2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Measuring Human-Robot Team Effectiveness to Determine an Appropriate Autonomy Level

Tobias Kaupp and Alexei Makarenko

***Abstract*— This paper proposes a methodology to measure the effectiveness of a human-robot team as part of an adjustable autonomy system. The effectiveness measure is aimed at determining an appropriate autonomy level prior to the system's deployment. Two competing goals need to be traded off: maximising robot performance while minimising the amount of human input. The relative importance of the two goals depend on the mission priorities and constraints which are taken into account. The proposed methodology is applied to a human-robot communication system developed for task-oriented information exchange. The robot uses a decision-theoretic framework to act autonomously and to decide when to request input from human operators. The latter is achieved by computing the value-of-information an operator is able to provide which is compared to the cost of obtaining the information. For our system, the cost parameter represents the autonomy level to be determined. We demonstrate how an appropriate autonomy level can be found experimentally using a navigation task. In our experiment, the robot navigates through a set of simulated worlds with human input being generated by a software component. The results are used to find appropriate autonomy levels for three example missions and a subsequent user study.**

I. INTRODUCTION

Fully autonomous robots are still a vision of the future. Combining the strengths of humans and robots to achieve a task cooperatively is becoming a popular paradigm [2][3][5]. Adjusting a robot's autonomy by allowing human input is one way to achieve the combination. The underlying assumption is that robot performance increases with more human input.

Robot performance and the amount of human input required to achieve the performance are two metrics used in this paper to measure *team effectiveness* as a function of robot autonomy. The objective is to find an appropriate autonomy level for an Adjustable Autonomy (AA) system prior to its deployment. The autonomy level acts as a design parameter which determines under what circumstances the AA system adjusts its autonomy online. To find an appropriate autonomy level, two competing goals need to be traded off: maximising robot performance while minimising human input.

The relative importance of the two goals is dictated by the priorities and the constraints of the mission. Examples for mission priorities are safety and efficiency, examples for mission constraints are the number of humans and robots available and communication bandwidth. To incorporate mission priorities and constraints, the two metrics mentioned above are weighted accordingly and summed up into a single scalar. The scalar can be interpreted as a mission-specific measure of human-robot team effectiveness. It can subsequently be used to find the appropriate autonomy level for a given mission scenario. We define *autonomy level* as a mission-specific parameter which determines the degree of autonomy the robot has during its deployment.

The methodology described above is split into five steps:

1) Identify robot performance metrics for a given task.
2) Experimentally obtain robot performance values and number of human-robot interactions as a function of the autonomy level.
3) Identify the constraints of the mission such as available resources (*e.g.* number of humans and robots), and mission priorities (*e.g.* safety, efficiency).
4) Combine robot performance metrics and the number of human-robot interactions into a scalar called *team effectiveness* using a weighted sum. Weights are chosen based on the constraints found in the previous step.
5) Select the autonomy level with the highest team effectiveness value for the mission.

The methodology is applied to a communication system developed for task-oriented information exchange [8]. The system fuses information from humans and robots to yield better informed decisions. Human operators are treated as a resource which needs to be managed carefully. Keeping the number of required interactions at a minimum has two advantages: humans will give higher quality input, and the number of robots that can be operated simultaneously will increase.

Our communication system makes use of probabilistic representations commonly used in robotics to model sensing and action uncertainties. Typically, sensors make observations to update the representation's probability distributions. We extend this by letting human operators contribute observations on a higher abstraction level [7] exploiting human cognitive abilities. An operator is thus treated as an additional information source.

Decision theory is used to compute decisions for the robot given all available evidence. Evidence is gathered from sensors and human operators: using Value-Of-Information (VOI) theory, it is possible to calculate the expected gain in consulting an information source which comes at a cost. This principle is used to decide whether an operator should be queried for an observation. In our system, the *cost* of obtaining information represents the aforementioned *autonomy level*. Both terms are used interchangeably throughout the rest of this paper.

The authors are with the ARC Centre of Excellence for Autonomous Systems (CAS), The University of Sydney, NSW 2006, Australia. {t.kaupp, a.makarenko}@cas.edu.au

978-1-4244-1647-9/08/$25.00 ©2008 IEEE.

The influence of the autonomy level on robot performance and the number of human-robot interactions is experimentally determined using a navigation task as an example. The robot navigates through a set of simulated worlds while higher-level human input is simulated by an external software component. The output of the experiment is a team effectiveness graph which is used to select an autonomy level for three example missions.

The remainder of the paper is structured as follows. Sec. II reviews related work. Sec. III describes the human-robot communication system which is demonstrated using a navigation task in Sec. IV. Results from applying the proposed methodology to that task are presented in Sec. V. Finally, Sec. VI concludes and lists future work.

II. RELATED WORK

Others have addressed the problem of finding metrics to evaluate human-robot systems more formally [3][13][10]. In their work, one of the goals of measuring the effectiveness of a human-robot team is to enable the prediction of an appropriate number of robots a user can effectively operate for a given task ("Fan-Out"). In contrast, we measure effectiveness to determine an appropriate autonomy level for an adjustable autonomy system.

Crandall *et al.* propose a set of *metric classes* applicable to all parts of a human-robot system: operators, individual robots, and the overall team [3]. The methodology proposed in the Introduction falls in the metric class of *Interaction Efficiency* measuring, among other things, how human input affects robot performance.

The fields of Adjustable Autonomy (AA) and Mixed Initiative Control aim at bridging the gap between full human control and full autonomy. The fundamental questions in these fields are how to decide when to relinquish control and on what criteria to base that decision [12]. In many AA systems, the human operators are in charge of switching between a set of predefined discrete modes which imposes a significant responsibility on the operator as pointed out in [2].

In contrast, the approach advocated here is to let the robot decide when to query operators for input based on the uncertainty in the robot's beliefs. This can be seen as a robot-initiated shift to lower autonomy at runtime. How often this shift occurs depends on the previously set autonomy level (a design parameter). How to find an appropriate autonomy level is experimentally demonstrated in Sec. V.

Our human-robot communication system is most closely related to Fong's *Collaborative Control* which, like our approach, treats humans as a resource for robots [5]. Bidirectional communication in the form of a human-robot dialog is used to exchange information of different types such as commands, queries and responses. The main difference to our work is the chosen approach: while Fong uses no specific underlying mathematical method, we cast the problem of collaborative control in a decision-theoretic formulation. This allows a quantitative analysis of the system as presented in this paper whereas Fong focuses on qualitative measures and usability [5].

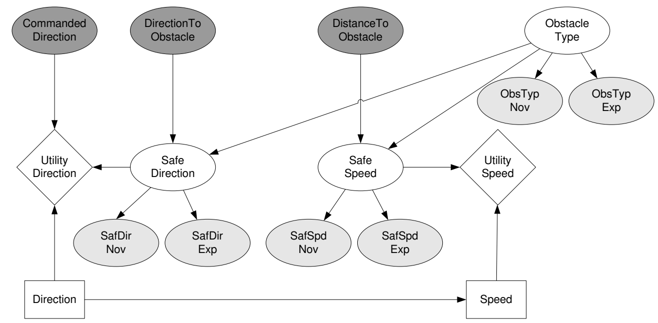


Fig. 1. Influence Diagram (ID) representation for a navigation task. Ovals are chance nodes, squares are decision nodes, and diamonds are utility nodes. Grey nodes represent information sources: dark/light grey nodes are observed by the robot/human.

Using VOI theory to decide what information source to query has been applied to a wide range of applications, *e.g.* distributed sensor management [9]. To the best of our knowledge, this work is the first to apply VOI theory to human-robot communication.

III. HUMAN-ROBOT COMMUNICATION

This section explains the mathematical background of our human-robot communication system. A probabilistic representation lies at the heart of the approach.

A. Decision-Theoretic Models

A probabilistic representation should encode the relationships between random variables qualitatively (model structure) and quantitatively (model parameters), and allow efficient inference. A class of graphical models fulfilling these requirements are *Bayesian Networks* (BNs) [11]. BNs encode beliefs about the world states represented as *chance nodes*. Chance nodes are either observed (in which case they are also called evidence nodes) or unobserved (in which case they contain probability distributions). Nodes that can be observed are referred to as *information sources* in this paper.

BNs can be extended to *Influence Diagrams* (IDs) to model decision making under uncertainty [6]. IDs are generally able to represent information about the current state, possible actions, the state resulting from the action, and the utility of that state [11].

IDs extend BNs by adding *decision* and *utility* nodes. Decision nodes represent choices available to the decision-maker (a set of possible actions). Utility nodes encode a utility function: the usefulness of the consequences of decisions using a scalar called utility.

An example of an ID is shown in Fig. 1 which is the model used for the experiment presented in Sec. V. It encodes a robot's low-level decision model which chooses driving actions based on the beliefs of the chance nodes, and the utility functions. In this example, two decisions are made: in which direction to move (decision node *Direction*), and with what speed (decision node *Speed*).

Each decision node has several discrete actions: choices for *Direction* are $\{$*left, right, straight*$\}$ while choices for

Speed are $\{stop, slow, fast\}$. The best decisions are the ones which maximise the expected utilities which are defined by the two utility nodes.

More formally, consider a decision node D and let e denote the set of observations (evidence). A rational decision-maker chooses the action d which maximises the expected utility (EU):

$$d = \underset{D}{\operatorname{argmax}}\, EU(D|e) \tag{1}$$

B. Value of Information

Rather than taking the action d in Eq. 1, a decision-maker might have the choice of consulting one of its information sources $\{I_1, ..., I_m\}$ in order to generate a more informed decision. Consulting an information source I_i is equivalent to obtaining the state of that chance node. We assume that only a single information source is consulted at any given time which is referred to as *myopic* information gathering [4]. In our example, grey nodes represent all information sources with light grey nodes representing human-observable nodes.

It is possible to calculate what we can *expect* to gain from consulting the information source *before* observing that node by using its belief given all current evidence $P(I_i|e)$ [4]. The expected utility of the optimal action (EUO) *after* having observed I_i is

$$EUO(I_i, D|e) = \sum_{I_i} P(I_i|e) \max_D EU(D|e, I_i) \tag{2}$$

The value of observing I_i is called the *Value Of Information* (VOI). It is calculated as the difference between the expected utility after having observed I_i and the currently available maximum expected utility:

$$VOI(I_i, D|e) = EUO(I_i, D|e) - \max_D EU(D|e) \tag{3}$$

For the representation shown in Fig. 1, the outcomes of the decisions are independent of the chance nodes in the model, *i.e.* decision nodes have no chance nodes as children. For this special case, it is straightforward to compute Eq. 3 (H denotes the parents of the utility node U):

$$\begin{aligned} VOI(I_i, D|e) = &\sum_{I_i} P(I_i|e) \max_D \left(\sum_H P(H|I_i, e) U(D, H) \right) \\ &- \max_D \left(\sum_H P(H|e) U(D, H) \right) \end{aligned} \tag{4}$$

Computing the VOI is relevant for intelligent information gathering systems where the goal is to maximise the amount of information collected [16]. Consulting an information source comes at a cost, so a sensible strategy is to consult that source only if the expected benefit is higher than the cost $C(I_i)$:

$$VOI(I_i, D|e) - C(I_i) > 0 \tag{5}$$

An experimental methodology to find an appropriate cost parameter C was proposed in Sec. I and is applied to a navigation task in Sec. V.

C. Using VOI to Adjust Autonomy

This section explains how VOI is used to query human operators and therefore adjust the robot's autonomy online. In our approach, humans are treated as information sources which can be queried for observations. VOI theory is used to determine under what circumstances to query operators which can be seen as an online shift to lower autonomy. Whether an autonomy shift is triggered during a mission depends on the robot's current probabilistic beliefs and the cost parameter of Eq. 5.

In the simplest case (as discussed here), the cost parameter is fixed prior to a mission. If the cost parameter value is high, the robot relies more on its own perceptual capabilities, asks fewer questions, and is more autonomous by definition. Setting the cost parameter can thus be seen as fixing the *autonomy level* which was defined in Sec. I.

IV. NAVIGATION APPLICATION

This section demonstrates the human-robot communication system using a mobile robot navigation task. The implementation was designed to provide a flexible testbed for multiple experiments, one of which is reported in Sec. V.

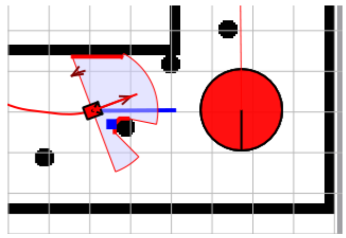
The navigation algorithm makes use of the representation shown in Fig. 1. Many well-researched methods in machine learning and knowledge engineering address the problem of how to construct such a representation [11]. It is not the focus of this work, and therefore both the structure and the parameters are handcrafted for our representation. Since all the nodes are discrete, the probability distributions and utility functions are represented as tables.

At robot run-time, the representation is instantiated at each time instance. Several steps are involved to yield driving actions which are described next.

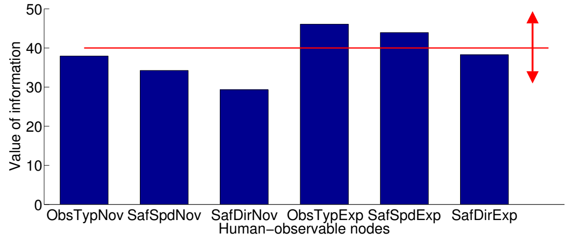
First, the robot determines the direction to the next waypoint which serves as evidence for the *CommandedDirection* node. Then, obstacle states are extracted from the current laser scan which serve as evidence for *DirectionToObstacle* and *DistanceToObstacle* representing the direction and distance to the closest obstacle (if any).

The next step is to compute the VOI for all human-observable nodes according to Eq. 3. Human-observable nodes are children of the *latent* (unobserved) variables *ObstacleType* (pushable or fixed) and *SafeDirection*/*SafeSpeed*. Two types of human-observable nodes (novice/expert) are used to represent the different levels of uncertainty in the operators' answers: in plain terms, how much the answers can be trusted. Mathematically, the nodes encode conditional probability distributions (CPDs) which we call *Human Sensor Models* (HSMs).

HSMs are important for the VOI analysis: the more an answer can be trusted, the more valuable it is. An example is visualised in Fig. 2: the robot is in a "critical" situation because there is an obstacle very close to its right which



(a) Robot navigating in a maze



(b) VOI analysis

Fig. 2. Robot in a "critical" situation: (a) it detected a close obstacle to its right which is also the direction it is supposed to move (blue marker points towards next waypoint, red arrows indicate current speed and turnrate); (b) VOI analysis for this situation showing the VOI for all human-observable nodes. The horizontal line represents the cost of consulting a human operator. How to set this design parameter appropriately is the topic of this paper.

is also the direction it is commanded to move (Fig. 2(a)). The VOI analysis for this situation is pictured in Fig. 2(b) showing the VOI for all six human-observable nodes. The figure also shows that information from an expert is more valuable than from a novice.

The horizontal line in Fig. 2(b) shows the variable cost of obtaining information. An appropriate cost level will be determined experimentally in Sec. V. In this scenario, two bars exceed the cost: *ObsTypExp* and *SafSpdExp*, *i.e.* Eq. 5 is fulfilled for these two information sources. Thus, the robot would consult an expert about the type of obstacle which is the most valuable observation it can possibly obtain.

When a human observation is received, it is incorporated into the representation as additional evidence. Finally, a driving action is selected according to Eq. 1.

V. EXPERIMENT

This section follows the five steps of the methodology proposed in Sec. I and applies it to the navigation task.

A. Performance Metrics (Step 1)

Two robot performance metrics are used here: the number of successfully completed navigation tasks and the time to complete each (successful) task. The two metrics are proposed in [13] to measure effectiveness and efficiency of a navigation task.

B. Experiment (Step 2)

1) Setup: To obtain a statistically relevant dataset, the following approach is taken. Human answers are simulated by running a software component which generates better informed observations than the robot itself is able to produce. This is achieved by restricting the robot's laser scan to a field

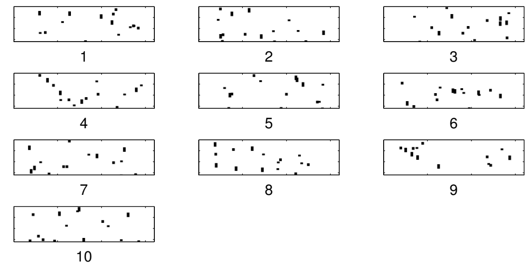


Fig. 3. Ten environments of size $64m \times 16m$ with randomly generated obstacles.

of view of only 40° but allowing the "simulated human" access to a 120° scan.

We argue that this is a valid approach since humans can often extract more information from the same sensor data by applying their rich perceptual capabilities. Humans can also apply background information not available to today's robots. For the purpose of this experiment, human and robot capabilities are artificially reduced.

The simulated human acts as an expert who, whenever requested, is able to produce *SafDirExp* and *SafSpdExp* observations based on a 120° laser scan. These two variables correspond to the 5^{th} and 6^{th} bar in Fig. 2.

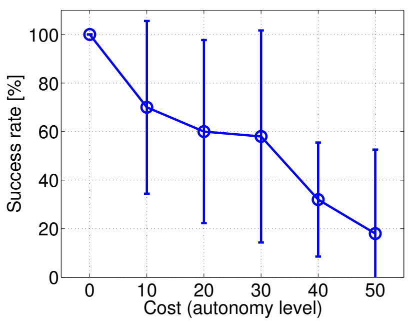
The experiment is conducted using 10 environments of size $64m \times 16m$ with randomly generated fixed obstacles as shown in Fig. 3. The robot is placed near the centre on the far left of the world and a goal waypoint is generated near the centre on the far right (with some randomness). An episode is declared successful if the robot manages to position itself within $1m$ of the goal before a timeout of $120s$ occurs. Timeouts occur if the robot hits an obstacle and as a result gets stuck.

The evaluation uses 6 cost parameters which are uniformly distributed over the range of possible VOI values. Each of the 10 environments is traversed 5 times resulting in a dataset of size 300. The experiment ran continuously for a total of 8.7 hours. All software components were implemented and deployed using the Orca software framework [1].

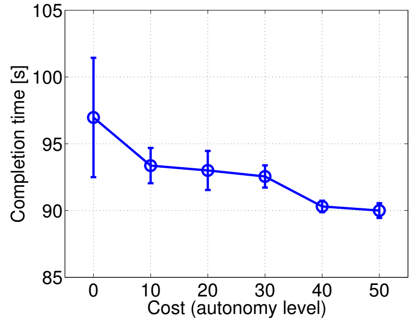
2) Results: Fig. 4(a) & 4(b) show success rate and completion time as a function of the cost parameter. The success rate drops off with increasing autonomy (fewer queries). At cost 0, when the simulated human is queried continuously, all 50 episodes are successful. At cost 50 the robot is autonomous at all times (no queries) and completes the course successfully in only 18% of the cases, on average. The plot shows that robot performance increases if input from an expert can be obtained.

The completion time is only measured for successful episodes, *i.e.* when the robot did not hit an obstacle. Completion time is highest for cost 0 when the simulated human is queried continuously. This is expected since the robot carefully avoids all obstacles. The high variance at cost 0 is also expected since more time is needed to navigate in more difficult environments.

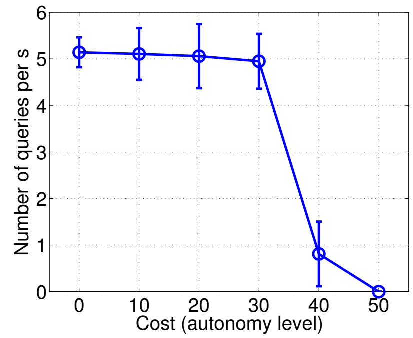
At higher autonomy, the robot takes less time to reach its



(a) Success rate



(b) Completion time



(c) Number of operator queries

Fig. 4. Team effectiveness metrics ((a),(b) robot performance; (c) number of interactions) as functions of cost, obtained experimentally. Variance in results is due to the varying difficulty levels of the environments. Error bars indicate one standard deviation.

goal which seems surprising at first. However, the number of successful episodes decreases as mentioned above, and only environments with few obstacles along the centre are traversed successfully. The robot is faster if it does not hit an obstacle because it steers "blindly" towards its goal without spending time to avoid obstacles.

Finally, Fig. 4(c) shows the number of operator queries per second as a function of the cost. The number of queries decreases gently first, then drops off steeply and finally goes to 0 which represents fully autonomous operation. The steep drop-off can be explained by the coarse representation used in this experiment: obstacles are either far, close, or very close. Situations in which obstacles are far away are much more numerous than situations in which they are close. At cost 30, the robot queries the human even in relatively safe situations when obstacles are far resulting in many more queries than at cost 40.

Even at high costs, the number of queries is too high for a practical deployment with real humans. Obstacle tracking and wall recognition algorithms can be used to avoid repeated questions about the same obstacle (as done for the user study reported in [8]). This functionality was turned off to simplify the experiment without having any impact on the conclusions to be drawn.

C. Identification of Constraints (Step 3)

Three mission scenarios are presented as examples for step 3 of the proposed methodology:

1) Multiple operators, robot is expensive: this scenario reflects the remote operation of an expensive robot, *e.g.* in a search and rescue mission. We assume the robot's environment is inaccessible to humans, so operators cannot physically intervene. It is also assumed that many human operators can be contacted, experts as well as non-experts, *e.g.* through a web application[1].

2) Single operator, multiple robots: the second scenario assumes a single operator who is in charge of multiple robots at the same time. It is assumed that the operator can only attend to one robot at the time. The operator's capacity is the limiting factor here.

3) Multiple operators, robot is expensive, limited communication bandwidth: this scenario is similar to scenario 1 with the additional assumption of a limited communication bandwidth. An example application is a remote planetary exploration mission.

D. Parameter Selection (Steps 4 & 5)

Step 4 requires the combination of success rate, completion time, and the number of queries into a single scalar which we call *team effectiveness*. The combination is achieved by a weighted sum which is subsequently scaled to a $[0; 1]$ interval. The weights represent the relative importance of each variable and are set according to the mission priorities and constraints identified in step 3.

For scenario 1, the completion time is of secondary importance but the robot could be lost if it hits an obstacle. A weight ratio of $3 : 1 : 1$ is manually chosen emphasising the importance of success over completion time and number of queries. Fig. 5(a) shows the team effectiveness for all cost parameters. The highest team effectiveness is achieved by setting the cost parameter to 0.

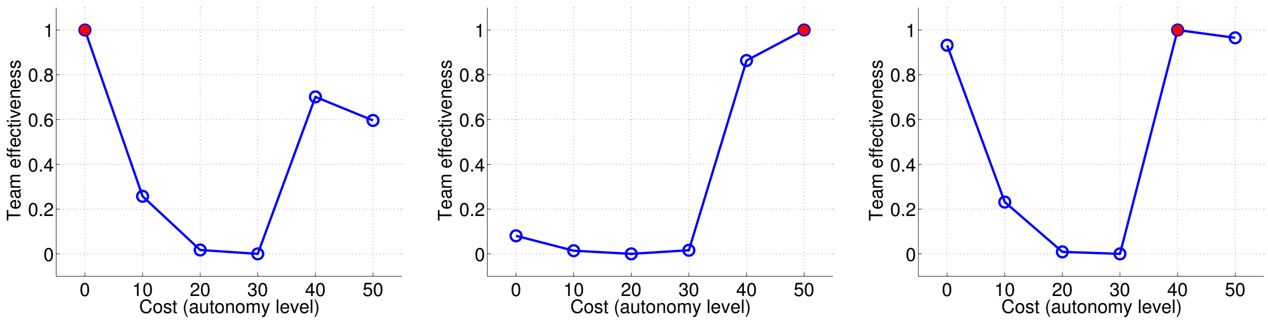
For scenario 2, the weight ratio is chosen to be $1 : 1 : 3$, reflecting the importance of minimising the number of queries and thus the operator workload. The cost parameter with highest team effectiveness is 50 for this case as shown in Fig. 5(b). Selecting this parameter implies the acceptance of a robot performance loss.

For scenario 3, the weight ratio is $3 : 1 : 1.5$ which is similar to scenario 1 but penalises human queries more to avoid excessive communication. Fig. 5(c) shows the result: the highest team effectiveness is achieved when cost 40 is selected.

VI. CONCLUSIONS AND FUTURE WORK

This paper addressed the difficult problem of determining an appropriate autonomy level for an Adjustable Autonomy (AA) system. The chosen autonomy level determines under what circumstances the AA system adjusts its autonomy online. We proposed a methodology to measure the effectiveness of a human-robot team prior to its mission deployment.

[1] Amazon's Mechanical Turk [14], which is similar to our communication system, was used in September 2007 to search for a disappeared plane [15]. Within three days up to 50,000 people joined in the effort (without success).



(a) Weight ratio $3:1:1$ (b) Weight ratio $1:1:3$ (c) Weight ratio $3:1:1.5$

Fig. 5. Team effectiveness graphs generated by weighted sums of 3 variables (success rate, completion time, number of queries). Team effectiveness values are scaled to $[0;1]$. The most appropriate autonomy level is found by selecting the cost where team effectiveness is 1 (indicated by the red markers). Three mission scenarios are shown: (a) multiple operators, robot is expensive; (b) single operator, multiple robots; (c) similar to (a) but penalising excessive communication.

Two competing goals were identified: maximising robot performance while minimising the amount of human input. To trade off these goals, the constraints and priorities of the mission were taken into account.

We applied the methodology to our human-robot communication system which was developed for task-oriented information exchange. The system makes use of common probabilistic robotic algorithms and decision theory to yield rational decisions. Value-Of-Information (VOI) theory is used to determine when it is worthwhile querying human operators who are regarded as information sources. How much value an observation adds has to be compared to the cost of obtaining the information. For our system, the cost parameter is equivalent to the autonomy level.

The communication system and the proposed methodology were demonstrated using a navigation task. The results were used to find an appropriate cost parameter (autonomy level) for three example mission scenarios. Results of this experiment were also used as a guide to set the autonomy level for a follow-up experiment – a user study [8]. In the future, more user studies will be conducted to verify the estimated autonomy levels shown in Fig. 5 on unseen environments.

Experiments to identify human-robot systems are expensive and time-consuming which was mitigated in this paper by simulating human input. Future work will make use of Design of Experiment (DoE) methods to determine which experiments are most relevant. Another option is to develop (partial) models of the overall human-robot system which will also be addressed in the future.

In this paper, the goal was to find an appropriate cost parameter offline which is subsequently fixed at mission runtime. An interesting research avenue is to investigate the online adaptation of the cost parameter based on a user model of workload and expected time delays. This would result in a system which adjusts its autonomy online based on both the environment's and the human operator's states.

While only one-to-one interactions were demonstrated here, we are confident that the communication system can be applied to multi-robot multi-user systems. Future reseach will investigate if the proposed methodology from this paper can be extended to these more complex systems.

VII. ACKNOWLEDGEMENTS

This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

REFERENCES

[1] A. Brooks, T. Kaupp, A. Makarenko, S. Williams, and A. Orebaeck. Orca: A component model and repository. In D. Brugali, editor, *Software Engineering for Experimental Robotics*, volume 30 of *STAR*, pages 231–251. Springer, 2007.

[2] D.J. Bruemmer, D.A. Frew, C.N. Nielsen, and M.C. Walton. Mixed-initiative control for collaborative countermine operations. *Submitted to IEEE Transactions on Robotics*, 2007.

[3] W. Crandall and M.L. Cummings. Developing performance metrics for the supervisory control of multiple robots. In *2nd Annual Conference on Human-Robot Interaction*, Washington, D.C., 2007.

[4] S.L. Dittmer and F.V. Jensen. Myopic value of information for influence diagrams. In *13th Conference on Uncertainty in Artificial Intelligence (UAI '97)*, 1997.

[5] T. Fong, C. Thorpe, and C. Baur. Robot, asker of questions. *Robotics and Autonomous Systems*, 42(3-4):235–243, 2003.

[6] R.A. Howard and J.E. Matheson. Influence diagrams. In R.A. Howard and J.E. Matheson, editors, *Readings on the Principles and Applications of Decision Analysis*, volume 2 of *Strategic Decisions Group*. 721–762, Menlo Park, California, 1984.

[7] T. Kaupp, B. Douillard, F. Ramos, A. Makarenko, and B. Upcroft. Shared environment representation for a human-robot team performing information fusion. *Journal of Field Robotics: Special Issue on Teamwork*, 24(11–12):911–942, 2007.

[8] T. Kaupp and A. Makarenko. Decision-theoretic human-robot communication. In *3rd ACM/IEEE Conference on Human-Robot Interaction (HRI'08)*, 2008.

[9] J. Manyika and H. Durrant-Whyte. *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. Prentice Hall, Upper Saddle River, NJ, USA, 1995.

[10] D.R. Olsen and M.A. Goodrich. Metrics for evaluating human-robot interactions. In *Proceedings of NIST's Performance Metrics for Intelligent Systems Workshop*, 2003.

[11] S. Russel and P. Norvig. *Artificial Intelligence – A Modern Approach*. Prentice Hall, 2nd edition, 2003.

[12] P. Scerri, D.V. Pynadath, and M. Tambe. Towards adjustable autonomy for the real world. *Journal of Articial Intelligence Research*, 17:171–228, 2002.

[13] A. Steinfeld, T.W. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich. Common metrics for human-robot interaction. In *2006 Human-Robot Interaction Conference*. ACM, March 2006.

[14] Wikipedia. Amazon Mechanical Turk, 2007.

[15] Wikipedia. Disappearance of Steve Fossett, 2007.

[16] S. Zilberstein and V. Lesser. Intelligent information gathering using decision models. Technical Report 96-35, Computer Science Department, University of Massachusetts, 1996.