# Analyzing the impact of human bias on human-agent teams in resource allocation domains

# (Extended Abstract)

Praveen Paruchuri, Pradeep Varakantham*, Katia Sycara, Paul Scerri
Robotics Institute, Carnegie Mellon University, {paruchur,katia,pscerri}@cs.cmu.edu
*School of Information Systems, Singapore Management University, pradeepv@smu.edu.sg

## ABSTRACT

As agent-human teams get increasingly deployed in the real-world, agent designers need to take into account that humans and agents have different abilities to specify preferences. In this paper, we focus on how human biases in specifying preferences for resources impacts the performance of large, heterogeneous teams. In particular, we model the inclination of humans to simplify their preference functions and to exaggerate their utility for desired resources. We then study the effect of these biases on two different problems, which are representative of most resource allocation problems addressed in literature.

## Categories and Subject Descriptors

I.2 [**Computing Methodologies**]: Artificial Intelligence

## General Terms

Human factors

## Keywords

Teamwork, DCOPs

## 1. INTRODUCTION

A range of exciting applications involve humans and agents working along side each other to achieve a complex objective. Domains for such applications include search and rescue[2], disaster response [5] and many others. Researchers envision automating the allocation of shared resources in these domains using the distributed constraint optimization problems (DCOPs)[4]. For example, access to satellites, robots, computation or space may be automatically assigned using DCOPs. Due to the computational load and communication intensity of these algorithms, humans in the team will need to communicate their preferences and utilities to a proxy that executes the algorithm on their behalf. If there are many resources these preferences may be communicated incompletely or approximately. However, agents participating in the allocation process will be able to precisely and completely specify their preferences for all resources. The question addressed by this paper is what happens to the quality of the overall resource allocation when human and agent preference specifications differ in this way.

Preference elicitation is a difficult problem that takes a lot of time and effort for humans. Many well known human biases come into play, either consciously or sub-consciously, when reporting the preferences, for instance the utility of a resource. In this paper, we use a combination of empirical and theoretical analysis to understand the impact of these biases using a DCOP algorithm.

We study the effect of two commonly known models of human biases [1] on two commonly studied resource allocation problems using two allocation algorithms. To be able to compute solutions to large problems quickly, we used the Distributed Stochastic Algorithm (DSA) [7]. Our expectation is that there will be very little change to the overall utility on average. This is because the solution provided by DSA is dependent on the local ranking of resources and not on the absolute utilities of the resources to the agent.

## 2. PROBLEMS AND MODELS

In this section, we provide a brief background of the motivating problems, models used to represent the problems and the approach employed to solve the models. Specifically, we start with describing two generic distributed resource allocation problems:

### 2.1 Problem 1: Discrete Resource Allocation

In this domain, resources need to be allocated to a group of agents and humans, $E$ based on their preferences [6]. For ease of explanation, we assume only one resource is allocated to one agent or human. We model this domain as a DCOP as follows: Each agent/human has a variable, represented as $e$. The values that the variable can take correspond to the resource allocated to the agent. This domain for variable $e$ is specified as $D_e$ and belongs to the set $1, 2, \cdots, R$. The utility for $e$ when a resource, $r_e \in D_e$, is allocated to $e$ is

$$u_e(r_e) = L * I_{r_e \neq 0} * (1 - \prod_{\tilde{e} \in E, \tilde{e} \neq e} I_{r_e \neq r_{\tilde{e}}}) + p_e(r_e), \quad (1)$$

where $p_e(r_e)$ is the preference value of agent $e$ for resource $r$ (soft constraint) and $L$ is a large negative number that represents the penalty for allocating one resource to two different entities (hard constraint) and $I_{condition}$ is 1 when $condition = true$ and is 0 when $condition = false$.

### 2.2 Problem 2: Distributed Event Scheduling

In this domain, $M$ meetings need to be scheduled for humans and agents (acting on behalf of humans). A meeting, $m$ can require multiple humans, $E_m$ and a human, $e$ could be part of multiple meetings, $M_e$. Furthermore, each human/agent, $e$ has preferences over the time slots, $p_e^t$ and meetings, $p_e^m$. Given this information, the goal is to compute a schedule which maximizes the utility of the team. This domain is modeled as a DCOP by [3] as follows:
(a) Each agent/human, $e$ has multiple variables, $M_e$, corresponding to all the meetings where $e$ is required.
(b) The values for each variable correspond to the time slots, where a meeting can be scheduled. This domain for variable $m \in M_e$ is specified as $D_m^e$ and belongs to the set $1, 2, \cdots, T$, where $T$ is the set of time slots available in which a meeting can be scheduled.

**Cite as:** Analyzing the impact of human bias on human-agent teams in resource allocation domains (Extended Abstract), Praveen Paruchuri, Pradeep Varakantham, Katia Sycara and Paul Scerri, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. 1593-1594

(c) If a meeting, $m$ for $e$ is scheduled at $t_e^m$, then the utility for $e$ is defined as

$$u_e^m(t_e^m) = L * (1 - \prod_{\tilde{t} \leq T, \tilde{t} \neq t_e^m, t_e^m \neq 0} I_{\tilde{t} \neq t_e^m}) * \\ (1 - \prod_{\tilde{e} \in E_m, e \neq \tilde{e}} I_{t_e^m = t_{\tilde{e}}^m}) + (p_e(m) + p_e^m(t_e^m)) \quad (2)$$

Here, $L$ is a large negative number that represents a penalty for scheduling the same meeting at two different time slots and other such cases. $(p_e(m) + p_e^m(t_e^m))$ represents the preferences of the humans (soft constraints).

# 3. BIASES CONSIDERED

## 3.1 Bias 1: Simplification of preferences

As shown in [1], humans tend to simplify preference values when faced with problems where multiple factors need to be considered. One popular way of simplifying preferences is thresholding. Formally, this involves approximating the preference function over a variable $v$, $p_e^v()$ as a function that has zero corresponding to all but the top "k" values in its range. In general, it could be a "multi-step" step function and defined as follows (with ordering of thresholds given as $thres_1 > thres_2 > \cdots$):

$$\begin{aligned} \tilde{p}_e^v(d) &= \max_{\hat{val}} p_e^v(\hat{d}), \quad if \quad p_e^v(d) > thres_1 \\ &= thres_1, \quad if \quad thres_2 \leq p_e^v(d) \leq thres_1 \\ &= \cdots\cdots \end{aligned}$$

Instead of a threshold, humans may simply limit themselves to specifying some number of top preferences. Specifically, we consider the scenario where humans specify preferences for only the most important resources while ignoring the rest. The modified function for this preference simplification is defined as:

$\tilde{p}_e^v(d) = p_e^v(d)$, if $d \in maxK(p_e^v)$
$= 0$, otherwise

## 3.2 Bias 2: Preference Exaggeration

This bias of humans arises due to exaggeration of the importance of certain features over others. This involves increasing the preference function value of a variable $v$, $p_e^v()$ for the top preferred value assignments. The modified function for this type of preference exaggeration is defined as:

$\tilde{p}_e^v(d) = \mathcal{S} * p_e^v(d) + \mathcal{A}$, if $d \in maxK(p_e^v())$
$= 0$, otherwise

$maxK(p_e^v())$ provides the top $k$ values in the range of the function $p_e^v()$, $\mathcal{S}$ and $\mathcal{A}$ are scaling and addition factors of exaggeration.

# 4. ALGORITHM FOR SOLVING DCOPS

Distributed Stochastic Algorithm (DSA) is the algorithm that we employ for solving DCOPs. It should be noted that DSA is the only algorithm that can realistically scale to the type of problems considered in this paper. The following key property of the DSA algorithm that will be very useful for our analysis.

PROPOSITION 1. *For DCOP problems 1 and 2, the solution provided by DSA is reliant on ranking of preference values and not on the actual preference values.*

# 5. IMPACT OF BIASES

In domains where the ordering of preference values remain the same as the original preferences after applying the human biases, we believe that according to Proposition 1 the impact will be zero. We further believe that even if there are disruptions to the preference order (due to human biases), DSA will provide solutions that are close to the optimal solution obtained with the original preference function. A key reason for this belief is that for large problem instances of interest, we expect to find that many solutions will have approximately the same utility.

We considered four different types of preference approximations in our experiments: Hch represents the problem where humans change their preferences using biases 1 and 2 (explained earlier) and Bch represents the case where both humans and agents change their preferences using bias models 1 and 2. The third approximation, EHch enhances Hch, where zero valued preferences are considered valid and would account for cases where preferences for a resource/time/event were decreased to zero from a positive value. Similarly, the fourth approximation, EBch enhances Bch.

Using these four preference approximations we performed different types of experiments. The parameters varied for the discrete resource allocation domain include: (a) Number of agents ($|E|$); (b)Number of resources ($R$); (c) Number of top choices reported by the human (corresponds to Bias 1, expressed as $maxK$); and (d) Exaggeration factor (corresponds to Bias 2, expressed using $S$ and $A$). Similarly, the parameters varied for distributed event scheduling are: the domain size (Number of agents/humans $E$, meetings $M$, maximum number of agents per meeting $E_m$, number of meetings per agent $M_e$ and time slots $t$), the number of top choices reported (Bias 1 captured using $maxK$) and the exaggeration factor (Bias 2 expressed using the terms $S$ and $A$). We performed a large scale simulation involving nearly 90 hours of computer simulation time. Both Uniform and Gaussian distributions were used to generate utilities for each agent across the various resources or meetings. Our main results can be summarized as follows:

(a) Across the various sizes of problem settings, the reward loss for the four variants of the DSA algorithm are within 10% of the DSA solution (within 7% for the enhanced versions) on an average while the random policy is atleast 30% away from optimal.

(b) Our experiments studying the effects of Bias 1 show that as long as humans specify their top few (>= 2) choices correctly, they will have negligible impact on the team utility.

(c) Our experiments studying the effects of Bias 2 show that exaggerating the preference values has little effect on team utility while having a significant effect on the individual subgroup utilities (i.e. agent and human subgroups). However, we could not find a systematic way to exploit this effect to favor a particular subgroup.

**Acknowledgments: This research is supported by MURI-7 award AFOSR FA955008-1-0356.**

# 6. REFERENCES

[1] J. Huber, D. Ariely, and G. Fischer. Expressing preferences in a principal-agent task: A comparison of choice, rating, and matching. *Organizational Behavior and Human Decision Processes*, 87(1):66–90, 2002.

[2] H. Kitano and S. Tadokoro. Robocup rescue: A grand challenge for multiagent and intelligent systems. *AI Magazine*, 22(1):39, 2001.

[3] R. Maheswaran, M. Tambe, E. Bowring, J. Pearce, and P. Varakantham. Taking DCOP to the real world: Efficient complete solutions for distributed multi-event scheduling. In *AAMAS*, pages 310–317, 2004.

[4] P. Modi, W. Shen, M. Tambe, and M. Yokoo. Adopt: Asynchronous distributed constraint optimization with quality guarantees. *Artificial Intelligence*, 161(1-2):149–180, 2005.

[5] R. Stranders, A. Farinelli, A. Rogers, and N. Jennings. Decentralised Coordination of Mobile Sensors Using the Max-Sum Algorithm. In *IJCAI*, 2009.

[6] M. Yokoo, O. Etzioni, T. Ishida, and N. Jennings. *Distributed constraint satisfaction: foundations of cooperation in multi-agent systems*. 2001.

[7] W. Zhang, G. Wang, Z. Xing, and L. Wittenburg. Distributed stochastic search and distributed breakout: properties, comparison and applications to constraint optimization problems in sensor networks. *Artificial Intelligence*, 161(1-2):55–87, 2005.