Cognitive and Probabilistic Models of Group Decision Making

Yuqing Tang*, Christian Lebiere*, Katia Sycara*, Don Morrison *, Paul Smart†, and Paul Stone‡
*Carnegie Mellon University, US
†University of Southampton, UK
‡IBM, UK

***Abstract*—We introduce an experiment designed to study trade-offs in collaborative decision making environments such as finding the best level of selectivity and abstraction in sharing information, and their impact on the time course and accuracy of group decisions. Two models of the experiment are presented: a cognitive model using the ACT-R cognitive architecture and a probabilistic argumentation model using Markov Random Fields (MARF). The cognitive model relies on memory mechanisms such as spreading activation, partial matching and blending to judge when to share information, which facts are relevant to a given question, and how to aggregate probabilistic evidence. MARF carries out real world reasoning after formal theory of human argumentation while at the same time being flexible to accommodate the deviations from the theory. MARF follows knowledge engineering paradigm aiming at reaching correct reasoning as much as possible. Representative results from the experiment are presented and compared to the results of the two models. Implications of the results and avenues for future work are discussed.**

I. Introduction

Decision making in distributed environments has become a ubiquitous part of our environment. Collaborative networked environments range from Google Docs to elaborate military command and control centers. The design of such environments is far from trivial: while more information is generally better, too much information can also be detrimental by overwhelming its users. Given various cognitive and attentional bottlenecks, decision makers face a fundamental trade-off in interacting with this type of environment. One could attempt to exchange as much information as possible with partners on the collaborative network but the obvious gains are limited by two factors: the limitations of our perceptual-motor capabilities, i.e., how fast we can enter information into the network (e.g., by typing) and parse information received from the network (e.g., by sorting through chat messages), and the limitations on retaining information gleaned externally (e.g., forgetting messages read earlier). Conversely, one could attempt to focus on one's own experiences, making the most of them by rehearsing them and performing as many inferences as possible, but at the potential cost of neglecting crucial information available externally. The effectiveness of a focus on internal (i.e., personal experience and memory) vs. external (i.e., group experiences shared through the environment) sources of information fundamentally depends upon the precise quantitative nature of our cognitive architecture, the statistical nature of the external environment, and the organizational structure of the information-sharing tool (Reitter & Lebiere, 2012).

In this paper, we assume some degree of information sharing through the information environment and focus on a related dilemma: what level of information to share between decision makers. One possibility is to share detailed information, making sure that all decision makers have all potentially useful information, at the possible cost of overwhelming them with irrelevant details. The alternative is to share a high-level, refined version of the information available, hoping to maximize the utility of the exchange while minimizing the perceptual-motor and attentional costs. The difficulty of that trade-off is that one does not always (or even most of the time) know *a priori* what is likely to be of interest to another decision maker. The only way to know in general would be to have access to all their information, creating a Catch-22 situation.

In the rest of the paper, we present the design of an experiment intended to address the issue, specifically by allowing decision-makers to share detailed facts or a high-level guess regarding a variety of questions that can be answered using the facts. We then present two computational models of decision-making for that experiment that are intended to quantitatively address the tradeoffs described. We then present an analysis of experimental results as well as preliminary results from the model simulations on the accumulation of information and its impact on the fluctuation of decisions. Finally, we draw some parallels between the two models and discuss some potential extensions of the work.

II. Experiment

The task used simulates, using textual information, an artificial world region with political unrest. It requires four cooperating subjects to discover a variety of details and draw conclusions regarding an impending terrorist attack. The data underlying this task are from ELICIT (Chan & Adali, 2012). Facts[1] from which these conclusions can be drawn are released to the subjects in stages over time. Each fact is given to only one subject, each subject receiving facts disjoint from those given to other subjects, and the subjects must decide which facts to forward to which other subjects. For each trial, 68 facts are distributed in three waves. Each wave contains

[1]ELICIT calls these statements "factoids"

roughly 1⁄3 of the 68 facts. Between two consecutive waves, the subjects have 5 minutes to process a wave of new facts. At the end of a trial, the subject must submit their best conclusions from the facts collected, 15 minutes after starting. In addition to running experiments with four human subjects, automated subjects (bots) were also implemented, and data were collected for single subjects playing with three bots, though without the human subjects knowing their teammates were automated.

Each of the four subjects is asked to answer a different question about the attack: who, where, what and when. For any group of subjects these four questions are distributed once to each subject on four different trials. The first is a training trial, the results of which are not used, followed by three experimental trials. The answer to the who question is the name of the group expected to conduct the attack; group names are colors, such as the “gold group” or the “violet group.” The answer to the where question is a country name; country names are derived from Greek letter names, such as “Chiland” or “Omegaland.” The answer to the what question is a kind of target, such as “embassy” or “military base.” The answer to the when question has a four-fold structure, consisting of month name, day of the month, hour on a twelve hour clock, and “AM” or “PM.” While not the subject of any of the questions, there are also individuals, who serve as links connecting some of the facts presented to subjects; individuals are named after animals, such as “the Lion” or “the Jackal.”

The facts delivered to the subjects are sentences. Some are simple and immediately useful, such as “The attack will be at 11:00.” Though even this fact is delivered to the “where” subject, and so must be forwarded by that subject to the “when” subject. Others are more complex, and must be combined with other information to be useful; for example, “The Azure and Brown groups prefer to attack at night,” or “The Lion is known to work only with the Azure, Brown, or Violet groups.” Some of the facts delivered are essential for constructing correct answers, others are helpful but not essential, and still others are mere noise, contributing nothing to correct answers.

The four subjects interact with the system and with each other through a web-based user interface, Figure 1, implemented with HTML and JavaScript. This interface is divided into several panes. One, on the right, summarizes the player’s current role (who, where, what or when), describes the names and roles of the other players, and allows access to the instructions for reference.

The most prominent pane of the interface is the inbox, to which new facts are delivered. These may be new facts, delivered by the system; or they may be facts forwarded by another subject. Facts are normally displayed here in a partially obscured form, with only a few keywords, such as “Yellow,” “Magenta” and “Green,” legible, the rest of the text being replaced with ellipses. The user can click on a fact to cause the full text to be presented. When the mouse pointer is moved off the fact it is partially obscured again; by recording the users’ mouse actions insight into the the users’ attention can be gleaned. Below the inbox is a pane multiplexed for three purposes: outbox, mylist and guessbox. When used as the outbox facts can be dragged to it, and forwarded to other subjects, in whose inbox they will appear. When used as mylist, facts can dragged to it for future reference; while users can use this for whatever purpose they choose, it is expected that those who do employ it will use it to consolidate facts they suspect are important for answering their own question. Facts in mylist, as in the inbox, are normally partially obscured, and must be clicked to be read in full. At several points in each round subjects are asked to make their best guess so far at the question they have to answer, along with their confidence in the guess, on a five-point scale. In this way, we can trace human subjects’ behaviors on accumulating facts and its impact on the fluctuation of decisions. These decision traces are then compared with the traces produced by the ACT-R model and the probabilistic argumentation model.

III. Models

Two different computational models of this task were implemented, and their results compared to the human data. The reason for using two different modeling paradigms is to study what each can contribute to understanding group decision making and draw lessons from any parallels or differences between models. (Lebiere, Gonzalez, & Warwick, 2009)

A. ACT-R Cognitive Model

The ACT-R model uses the ACT-R cognitive architecture (Anderson & Lebiere, 1998) and in particular leverages the activation calculus in declarative memory. This modeling approach reflects the fact that performance in this task heavily relies on retrieval of information from memory, and the activation processes in ACT-R declarative memory provide powerful mechanisms to guide information retrieval as well as embody limitations on the storage and retrieval of memories under demanding conditions.

The task is decomposed into three component sub-tasks: information sharing, inductive inference, and probability estimation. The goal of the information sharing subtask is to determine which facts to share, and if so with which of their teammate(s). The basic approach is to share facts that are semantically related to the question domain of the teammate, e.g., share facts containing location information with the person in charge of the ’who’ question. The implementation leverages the ACT-R partial matching mechanism that retrieves chunks in declarative memory by combining activation with semantic similarities.

The goal of the inductive inference subtask is to determine the relevance and applicability of various facts to the specific question. The approach is to activate facts whose context is associated with a specific guess or answer. The implementation leverages the ACT-R spreading activation and base-level learning mechanisms. Finally, the goal of the probability estimation task is to determine the probabilty of each candidate answer given various facts and their activation. This approach is grounded in the assumption to represent each fact as a rough probability estimate over the given options. The probability

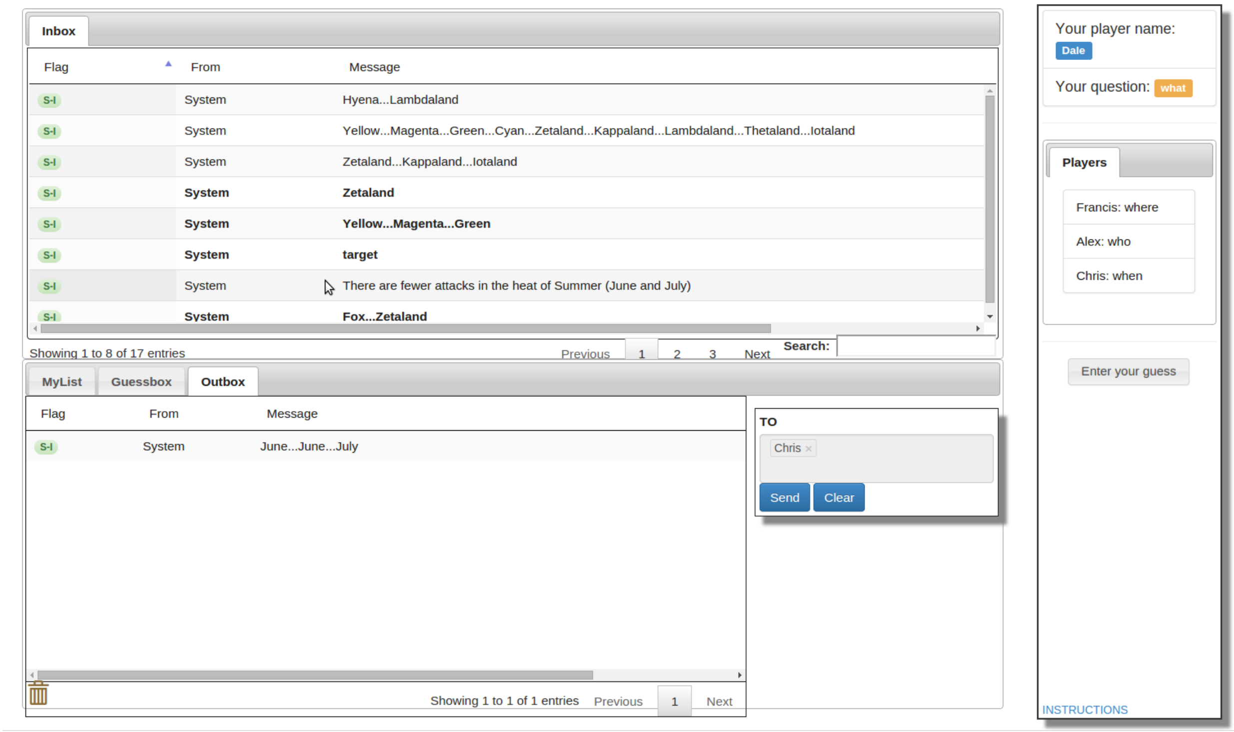

Figure 1. The user interface of the experiment

estimation process leverages the blending memory retrieval mechanism to generate aggregate estimates.

The information sharing model assumes that people use primarily simple heuristics in determining whether information is relevant to another decision maker. In this case, the assumption is that subjects share a fact if they contain information of the same semantic domain as that person's question (e.g., location information if the given question is 'who'). While this approach is well short of optimal it has the advantage of being highly efficient and avoids assuming knowledge of the other decision makers processes that is unlikely to be available. To avoid implementing this approach using large numbers of ad hoc heuristic rules (e.g., one for each combination of question and answer), pattern matching processes in memory are used instead. Each fact is encoded as a set of semantic keywords reflecting the key information contained in the sentence. For each subject and given question, each keyword is matched against the various questions to determine if it contains information related to that question's domain. The ACT-R partial matching mechanism is used by setting high semantic similarities between concepts of a common domain (e.g., locations like 'psi' and 'chi', and a question like 'where'). This will result in the relevant question(s) retrieved for each fact, indicating the relevant decision maker with whom to share that fact. The mismatch penalty scaling factor controlling the partial matching process will determine the selectivity of the process. Thus varying the scaling factor can determine the overall willigness to share information, leading to individual differences reflected in more conservative or widespread sharing.

The inductive inference model is used to reflect dependencies between answers to different questions. For instance, an answer (or guess) of a given location to a 'where' question (e.g., 'psi') will raise the relevance of facts mentioning that location when answering other questions (e.g., 'who'). The most natural way to implement that dependency process is to use the spreading activation mechanism. To do that, each fact is associated with a context element representing its dependency upon another answer (guess). When a given answer is processed, related facts are retrieved from memory by spreading activation from that answer to the facts including that answer as context. Those facts receive a boost in activation from the base-level learning process, making them more salient in the subsequent probability estimation process. This boost can result in well-known cognitive biases such as availability bias. This approach is similar to the model of the impact of memory availability in the model of sequential diagnostic reasoning of Melhorn et al (2011).

The probability estimation model follows the instance-based learning (IBL) modeling methodology (Gonzalez, Lerch, & Lebiere, 2003). To provide for finer discrimination in judgment and ensure the ability to gradually accumulate evidence from a stream of individual facts, the basic problem of determining the most likely candidate answer for each question is formulated as a goal to assign a probability to each potential answer. The goal is defined as a chunk of type hypothesis that contains three slots:

- Question: the representation of the question, i.e., who,

what, where and when
- Answer: the representation of each possible answer, e.g., various groups for who
- Probability: a probability value assigned to the question-answer pair

This representation follows the general IBL pattern of context (question), decision (answer) and outcome (probability). In keeping with the instance-based methodology, this representation is used both for facts as well as goals. Specifically, most facts are transformed into chunks of this type if they make a strong assertion about a given question. For instance, if the fact rules out a particular group, a hypothesis chunk will be created (or reinforced if it already exists) stating (who, group, 0). Conversely, if it strongly implies a group's involvement, the chunk (who, group, 100) will be created. If the fact mentions the possible involvement of n groups, then a separate hypothesis chunk is created for each group with a probability of $1/n$, reflecting mutually exclusive participation.

Of course, those assertions are not literally correct–rather the intent is to provide the basis for a rough estimate of relative probabilities based on the information provided. More precise facts (e.g., stating actual probabilities, or using qualifiers such as likely or probably) could be used to create more accurate chunk encodings. When the model is asked to generate a guess to a question, it iterates through all the possible answers (e.g., all the groups for a who question) and generates a probability estimate for each using the blending mechanism used for memory retrievals (Lebiere, 1999). During memory retrievals, each chunk in memory has an activation that reflects factors such as recency, frequency, and degree of match to the requested pattern. Recency is factored through a power law decay from the time that the chunk is created. Frequency reflects a power law of practice of the numbers of times that a chunk is strengthened following rehearsals. For degree of match, we assume for simplicity that each question and answer are distinct and no similarities are defined. Blending retrieval then assigns for a given question-answer pair a probability to each chunk matching that request (in general, there will be several) reflecting a softmax (Boltzmann) distribution of chunk activations given a certain amount of noise. Those probability estimates for each chunk associated with the question-answer pair are then blended according to a weighted average of the chunk probabilities (assuming linear similarities over the probability space (Lebiere et al., 2013)). The probability estimates are not normalized but instead the largest one is selected to generate the guess. All parameters controlling the behavior of the model are left at their default values: the base-level decay rate is 0.5, the mismatch penalty is 2.5, the activation noise is 0.25, and the blending temperature is 0.4.

Note that, as mandated by the ACT-R theory, the hypothesis goals generated to provide the guess become themselves chunks in memory, as are guesses received from other agents. This can give rise to cognitive biases such as confirmation bias, where a strong initial estimate leads to overoptimistic estimates later despite contradictory evidence.

B. Probabilistic argumentation model

We developed the Markov Argumentation Random Field (MARF) (Tang, Toniolo, Sycara, & Oren, 2014), which is a combination of formal theory of human reasoning in argumentation and Markov random fields. The formal theory of argumentation (Dung, 1995) formalizes the essentials of human reasoning about inconsistent, uncertain and incomplete information in the course of argumentative dialogues. However, in real world scenarios deviation from the formal theory is unavoidable. MARF is a probabilistic model which carries out real world reasoning after the formal theory of human reasoning while at the same time being flexible to accommodate the deviations from the theory. Unlike the ACT-R model which focuses on revealing the cognitive process of human reasoning, MARF follows the knowledge engineering path aiming at reaching correct reasoning as much as possible.

A Markov random field (Kollar & Friedman, 2009) is a graphical model which encodes local Markov properties — a random variable is independent of all other variables given its neighbors — as an undirected graph to establish probabilities of all valuations to the variables. Echoing the local Markov properties, Dung's argumentation semantics (Dung, 1995) can be recovered by applying a list of acceptability rules based on a graphical model of argument interaction. For example, "A is labeled IN (accepted) if all its attackers are labeled OUT (rejected)" (Caminada & Gabbay, 2009). Such rules, which assign acceptability to an argument given the status of its neighbors, also satisfy local Markov properties. Moreover, the construction of arguments as proof networks (Tang, Cai, McBurney, Sklar, & Parsons, 2011) also admit the local Markov properties — the establishment of a conclusion is independent of all other rules given the premises of the rules for the conclusion. These two observations allow us to construct Markov Argumentation Random Fields (MARF).

MARF compiles the argumentative knowledge and received information into a mathematically rigid Markov Random Field. The resulting MARF is able to track both supporting links and conflicting links (argumentative defeats) among the outcomes, the applied knowledge and the received information. It can compute the most probable argumentation for the outcomes and identify the pieces of knowledge or received information that would render the premises or outcomes unreliable or reverse the outcome dramatically.

For example, the MARF in Figure 2 is compiled from the following facts in the ELICIT tasks[2]: (1) The Lion is involved; (7) The Chartreuse group is not involved; (9) The Purple or Gold group may be involved; (10) All of the members of the Azure group are now in custody; (12) There is a lot of activity involving the Violet group; (13) The Brown group is recruiting locals - intentions unknown; (16) Members of the Purple group have been visiting Omega; (18) The Azure group has a history of attacking embassies; and a domain constraint (S-1) there is only one answer for the who question: either Brown, Violet,

[2]The numbering of the facts are same as it is in the ELICIT fact set coded as 1aGMU.

Chartreuse, Purple, Gold, or Azure.

In Figure 2, Oval nodes are *variable nodes* tracking the acceptability status (i.e., *accepted*, *rejected*, *undecided*) of predicates (including equality assertion, e.g. who? := "Brown"). Square nodes are *factor nodes* modeling how predicates acceptability status interrelate with each other regarding the meaning of facts. For example, fact "10) All of the members of the Azure group are now in custody" relates acceptability of predicates inCustody("Azure") and the equality assertion who? := "Azure" (the answer to who is "Azure"). If inCustody("Azure") is accepted, then who? := "Azure" is likely to be rejected. Every factor node is associated with a weight to reflect how much such a factor should be taken into account when evaluating the probability of an acceptability assignment to predicates via an exponential family distribution parameterized by the weights of the facts:

$$\Pr(\vec{x}) = \frac{1}{Z} \prod_{F_j \in \mathscr{F}} \exp\left(\langle \vec{W}_j, \phi_j(\vec{x}_j) \rangle\right)$$

where $\langle \vec{W}_j, \phi_j(\vec{x}_j) \rangle$ is the inner product of the weights and the argumentative features $\phi_j(\vec{x}_j)$ of the acceptability variables vector $\vec{x}_j$ of a fact F_j. Z is a normalization constant to ensure that $\Pr(\vec{x})$ is a probability distribution over all possible acceptability assignments. Argumentative features $\phi_j(\vec{x}_j)$ reveal elements that are essentials in evaluating the meaning of the fact according to the formal argumentation theory. The higher the validity of an acceptability assignment is, the higher the probability of such an assignment will be; the higher the weight of a fact is, the higher the probability will be for an acceptability assignment that conforms with the meaning of the fact.

With MARF, we can model the interactions of premises, conclusions, inference rules, and argument attacks quantitatively through potential functions. Simple operations on these potentials facilitate the computation of a coherent probabilistic interpretation of the argumentation outcome — the argumentation structure along with the acceptability status assigned to premises, conclusions, inference rules and arguments. In addition, MARF provides a computational framework to learn probabilistic evaluation functions of the premises and outcomes following data revealing human reasoning.

IV. Results

In this section, we will compare the results of human experiments, the ACT-R model, and the MARF model running on the same ELICIT task.

A. Human experiment results

Sixty subjects, divided into twelve groups of five, were recruited and finished the task. While they did not know how they were divided, four of each five worked cooperatively together, and the fifth worked separately, with three bots. Among the 60 subjects who participated in our experiments, 15 of them (including the subjects who worked with bots) answered the "who" question for the fact set "1aGMU17". The results are depicted in Figure 3. Among all these participants, 50% of them reached to the correct answer, "the Violet group", after seeing the first wave of facts. After seeing the second wave of facts, 100% percentage of the participants reached the correct answer. However, after the third wave, about 40% of the participants were confused by the new facts and changed their answers from the correct one.

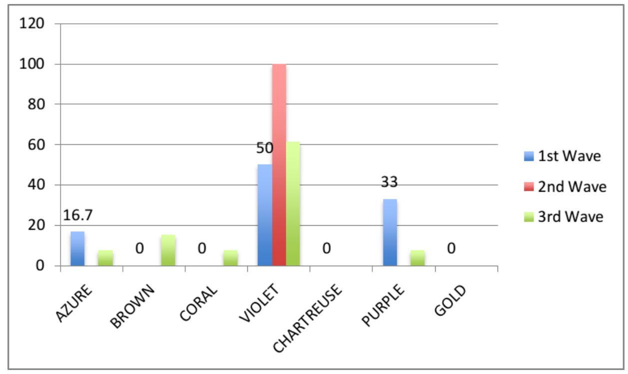

Figure 3. The human experiment result

B. ACT-R Results

Sample results for the "who" question are presented in Figure 4 of fact set "1aGMU17". Probability estimates for each possible answer (i.e., all groups) are presented for each of three waves of facts. Note that those are the unnormalized (non-exclusive) probability estimates rather than actual (exclusive) forced-choice answers. The initial estimate for the violet group (the correct answer, as it turns out) is the highest following the first and second batch of facts, making it the preferred choice in these two phases as for the human subjects. However, the estimate for the violet group falls to third-highest after the third batch of facts due to a dilution effect from a number of facts mentioning other possibilities.

Note that these results were generated without reflecting the effect of previous guesses on later phases. This would be a case where confirmation bias could actually lead to a correct final answer by strengthening the correct guess based on the effect of early evidence.

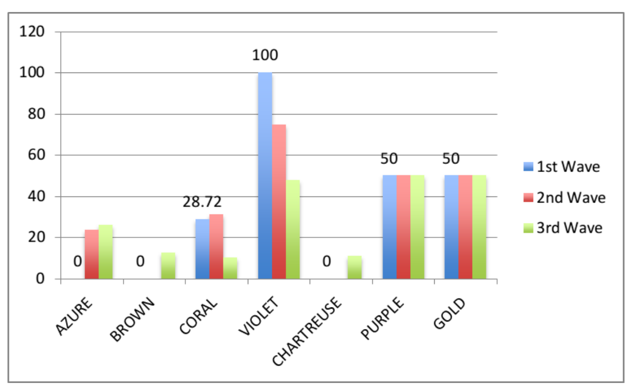

Figure 4. The ACT-R result

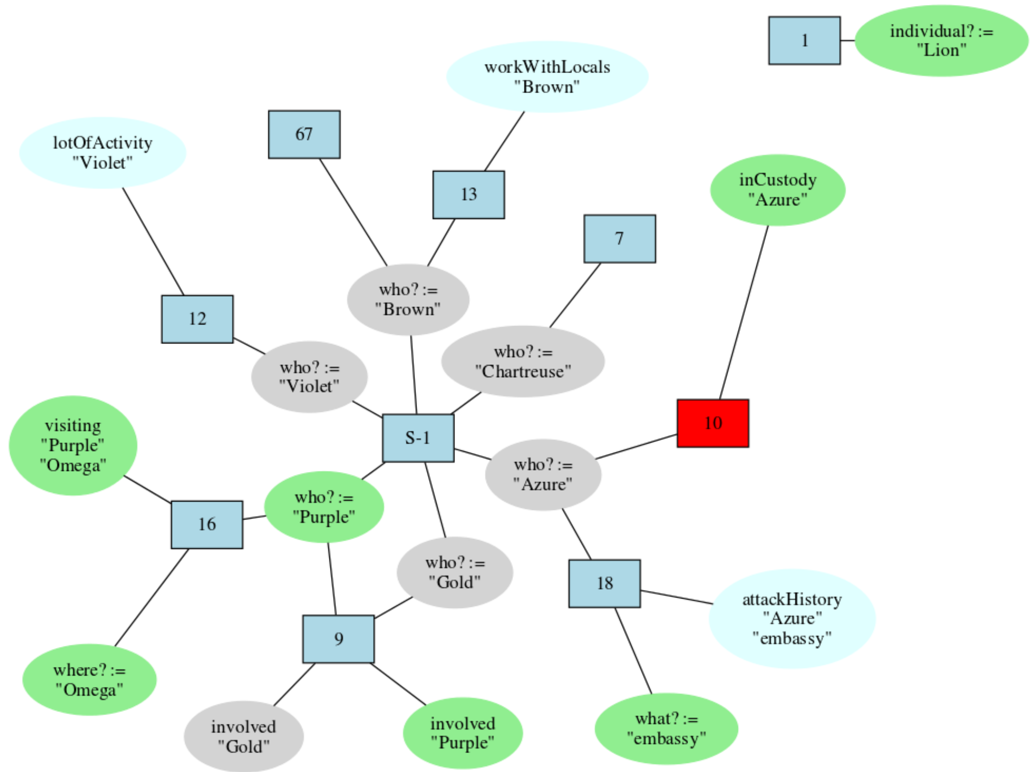

Figure 2. A compiled Markov Argumentation Random Fields after the first wave of facts (light green oval nodes are accepted predicates; grey oval nodes are rejected predicates; light blue oval nodes are undecided predicates; red square nodes model the argumentative conflicting relationships among predicates)

C. Probabilistic argumentation results

The results MARF over the same ELICIT task (the fact set “1aGMU17”) is depicted in Figure 5. The three waves of incoming facts are compiled into three MARFs. Figure 2 is the MARF compiled from the first wave of facts. In this first wave MARF, the fact “(1) The Lion is involved” is disconnected from other facts because the meaning of this fact is disconnected from other facts. After the second wave of facts, fact (1) is connected to the majority of facts (as depicted in Figure 6); however, there is a new disconnected fact (15). As more and more facts becomes available, the MARFs are able to consider more connected facts to evaluate the acceptability of the predicates underlying the meanings of these facts. After 3 waves of facts, the MARF is able to evaluate acceptability status of each answer as marginal probabilities, i.e., the probability of accepted, rejected and undecided, considering all available facts. To align with the results of human experiments and the ACT-R model, Figure 5 plots the probabilities of accepting the answers omitting the probabilities of rejected and undecided status of these answers where weights of all the facts are set to 10 and the weight of the domain constraint is set to 100. In the first two waves, MARF decides that the “Purple” group is the answer with probability of 94%. However, after the receiving the third wave, MARF changes its opinion sharply to the “Violet” group. This is the case because different from human and ACT-R model, the MARF is constructed to be decisive separating the accepted answers and the rejected answers as much as possible while following meanings of the available facts as closely as possible.

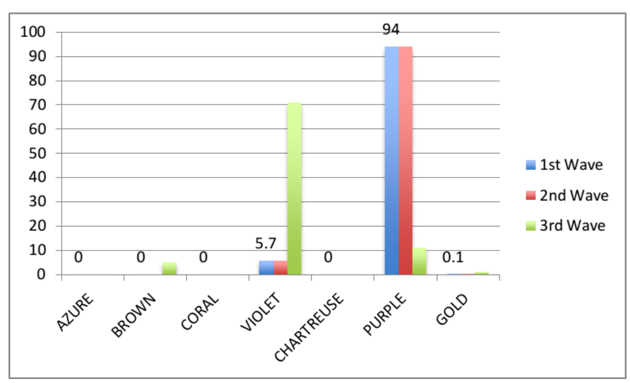

Figure 5. The probabilistic argumentation result

V. Conclusions and Further Work

We present an experiment and two computational models of group decision making. While both data analysis and model development are preliminary, they highlight interesting emerging effects. Rather than following a linear path, the deductive processes faced with a constant stream of facts induce a fluctuation in beliefs that reflect a potentially rich dynamic. Both computational models capture some aspects of the human

Figure 6. The compiled Markov Argumentation Random Field after the second wave of facts (light green oval nodes are accepted predicates; grey oval nodes are rejected predicates; light blue oval nodes are undecided predicates; red square nodes model the argumentative conflicting relationships among predicates)

data but not others. While both include a representation of the deductive process (e.g. a question activates its relevant facts and answers in the ACT-R model; an MARF factor models a logical deductive rule) and its constituent facts and conclusions, the processes reflect distinct assumptions regarding the parallel vs sequential nature of inference processes, the implicit vs explicit nature of probabilistic information, and whether those processes are fundamentally optimizing or satisficing. Still, those models share many representational assumptions regarding the nature and structure of the problem representation, which will allow us to formally examine the implications of their assumptions.

ACT-R uses IBL to drive decision making; while MARF uses potential factors to relate facts and answers. If the given facts and its symbolic representation truly reflects logical relation among the facts and answers, by design MARF will produce the right answers. Furthermore, as factors in MARF are interrelated through an undirected graphical model, MARF is not sensitive in the order of receiving facts but sensitive in the availability of facts. On the other hand, since the ACT-R model uses IBL activation, it is more sensitive to the order of receiving facts. Therefore the ACT-R model follows human decision making closely while MARF follows closely the logical relationships of information embedded in the facts to estimate the answers.

Numerous avenues of work are possible for both data analysis and model development. We will examine whether learning processes can improve decision making with experience, develop models of judgments for information sharing, and analyze various experimental conditions to determine the answer to our initial question as to whether information is best shared at the most detailed level of basic facts or in the form of refined, high-level conclusions.

ACKNOWLEDGMENTS

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defense 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 Defense 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

Anderson, J., & Lebiere, C. J. (1998). *The atomic components of thought*. Mahwah, N.J.: Erlbaum.

Caminada, M., & Gabbay, D. (2009, December). A logical account of formal argumentation. *Studia Logica*, *93*(2), 109–145.

Chan, K., & Adali, S. (2012). An agent based model for trust and information sharing in networked systems. In *Cognitive methods in situation awareness and decision support (CogSIMA), 2012 IEEE international multi-disciplinary conference on* (pp. 88–95).

Dung, P. M. (1995, September). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, *77*, 321–357.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*, 591–635.

Kollar, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. The MIT Press.

Lebiere, C. (1999). The dynamics of cognitive arithmetic. *Kognitionswissenschaft Special issue on cognitive modelling and cognitive architectures, D. Wallach and H. A. Simon (eds.)*, *8*, 5-19.

Lebiere, C., Gonzalez, C., & Warwick, W. (2009). Convergence and constraints revealed in a qualitative model comparison. *Journal of Cognitive Engineering and Decision Making*, *3*, 131-155.

Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., & Anderson, J. R. (2013). A functional model of sensemaking in a neurocognitive architecture. *Computational Intelligence and Neuroscience*.

Reitter, D., & Lebiere, C. (2012). Social cognition: Memory decay and adaptive information filtering for robust information maintenance. In *Proceedings of the twenty-sixth aaai conference on artificial intelligence.*

Tang, Y., Cai, K., McBurney, P., Sklar, E., & Parsons, S. (2011). Using argumentation to reason about trust and belief. *Journal of Logic and Computation*.

Tang, Y., Toniolo, A., Sycara, K., & Oren, N. (2014). Argumentation random field. In *Eleventh international workshop on argumentation in multi-agent systems.*