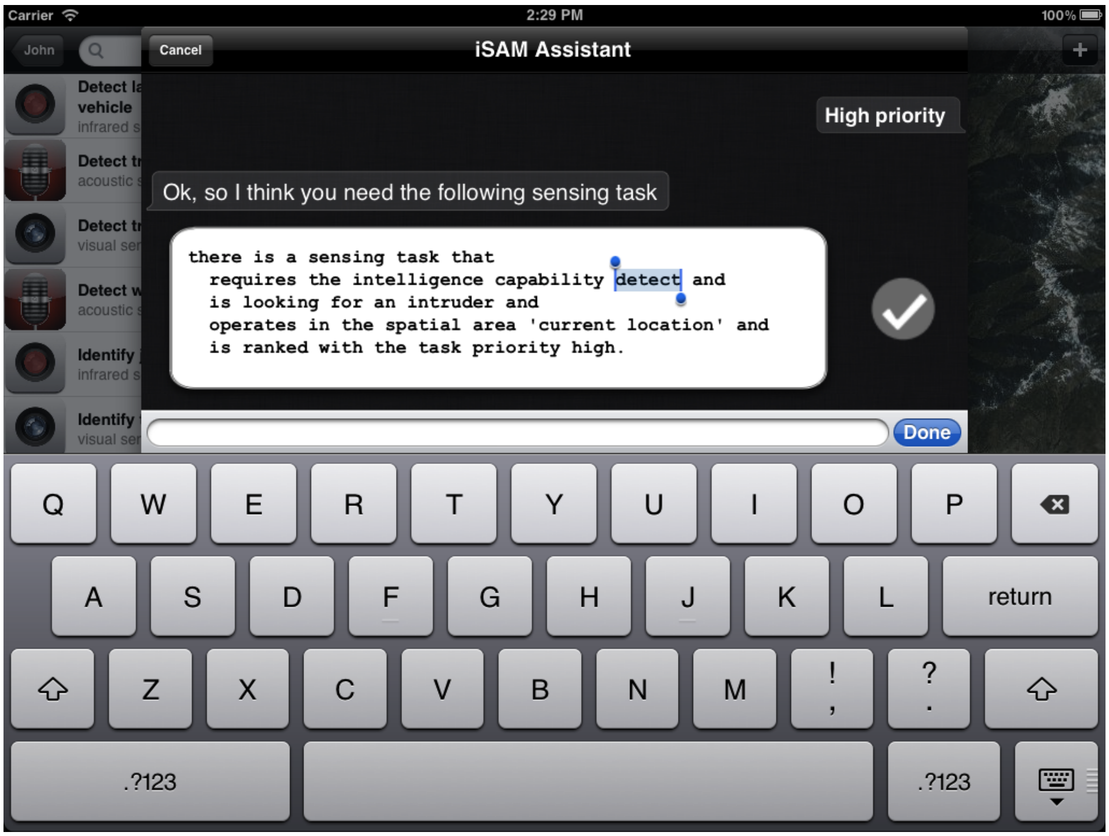

Figure 6. CE-summarisation of ISR information need, for user confirmation.

Following the steps above, CE-SAM continues by asking for additional details about the ISR request via the conversation. In this case, the duration for which this ISR request will be needed by the user and the priority or importance of this information need, both of which are then used to select the best data and sources of data.[2] After that the system concludes it has gathered enough information about the information need, CE-SAM presents back to the user the refined request summarised in Controlled English (CE) as represented in Figure 6. This CE task summary was generated by the system through the conversational interaction. By presenting the representation of the NL user request back to the user in CE, we aim at decreasing possible misunderstandings between the system and the user. This is, in fact, the same CE representation of the ISR task that the system will use internally in order to carry on the matching between request and ISR assets.[2]

CE-SAM allows the users to optionally edit the information need directly in CE if it was not correctly interpreted by the system. For example the user can edit the intelligence capability that the system mapped to the user need, in our case instead of "detect" (by saying "I am looking for") the user might have wanted to "localise" (thus track) an intruder. This is achieved by simply modifying the CE sentence to reflect the changes and does not require a re-run of the conversation. Of course the issue here is to ensure that the edits keep the consistency of the CE syntax. In order to achieve this we constrain the users in the edits that they can make, and we plan on assisting them further with autocompletion suggestions when editing pieces of the CE description so that they adhere to the underlying CE-model without requiring specifying training to do so. As an example, in the case of intelligence capability we can provide them with the "n" different intelligence capabilities which the knowledge base of the system contains (e.g. detect, localise and identify). Once the user is happy with the CE description of their information need, they can confirm the choice by tapping on the confirmation check button. For a more detailed overview of CE and its capabilities and future directions we refer to the latest paper describing such technology.[22]

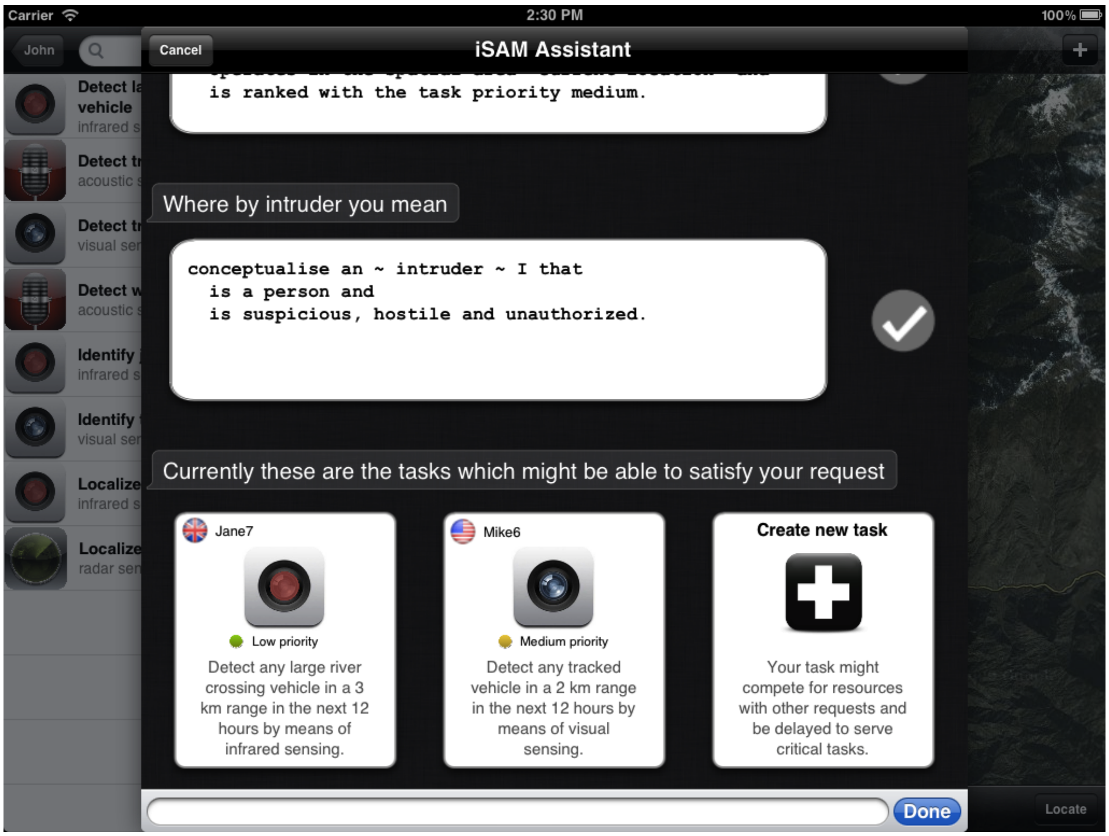

Figure 7. CE-representation of the new "intruder" concept created in the knowledge base.

In Figure 7, the system presents back to the user the additional data about the information need which it has learnt via the conversational interface. In particular, here CE-SAM presents the new detectable thing that will be added to its knowledge base and that it will use to complete the meaning of the previous CE task description. In this case, the system presents back to the user the relationship of new concept "intruder" to the concept person which is already in the knowledge base (and for which it knows which ISR sensing assets might satisfy the information need). The new concept contains also the stated specialisation characteristics which might help to distinguish an "intruder" from a normal "person". Note again that these might be used in the system to do text mining in textual reports or could have associated some routines that will be able to identify a person who is an intruder for example via image processing. Again the user is allowed to edit the description in CE of the new concept.

Once the user has confirmed the interpreted information is correct, CE-SAM suggests to the user the current tasks that could be used to satisfy the information need of the user for the task just described. The system, in fact, now can suggest existing ISR tasks that are currently using ISR assets able to satisfy the same information need just stated by the user. Alternatively, the system can also suggest ISR tasks which are similar to the one specified by the user in terms of similar detectables or similar intelligence capabilities – thanks to the linkage of the new knowledge to previous knowledge in the ontology. This is helpful in terms of information reuse; that is, other users belonging to the same coalition might decide to subscribe to the same information need rather than creating a new one. This clearly improves the situation in terms of competition for ISR sensing assets, in fact by subscribing to a pre-existent ongoing task the users will avoid the creation of new requests which might compete for the same resources thereby increasing the number of overall unsatisfied information needs.[4]

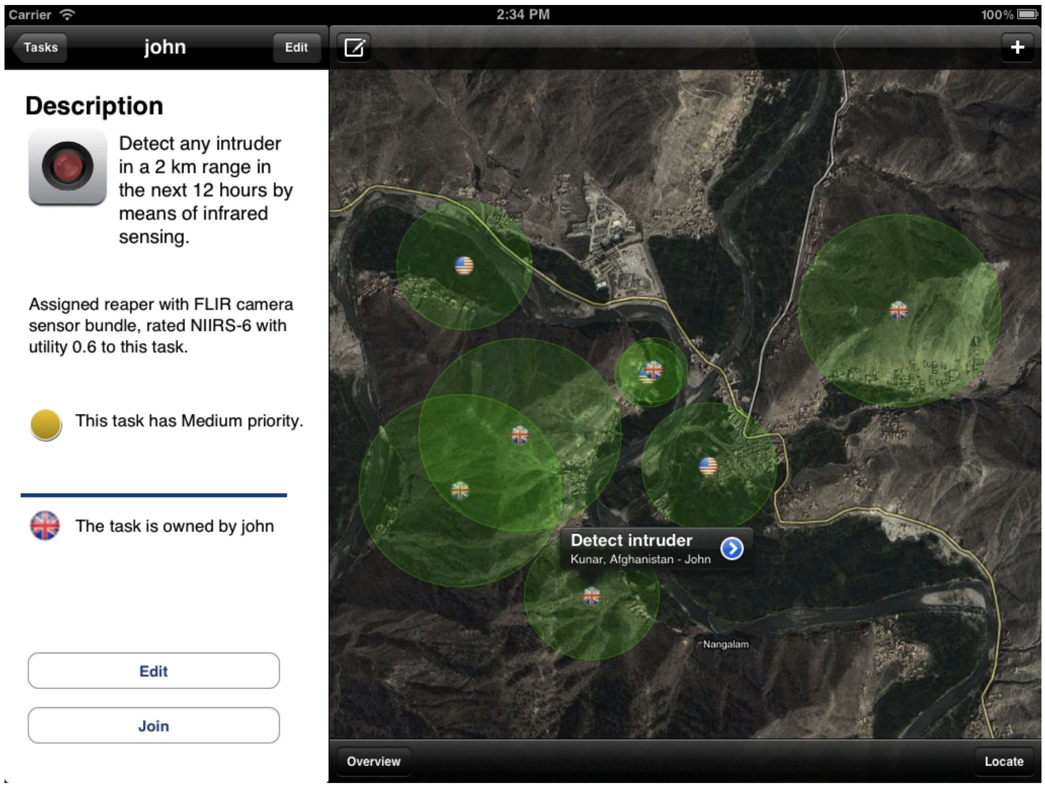

Figure 8. ISR need details expressed in Natural Language via CE-based description.

If the user decides instead to create a new sensing task (i.e. submit a request for resources in order to satisfy their information need), CE-SAM will decide if there are enough resources to satisfy that new task by taking into account other tasks competing for the same ISR assets. This is achieved by using allocation algorithms[4,23] and a semantic matchmaking process[2] as previously mentioned.

This scenario has explained how the making of ISR requests is being successfully supported by this system. In Figure 8 we show the description of the corresponding ISR task in natural language, together with other parameters which identify more in detail the sources collecting the data. This NL summary was generated from the CE task description and the CE representation of ISR assets which the system uses in order to perform the allocation as explained in our previous research.[2,4] In this particular example the task was assigned a reaper with a FLIR camera sensor. In addition, we provide details of the priority of the task and the level of confidence that the ISR assets allocated to the task by CE-SAM manage to provide. Note that the level of confidence is defined by two measures. The first is the utility provided by the ISR assets to the specific task; this goes from 0 to 1 and can be calculated via specific utility functions based e.g. on the distance of the ISR assets and degradation of signal.[1,4] The second measurement is the NIIRS value, which CE-SAM uses in oder to match resource types with the appropriate ISR information needs.[2,12]

In addition, the new concept created via the conversational interaction (intruder) could also automatically be shared by the system with other users in the same coalition. Different strategies and change control mechanisms can be applied in order to maximize knowledge usefulness and minimise possible semantic-divergences in each use case. For example, we might decide to formally include a new concept (e.g. "intruder") into the CE-SAM knowledge base by using a frequency-based approach: i.e. by presenting by default this new term among the higher-level detectable things only after a number of users have refined their information needs towards the same term, by using similar definitions of "intruder" during the conversational interactions.

5. CONCLUSION & FUTURE WORK

In this paper, we have proposed CE-SAM, a system based on natural language conversation interaction to help users refining and satisfying information needs in the context of Intelligence, Surveillance, and Reconnaissance (ISR) operations. The increasing availability of sensing assets and other information sources in dynamic environments requires effective ISR resource management. The conversational mode allows the resting of ISR requests to pre-existing concepts in an ISR knowledge base represented in Controlled English (CE). Users interact with CE-SAM using natural language, which is converted to CE for user confirmation and subsequent automated reasoning. CE-SAM helps to identify relevant ISR assets and assists to fulfil the stated information requirement and allows extension of the CE knowledge base with new concepts.

In terms of future work we plan to explore how "high level" a user can go when expressing their information need but still allowing the system-guided conversation to translate the request into "low level" CE-based ISR needs. We plan to achieve this with a combination of the current confirmation mechanism, via which the system asks for user approval presenting back the information request translated in CE, together with the three conversation parsing/guiding techniques described above. One of the main challenges to overcome is to ensure that while creating new concepts in the KB of the system, via the user-machine conversation, ontology consistency is maintained; we plan to explore this as part of our future work and we believe a good starting point is to explore available techniques in the wide literature for ontology editing via NL and CNLs.[24] We also plan to evaluate the CE-SAM tool with human-subjects, conducting experiments on how helpful such a system could be for actual users, in specific compared to the current practices for selecting ISR resources for ISR military missions.

Finally, CE-SAM could also act as a tool to formalize access and sharing policies among members within a coalition. We aim at developing a CE-based negotiation mechanism, which relaxes policies constraints through a conversational interaction with the user. We believe, in fact, that expressing policies via CE at a high-level allows non-technical users at or near the organization edge to negotiate access to resources at a horizontal level instead of going up the chain of command making the coalition collaboration more agile.

ACKNOWLEDGMENTS

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 autho- rized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. We acknowledge Cheryl Giammanco from the U.S. Army Research Laboratory for her support in the project, her feedback on the user interface and interactions and her insightful comments on this paper.

REFERENCES

[1] H. Rowaihy, M. P. Johnson, D. Pizzocaro, A. Bar-Noy, L. Kaplan, T. La Porta, and A. Preece, "Detection and localization sensor assignment with exact and fuzzy locations," in *Proceedings of the 5th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS 2009)*, pp. 28–43, Springer-Verlag, (Berlin, Heidelberg), June 2009.

[2] A. Preece, D. Pizzocaro, D. Braines, and D. Mott, "Tasking and sharing sensing assets using controlled natural language," *Proceedings of SPIE Defense, Secuirty and Sensing* **8389**, pp. 838905–14, 2012.

[3] B. Gerkey and M. Mataric, "A Formal Analysis and Taxonomy of Task Allocation in Multi-Robot Systems," *The International Journal of Robotics Research* **23**(9), p. 939, 2004.

[4] D. Pizzocaro, A. Preece, F. Chen, T. La Porta, and A. Bar-Noy, "A distributed architecture for heterogeneous multi sensor-task allocation," in *2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS 2011)*, pp. 1–8, IEEE, June 2011.

[5] D. Pizzocaro, "Instantaneous multi-sensor task allocation in static and dynamic environments." http://orca.cf.ac.uk/31333/, 2011.

[6] L. Kaplan, "Global node selection for localization in a distributed sensor network," *IEEE Transactions on Aerospace and Electronic Systems* **42**, pp. 113–135, January 2006.

[7] L. Kaplan, "Local node selection for localization in a distributed sensor network," *IEEE Transactions on Aerospace and Electronic Systems* **42**, pp. 136–146, January 2006.

[8] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Processing Magazine* **19**, pp. 61–72, March 2002.

[9] T. W. Sandhlom and V. R. T. Lesser, "Coalitions among computationally bounded agents," *Artificial Intelligence* **94**, pp. 99–137, July 1997.

[10] O. Shehory and S. Kraus, "Methods for task allocation via agent coalition formation," *Artificial Intelligence* **101**, pp. 165–200, May 1998.

[11] E. Greengrass, *Information Retrieval: A Survey*, 2000.

[12] M. Gomez, A. D. Preece, M. P. Johnson, G. de Mel, W. Vasconcelos, C. Gibson, A. Bar-Noy, K. Borowiecki, T. F. La Porta, D. Pizzocaro, H. Rowaihy, G. Pearson, and T. Pham, "An ontology-centric approach to sensor-mission assignment.," in *EKAW'08*, *Lecture Notes in Computer Science* **5268**, pp. 347–363, Springer, 2008.

[13] J. C. of Staff, "Joint and national intelligence support to military operations," tech. rep., Joint Publication 2-01, January 2012.

[14] H. Heathfield, "The rise and fall of expert systems in medicine," *Expert Systems* **16**(3), pp. 183–188, 1999.

[15] D. A. Norman, *The Design of Everyday Things*, Basic Books, first edition ed., Sept. 2002.

[16] C. D. Wickens, S. E. Gordon, and Y. Liu, *An Introduction to Human Factors Engineering*, Longman, US ed ed., Dec. 1997.

[17] O. Wegwarth, W. Gaissmaier, and G. Gigerenzer, "Smart strategies for doctors and doctors-in-training: heuristics in medicine," *Medical education* **43**, pp. 721–728, Aug. 2009. PMID: 19573016.

[18] D. Mott, "Summary of controlled english," *https://www.usukita.org/node/1411* , May 2010.

[19] "The stanford parser, a statistical parser," *http://nlp.stanford.edu/software/lex-parser.shtml* .

[20] P. Xue, D. Mott, D. Braines, S. Poteet, A. Kao, C. Giammanco, T. Pham, and R. McGowan, "Information extraction using controlled english to support knowledge-sharing and decision-making," *17th ICCRTS – Operationalizing C2 Agility* , June 2012.

[21] A. Preece, D. Pizzocaro, D. Braines, D. Mott, G. de Mel, and T. Pham, "Integrating hard and soft information sources for D2D using controlled natural language," in *2012 15th International Conference on Information Fusion (FUSION)*, pp. 1330–1337, 2012.

[22] D. Braines, D. Mott, S. Laws, G. de Mel, and T. Pham, "Controlled english to facilitate human/machine analytical processing," *Proceedings of SPIE Defense, Secuirty and Sensing* **8758**, pp. 8758–13, 2013.

[23] F. Chen, T. La Porta, D. Pizzocaro, A. Preece, and M. B. Srivastava, "A system architecture for exploiting mission information requirement and resource allocation," *Proceedings of SPIE Defense, Security and Sensing* **8389**, pp. 838908–9, 2012.

[24] P. R. Smart, "Controlled natural languages and the semantic web." http://eprints.soton.ac.uk/265735/, July 2008.