CogSIMA 2011, Miami Beach, FL

Augmenting the Analyst via Situation-Dependent Reasoning with Trust-Annotated Facts

Brian Ulicny, Gerald M. Powell, Chester F. Brown III, Mieczyslaw M. Kokar, Christopher J. Matheus, Jerzy Letkowski

***Abstract*— We say that a computer program augments the analyst if it can infer facts that are implicit in existing information, but that may be relatively difficult for a human to infer. Among a multitude of reasons, the analyst's task is difficult because (1) reported information to be analyzed and reasoned about often cannot be completely trusted (requiring verification attempts via further collection of information, corroboration where verification is not possible, and/or assumption-based reasoning) and (2) evaluation of trust (reliability, credibility) is context (or situation) dependent. These two sources of difficulty, and the possibility of augmentation-generated false alarms, imply that if one wants to employ the capabilities of an automatic reasoner, the reasoner must be able to deal with these kinds of complexities.**

***Index Terms*—Trust, Reliability, Credibility, STANAG 2022, Situations, Intelligence Augmentation**

I. INTRODUCTION

In this paper we describe scenarios in which the truth-value of a statement depends on both the context and the trust associated with reported information and the source of the information. We will use “trust” here as a shorthand for any of a cluster of several related epistemic notions, including trustworthiness, reliability, credibility, and confidence. We use Situation Theory as developed by Barwise and Perry and expressed in an OWL-based Situation Theory Ontology (STO) previously developed by some of the authors. We also employ the NATO STANAG (Standardization Agreement) 2022 specification for evaluating the level of trust in reported information by means of evaluating information credibility and source reliability [9]. We use an ontological representation of information, sources and trust implemented as an extension of STO. To facilitate reasoning, we incorporate automated policies (rules) for reasoning about trust-annotated information according to STANAG 2022 using a formal inference engine, BaseVISor [10]. We show how one might formally encode and make inferences in a variety of scenarios and explore various issues concerning reasoning with trust-annotated information.

II. SITUATION THEORY AND STANAG 2022

A. Situation Theory

Situation Theory, as initiated by Barwise and Perry [3] and developed by Devlin [4], is a theory of information flow among cognitive agents, particularly by means of language. Barwise and Perry begin with the assertion that people use language to talk about (i.e., exchange information about) limited parts of the world, which they call situations. (For example, scenes are situations that are visually perceived by some observer.) Abstract and concrete situations are partial possible worlds, and the information an agent has about a given situation at any moment is limited to information about elements of the situation.

In situation theory, information about a situation is expressed in terms of infons. Infons are written as

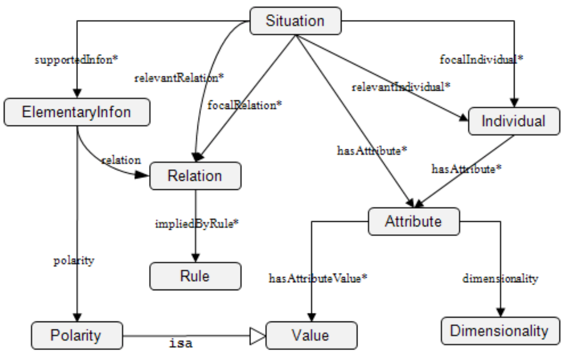
$$\langle\langle R, a_i, \ldots, a_n, 0/1 \rangle\rangle$$

where R is an n-place relation and $a_1, \ldots, a_n$ are objects appropriate for R. Situations support (|=) infons. Infons have slots for time and location parameters, which, by convention, are encoded as the two slots before the final polarity parameter (0 or 1). Each slot is associated with a type of individual that must fill it (e.g. times, locations, persons). A polarity of 1 indicates that the situation contains the described state of affairs. A polarity of 0 indicates the opposite. Infons may be recursively combined to form compound infons.

In [8], a computer-processable semantics for Situation Theory was developed that is compatible both with Barwise and Perry and with Endsley’s model of human situation awareness [11]. To achieve this, Situation Theory was encoded using a formal ontology in OWL (Figure 1). An ontology-based approach to situation awareness supports the inference of new facts about the situation from the encoded facts.

There are three basic types of situations: Focal Situations (what an utterance is about), Utterance Situations (where, when and by whom an utterance takes place) and Resource Situations (other situations that contribute to the Focal Situation).

Figure 1 Situation Theory Ontology (STO)



It has been shown that the OWL ontology encoding Situation Theory can be used to model and track situations as they unfold [8]. The following is an example of the OWL encoding of the following two propositions in OWL Abstract Notation:

```
BlueThreatenRed |= <<warns;Blue;Red;0>>
BlueThreatenRed |=  <<threaten;Blue;Red;1>>

Individual(sto:BlueThreatenRed
   type(sto:Situation)
   value(sto:supportedInfon
    Individual(  type(sto:ElementaryInfon)
                  value(sto:anchor1 ex:Blue)
                  value(sto:relation  ex:warns)
                  value(sto:anchor2 ex:Red)
               value(sto:polarity sto:_0)))
   value(sto:supportedInfon
    Individual( type(sto:ElementaryInfon)
                  value(sto:anchor1 ex:Blue)
                  value(sto:relation  ex:threaten)
                  value(sto:anchor2 ex:Red)
                  value(sto:polarity sto:_1))))
```

Implicit features in Situation Theory notation are made explicit in the STO OWL notation.

B. *Information Evaluation*

The Annex to NATO STANAG (Standard Agreement) 2022 “Intelligence Reports” [9] states that where possible, “an evaluation of each separate item of information included in an intelligence report, and not merely the report as a whole” should be made. It presents an alpha-numeric rating of “confidence” for each piece of information which combines an assessment of the reliability of the source of the information and an assessment of the credibility of a piece of information “when examined in the light of existing knowledge”. The alphabetic Reliability scale ranges from A (Completely Reliable) to E (Unreliable) and F (Reliability Cannot Be Judged). A similar numeric information Credibility scale ranges from 1 (Confirmed by Other Sources) to 5 (Improbable) and 6 (Credibility Cannot Be Judged). See Table 1 for descriptions of the criteria.

III. ENCODING REPORTING SITUATIONS

We extend our STO Ontology to facilitate the formal representation of information reports and the STANAG 2022 metrics. We illustrate this approach with the following scenario: two (human) subjects, A and B, are in a bar, and A is trying to find out the score of a baseball game by asking patrons. Let *Bar* be the name of this situation, at location BarLoc and at time t. *Bar* is a discourse situation and contains utterance situations. When A asks B, “What is the score of the Red Sox game?” we represent this as:

UttSit1 |= <<utters, A, “What is the score of the Red Sox game?”, BarLoc, T, 1>> (1)

UttSit1 is about focal situation *Game*, the referent of A’s singular term “the Red Sox game”. It is not about the immediate environmental situation *Bar*, or any other resource situation (2):

UttSit1 |= << focalSituation, Game, BarLoc, T,1>> (2)

Table 1 NATO STANAG 2022 Rubric (Adapted)

Reliability (Source)	Credibility (Reported Information)
A: Completely reliable. A tried and trusted source that can be depended upon with confidence.	**1: Confirmed by Other Sources.** The reported information originates from another source than already existing info on subject.
B: Usually reliable. Source has been used in the past, but some element of doubt in particular	**2: Probably True.** The independence of the source of any item of information cannot be guaranteed, but based on

cases.	previous reports, its likelihood is regarded as sufficiently established
C: Fairly reliable. Source has occasionally been used in the past; some degree of confidence.	**3: Possibly True.** Insufficient confirmation to establish higher degree of likelihood. A freshly reported item of info that does not conflict with previous reports.
D: Not usually reliable. Source has been used. More often than not has proved to be unreliable.	**4: Doubtful.** An item of information which tends to conflict with the previously reported behavior pattern of intelligence target.
E: Unreliable. A source which has been used in the past and has proved unworthy of any confidence.	**5: Improbable.** An Item of information that contradicts previously reported information or conflicts with the established behavior pattern of intelligence target in a marked degree.
F: Reliability cannot be judged. A source not used in past.	**6: Truth of information cannot be judged.**

Game is a situation of type *BaseballGame*, which we could encode in our domain ontology. It may support many infons, but we are concerned with just the fact that it has a time, place and a score.

In saying that A asks B the score of *Game*, we need to represent the question that A asks B. Following Barwise & Perry, we represent questions as abstract propositions, with a question variable corresponding to the queried role (or to the polarity, for Yes/No questions). A asks B question (Q):

Game |= <<score, ?, Fenway, t, 1>> (Q)

where the ? indicates that it is the score of Game that is being asked for.

In reply, B says (UttSit2), "Red Sox 2 Yankees 0". By saying this, B says of the focal situation *Game* that its score is *Red Sox 2, Yankees 0* because of the discourse context in which it was given: it is interpreted with respect to the Focal Situation. Thus, B asserts Propositon (P):

Game |= <<score, Red Sox 2 Yankees 0, Fenway, t, 1 >> (P)

The form of a Proposition is s |= σ. A proposition conveys that a particular state of affairs σ is made true (polarity 1), or not made true (0), by a situation s.

As a stranger to A, B has reliability *F: Reliability Cannot Be Judged* because B has no previous track record of true assertions.

Since B's report is a "freshly reported item of information that does not conflict with previously reported behavior", B's report (P) has credibility *3: Possibly True*. Future reports of the same score by independent informants would upgrade the assigned credibility. Inconsistent reports about the same situation would downgrade the assigned credibility.

Since the STANAG 2022 Credibility rubric involves independent assertion of the same proposition, we can implement rules such as CredIncrement to aggregate assertions of the same proposition by independent sources:

If Proposition ?p hasCredibility 3,
  and source ?s1 asserts ?p,
  and source ?s2 asserts ?p,
  and ?s1 owl:differentFrom ?s2,
  and ?s1 independentOf ?s2,
then
  Proposition ?p hasCredibility 2.
  (CredIncrement) (? Indicates variable)

Similar rules can be implemented to upgrade or downgrade the Reliability of a speaker based on the ratio of verified propositions they assert to the total number of evaluated propositions they assert.

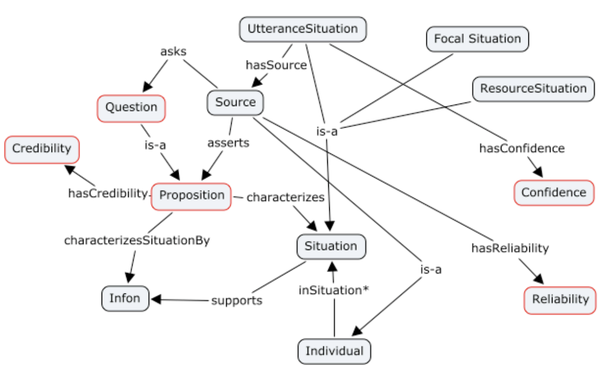
STANAG 2022 says that we combine the alphabetic reliability score with the numeric credibility score to form the alpha-numeric Confidence metric (5). We say that *UttSit2* has Confidence F3 corresponding to the reliability of its speaker, F, and the credibility of its content, 3.

The STANAG 2022 rubric specifies that the reliability of any informant not previously encountered must not be judged, but this is obviously unrealistic. Humans constantly judge strangers to be reliable based on background clues like their facial expressions [19], clothes, manner, and social role (clergy, uniformed police, fellow Red Sox fan). Clearly, these introduce cognitive heuristics and biases that have been shown to be adaptive or problematic for decision-making depending on the situation [20,21]; they cannot be assumed away in realistic situations. Interpretations are known to be influenced by expectations. This is part of the context-dependence of interpretation [16]. Various cognitive, emotional and other factors have been identified as relevant to judgments of trust and trustworthiness [22]. On the Web, unfamiliar authoritative websites are ranked by network centrality metrics, (e.g. Google PageRank [15]).

NATO and Army doctrine provide no methodology for establishing or reasoning about source (in)dependence in any systematic way. Without provenance[17], it may be very difficult for the analyst to determine whether information sources are independent. Becker and Corkill [13] show that multi-agent systems are very error-sensitive to incorrect confidence-integration based on simplified assumptions about source independence.

Figure 3 shows extensions to the STO Ontology made in order to facilitate reasoning about informant reliability, information credibility, and report confidence. The new classes are outlined in red. The STO ontology is simplified here for presentation purposes. Credibility is associated with Propositions, rather than utterances because the same utterance can express different propositions in different contexts (situations). Conversely, different utterances can express the same proposition in different utterance contexts. Propositions express the meaning of what has been asserted about a (focal) situation, and it is they that are true or false. Confidence is associated with an Utterance Situation since it is related to both the speaker and the proposition expressed.[1]

Figure 2 STANAG 2022-extended STO Ontology



IV. STANAG 2022 Reasoning in Context

The US Army and NATO doctrine require evaluating all pieces of intelligence as described. However, the Army does not prescribe how to combine information across evaluated information to generate confidence (reliability, credibility) ratings for information from multiple reports (conflicting or not), or within a given document. There could, in principle, be highly divergent confidence ratings dependent on the analysts' methods and dependent on the context of the ratings.

A. Bayesian Approaches

Bayesian methods calculate the increase or decrease in the probability of an event given the prior probabilities of contributing events. Bovens and Hartmann [5] provide a Bayesian account of the confidence which one should attribute to a conjunction of reports produced independently by sources of varying reliability. For example, the posterior probability that a fair coin toss is Heads, given that three informants, each with reliability of 80%, report it as Heads, is 0.985. That is, the posterior probability of three identical reports by informants each with individual reliability 0.8 is more likely to be true than a single report by an informant with reliability 0.98 by Bayesian reasoning. Bayesian methods predict subjects' assignments of subjective certainty in informal, non-mathematical settings [1].

1.Discussion. In order to calculate the Bayesian posterior probability for n reports, it is necessary to know 2^n joint and disjoint prior probabilities: for example, the prior probabilities of: "The subject is wearing a hat", "The subject is driving a white Ford", "The subject is driving east to Samarra", and their joint probabilities. This is what makes Bayesian approaches infeasible in practice. On the battlefield, priors are typically lacking. There are many challenges, including context-dependence, involved in practical and accurate methods for collecting, storing and distributing this information.

B. The Weighted-Majority Algorithm

Perhaps because of the difficulty of obtaining prior probabilities, Laurent Cholvy proposes ([6], [7]) a formal method for integrating potentially conflicting information into a fact-base in a non-Bayesian way, where that information has been evaluated for reliability and credibility.

In Cholvy's Weighted-Majority algorithm, a set of evaluated propositions is to be merged into a factbase. First, sets of propositions are identified that are consistent with a set of integrity constraints. Then the reliabilities of reports contrary to each consistent set of propositions are aggregated. That is, if one of the consistent sets is {a, -b}, then the

[1] OWL ontologies corresponding to this situation are available at http://vistology.com/ont/2010/SIMA/

reliability of independent reports of –a and b are summed. This total is based on assigning weight 5 to *A: Completely Reliable* informants, down to weight 1 to informants judged *E: Unreliable*. The set of consistent statements with the least total reliability assigned to contrary observations is then taken to be factual and added to the fact base.

For example, if there are two sources and one report the state of a coin as Heads and the other Tails, then, since they both cannot be right, one should accept the observation of the observer with the higher reliability, or, equivalently, the report that is contradicted by the observer(s) with the lower reliability.

Cholvy provides the following, more complex, example [6]: suppose Observer 1 has reliability A (weight 5); Observer 2 has reliability B (weight 4); Observer 3 has reliability C (weight 3), and Observer 4 has reliability D (weight 2).

Observer 1 reports b: an observed object is a helicopter.

Observer 2 reports a: the object is a plane

Observer 3 reports a: the object is a plane

Observer 4 reports c: object has altitude < 3km

Observer 1 report c: object has altitude < 3km

The integrity constraints at work here are: it is impossible that both a and b are true, and it is impossible that both a and c are true. This means that the consistent sets of propositions (models) are: {a, -b, -c}, {-a, b, c}, {-a, b, -c}, {-a, -b, c}, and {-a, -b, -c}.

The combined reliability of assertions contrary to these sets of propositions are: 12, 7, 14, 12, 19, respectively. Since the combined reliability of observations inconsistent with {-a, b, c} (i.e., the combined reliability of observations by Observers 3 and 4 (= 4 + 3)) is the least out of all five sets, then the set of propositions {-a, b, c} is taken to be the best integration of the reports.

1.Discussion. Cholvy does not indicate when these calculations should be performed. If it is after every report, then every report will be factual, since there is no reliable evidence to the contrary. On the other hand, it might make a great deal of difference whether one performs this calculation after *n* reports have been received, or *2n* reports on the same topic. The result may be quite different.

In addition, it should be observed that it takes three consistent sources of reliability .8 to outweigh a single report of a fair coin toss by an informant of reliability 0.98 in the Bayesian account, but on Cholvy's account, it only takes three reports by informants considered *D: Not Usually Reliable* to outweigh an informant considered *A: Completely Reliable*. Moreover, a single *A: Completely Reliable* report is outweighed by six independent reports from *E:Unreliable* informants, on Cholvy's account, no matter what the prior probabilities are. While Cholvy does not associate reliability measures with probabilities, these observations provoke skepticism.

C. *Informant Interests*

Cholvy's integration method breaks down, however, in certain identifiable conditions of self-interested reporting.

Consider this situation (adapted from [12]). An absentminded Professor has forgotten to record the grades for three exams he has already returned. For whatever reason (e.g, dormitory fire, overzealous recycling), the Professor can't simply ask for the papers back. He must ask the students to report their grades. Without hesitating or consulting one another, each student says, of course, "We each got 100%" (call this proposition 100-Percent). Analogous situations of self-interested reporting could have serious consequences in lives or equipment in battlefield informant situations.

Since the Professor has recorded the total of all the scores, including the missing ones, the Professor (but not the students) knows that the average of the missing scores is 80% (call this proposition 80-Average).

The proposition 100-Percent and 80-Average cannot both be true. Suppose that the Professor has a Reliability of A (weight 5) and the students in question each have a reliability of D (weight 2). Then the possible sets of consistent statements are: {100-Percent, -80-Average}, {-100-Percent, 80-Average}, {-100-Percent, -80-Average}. These have aggregated reliabilities of: 5, 6 and 11, respectively. Cholvy's method thus supports the students' claim over that of the Professor. That is, the professor should override his previous total and record the three 100% grades into his grade book.

Clearly, this is wrong, however.

Suppose further that the Professor's average grade clause is upgraded from being just another observation to being an integrity constraint on the reports. Further, let us imagine that two of the students (I, II) have reliability C (weight 3) and one (III) has reliability D (weight 2). The students report their grades, once again as 100%. Given the constraints for these reports, models that satisfy the integrity constraints are: {100, 100, 40}; {100, 40, 100}; {40, 100, 100}; {80, 80, 80} … and so on. Since the first model has the lowest reliability contrary to it (because the third student has the lowest reliability), and the others are not attested at all, the professor is obliged to give the first two students perfect grades and the last one a 40% grade (failure), according to Cholvy's algorithm. The other possible models of the situation don't have as much reliable support. Clearly, this too potentially over-rewards (and over-punishes) the sources (students) based on their (low) prior reliability.

To return to Bayesian analysis momentarily: under a Bayesian approach, the posterior probability that all three students and the professor report the truth is zero because they are jointly inconsistent. However, we can calculate the posterior probability that any three of the reports are correct. That is, we can calculate the posterior probability that all three students correctly report their grade, given the appropriate priors and the reliability of each of the students. We can also calculate the probability that any two of the students report their grade correctly and the Professor reports the average correctly, (i.e. so the third scored 40%). Finally, we can calculate the probability that the Professor reports the average of the scores accurately, given prior probability of the average of any three grades being 80% and the past reliability of the Professor. This all assumes that the Professor can provide accurate prior probabilities for all these events, perhaps based on the prior performance of the students and class. These calculations will not necessarily yield the truth, however.

As Balaban and Yost point out, there are ways for the Professor to structure this situation so as to be extremely confident that the students report the truth. In these ways of structuring the situation, there is a Nash equilibrium (a strategy that it is in the best interest of all the participants to follow) that provides an incentive for all the students to report the truth in order to achieve the highest payoff. For the Professor, this means getting as close as possible to the actual grades, and for the students, we assume that this means getting the highest possible grade. In the unconstrained situation, the Nash equilibrium for the students, assuming they care only about receiving the highest grade, is to report 100%.

To get as close to the truth as possible, the Professor announces that the students will get the grade they report independently, according to the following payoff function g_i, where $\bar{s}$ = 80 and $\bar{r}$ is the average of the reported grades (Equation 1).

Equation 1 Credibility-Enhancing Payoff Function

$$g_i = r_i - \begin{cases} 0, & if \quad \bar{r} \le \bar{s} \\ k(\bar{r} - \bar{s}), & if \quad \bar{r} > \bar{s} \end{cases}$$

$$\text{where } \bar{r} = \sum_{i=1}^{n} \frac{r_i}{n} \text{ and } k > n$$

That is, the students get the grade g_i they report if the reported average is less than the observed average (under-reporting). On the other hand, the students are each penalized by an amount related to the difference between the reported average and observed average if the reported average is greater than the observed average (over-reporting).

In the original, unconstrained report scenario, the students' payoff function was $g_i = r_i$: the student received the grade he or she reported. As such, it was in the student's interest to report the maximum grade. The student only has to know the payoff function to understand what he or she should report, independently of consulting with the others. This is the Nash equilibrium in the original scenario.

As such it seems clear that the algorithm suggested by Cholvy, and even Bayesian methods relying on prior probabilities derived from historical data, are inappropriate in at least one class of situations: when the reporter is self-interested and the reporter's payoff function depends only on what he/she reports. In such situations, it is obvious to the informant what he or she should independently report out of self-interest, as long as he/she knows the payoff function in advance, and he/she need not coordinate in advance to achieve a jointly optimal strategy.

The solution seems to require that our representation of the situation includes not only a representation of the information Source in an Utterance Situation, the source's Reliability, and a representation of the Proposition expressed and its Credibility, but also a representation of the Payoff Function of a source, for human informants, and the variables the payoff function depends upon (Figure 5).For example, based on the extended STO Ontology represented in Figure 5, one could implement a formal rule (represented here in pseudocode) stating that, if a proposition is judged more credible than *4:Doubtful*, but the payoff function for the Source depends only on the proposition asserted, then the proposition is *4:Doubtful* (Self-Interest).

If Proposition ?p hasCredibility *?c,*
And ?c moreCredibleThan *4:Doubtful,*
And ?p isAssertedIn UtteranceSituation ?u,
And ?p isAssertedBy Source ?x,
And ?u isPartOf Situation ?s,
And Situation ?s hasPayoffFunction ?f,
And ?f isPayoffFor ?x,
And ?f dependsOn Variable ?r
And not
  ?f dependsOn ?s, ?s ≠ ?r
And ?r functionOf ?p
Then assert ?p hasCredibility *4:Doubtful* (Self-Interest)

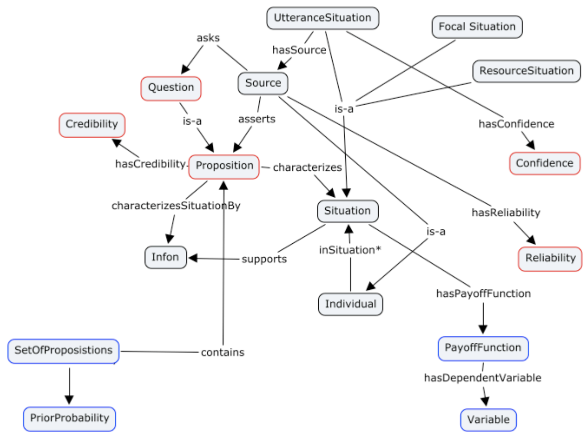
That is, a Source's assertion of a proposition should be taken to be *4:Doubtful*, if the Source is rewarded in the situation simply based on what the Source reports alone.

V. Discussion

In today's military missions, information must be collected and analyzed across the entire Political, Military, Economic, Social, Infrastructure and Information spectrum in order to inform military intelligence and operations. This information often originates outside of the intelligence process, but nevertheless must be evaluated and incorporated into military decision -making [2].

Whether information originates from within or outside the intelligence collection process, it is always possible that context-dependence, heuristics and biases all play a part in determining what information is selected for a report, what questions are asked in interviews, which colleagues are consulted, and so on. Furthermore, without complete knowledge of the networks (social and otherwise) that sources exist in, it is impossible for the reporter and consumers of this information to identify dependencies between information sources.

Figure 3 STO Ontology extended with Probabilities and Payoff Functions



Consumers of reported information are all subject to order effects (the order in which information is received can affect its interpretation [21]), representation errors (when information contradictory to one's hypotheses is dismissed or downplayed), and confirmation bias [23] (information is sought which tends reinforce one's hypotheses). Other aspects of context-sensitivity such as knowledge of recent events, the nature of the enemy and its current tactics, techniques and procedures can also influence which information is believed, sought, and how it is interpreted, and how reliability (i.e., which involves trustworthiness and expertise among other factors[18]) and credibility judgments are made. As such, there are many simplifications within STANAG 2022, which as it stands leaves much room for interpretation and formalization.

Once a report with its associated STANAG 2022 evaluations enters the information stream, there is no Army-prescribed (let alone overarching normative) theory for reasoning with this knowledge. The factors involved are multifaceted, and the disciplines studying reasoning with reliability and credibility tend to focus on isolated factors. Automating such reasoning is likely to produce false inferences in many cases and could thus be rejected by its potential users.

Further research is required to determine whether it is possible for automated systems to generate a range of likelihoods that the inference is false. If the range is typically too wide, intended users may not use the automation. If we assume the range isn't typically too wide to be of any utility to an analyst, it may be possible to direct the user's attention to where the most important potential credibility and reliability problems exist in the information and inferences made thus far. Perhaps the user of such a future system could ask for human corroboration and/or initiate a new collection task, or look for information in a database (or open source) not used thus far. Perhaps the system could provide visualizations for the user of aspects of the contextual situation that the augmentation believes are highly relevant to the reasoning.

VI. Conclusion

We have described ways in which reasoning about source reliability and information credibility is context- and situation-dependent, and shown how it can be formally represented in the unifying framework of Situation Theory. We have outlined some ways forward in formally reasoning about situations with source reliability and information credibility by means of our STO ontology, which is extended in various ways to account for the STANAG 2022 metrics, Bayesian probabilities and game-theoretic payoff functions.

We said that Bayesian reasoning is impractical in most military situations because the required prior probabilities are unavailable. Cholvy's non-Bayesian weighted majority algorithm was shown to yield the wrong predictions in situations in which the source is constrained only by self-interest. In such situations, it is not appropriate to trust the informant's reports despite their prior reliability and despite the independent confirmation of what they say. We have outlined how rules can be implemented for detecting such situations formally in an inference engine by extending the STO ontology. Completely reliable automated reasoning with trust-annotation judgments is far from completely implementable at this point. However, given all the considerations outlined here, a unified formal representation of the factors involved might have utility within the context of tools to assist the analyst in checking for context-dependent biases in the near term.

References

[1] A.J.L. Harris; U. Hahn. Bayesian Rationality in Evaluating Multiple Testimonies: Incorporating the Role of Coherence. J. of Exp. Psych: 2009, Vol. 35, No. 5, 1366–1373..

[2] Flynn, MG M. T., Pottinger, Capt. M., USMC, Batchelor, P. D., "Fixing Intel: A Blueprint for Making Intelligence Relevant in Afghanistan", Ctr for a New American Security (CNAS) Working Paper. Jan 4, 2010.

[3] J. Barwise, J. Perry, Situations and Attitudes, MIT Press, 1983.,

[4] K. Devlin, Logic and Information, Cambridge U. Press, 1991.

[5] L. Boven, S. Harmann. *Bayesian epistemology*: Oxford U. Press. 2003.

[6] L. Cholvy. Information Evaluation in fusion: a case study In Proc. of 10th Int'l Conf. on Info Processing and Mgt of Uncertainty in Knowledge-Based Systems (IPMU'04), Perugia, July 2004.

[7] L. Cholvy. Modelling Information Evaluation in Fusion. Proc. of 10th Int'l Conf. on Info Fusion (FUSION'2007), July 2007.

[8] M. Kokar, C. Matheus, and K. Baclawski, "Ontology-based situation awareness," Information Fusion, vol. 10, no. 1, pp. 83-98, January 2009.

[9] North Atlantic Treaty Organization (NATO) Standardization Agreement 2022 (Edition 8) Annex. Essentially the same matrix is presented in Appendix B "Source and Information Reliability Matrix" of US Army FM-2-22.3 "Human Intelligence Collector Operations" (2006),

[10] C. Matheus, B. Dionne, D. Parent, K. Baclawski and M. Kokar. BaseVISor: A Forward-Chaining Inference Engine Optimized for RDF/OWL Triples. In Digital Proceedings of the 5th International Semantic Web Conference, ISWC 2006, Athens, GA, Nov. 2006.

[11] M.R. Endsley, Theoretical underpinnings of situation awareness: a critical review, in: Situation Awareness Analysis and Measurement, Lawrence Erlbaum Associates, Mahawah, NJ, USA, 2000.

[12] R. Balaban. J. Yost. Professor Gump's Dilemma: A Classroom Exercise in Game Theory, J of Econ and Finance Ed. Summer, 2006. , pp. 8-13.

[13] R. Becker. D. Corkill. Determining Confidence When Integrating Contributions from Multiple Agents. In *6th Int'l J Conf on Autonomous Agents and Multi-Agent Systems* (AAMAS 2007), Honolulu, ,May 2007.

[14] M. Jensen; J. Burgoon. J. Nunamaker. Judging the Credibility of Information Gathered from Face-to-Face Interactions. *J. Data and Info. Quality* 2, 1 (Jul. 2010), 1-20.

[15] L. Page. S. Brin. R. Motwani. T. Winograd,. The pagerank citation ranking: Bringing order to the web. Report, Stanford Digital Library Technologies Project. Stanford, CA. 1998.

[16] G. Powell, C. Matheus, M. Kokar and D. Lorenz, Understanding the Role of Context in the Interpretation of Complex Battlespace Intelligence. In Proceedings of FUSION'06, Florence, Italy, July, 2006

[17] C. Matheus, D. Tribble, M. Kokar, M. Ceruti and S. McGirr, Towards a Formal Pedigree Ontology for Level-One Sensor Fusion. In Proceedings of the 10th ICCRTS. McLean, VA, June 2005

[18] B. Hilligoss. S. Y. Rieh. "Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context." *Information Processing & Management*, 44(4), (2008). 1467-1484

[19] Todorov, A. Evaluating faces on trustworthiness- An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Year in Cognitive Neuroscience 2008* Vol. 1124, pp. 208-224.

[20] G. Gigerenzer, H. Brighton. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1). 2009.

[21] A. Tversky. D. Kahneman. Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1974. Pp. 1124–1131.

[22] F. Schoorman, R.Mayer, J. Davis, An integrative model of organizational trust: past, present and future. *Academy of Management Review*, Vol. 32, No. 2, 2007. Pp. 344–354.

[23] R. Nickerson, Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, Vol. 2, No. 7, Pp. 175-220, 1998.

Brian Ulicny, Ph.D., VIStology, Inc., Framingham, MA. Gerald. M. Powell, Ph.D., U.S. Army Research Laboratory HRED, Fort Monmouth, NJ. Chester F. Brown, III, U.S. Army Intelligence Center of Excellence, Fort Huachuca, AZ. Christopher J. Matheus, Bell Labs, Ireland. Mieczyslaw M. Kokar, Ph.D., Northeastern University, Boston, MA. Jerzy Letkowsk, Ph.D., Western New England College, Springfield, MA.
Research performed under US Army contract W15P7T-09-C-S031.