Task Fidelity: a new metric for measuring task complexity involving robots

Michael Vallance, Takafumi Yamamoto, Yuta Goto, and Kenta Ibayashi

***Abstract*— There is no consensus regarding a common set of metrics for robot task complexity and associated human-robot interactions. In our research, tasks involving students in Japan and UK interacting in a 3D virtual world to collaboratively program robots to solve maze problems have resulted in quantitative data of immersion, Circuit Task Complexity and Robot Task Complexity which have subsequently been collated to create a proposed new metric for tasks involving robots, which we have termed Task Fidelity.**

***Index Terms*—** collaboration, metrics, RMI, robot task complexity, task fidelity, virtual world.

I. INTRODUCTION

In any task design it is important to consider its difficulty for the intended learners. Task designers such as teachers and Higher Education practitioners need to provide tasks commensurate with the expected successful outcomes that will, it is anticipated, be developed by the learners. In this paper we will demonstrate how tasks can be quantified within the particular context of communicating the programming of robots in a 3D virtual world. Circuit Task Complexity and Robot Task Complexity will be calculated alongside immersivity to determine a new metric for measuring tasks involving robots, which we have termed Task Fidelity.

Common metrics allow for benchmarking within a particular domain. For instance, road transportation such as cars, motorcycles and trucks can be compared by the metrics of top speed, acceleration, engine capacity, fuel economy, transmission and price. However, this is not the case with robots. Although there are a number of metrics which can be related to robot-related tasks and the complexity of the domain where robots are utilized, Steinfeld *et al* [1] state that "the primary difficulty in defining common metrics is the incredibly diverse range of human-robot applications" (p.33). When discussing robots that undertake specific maneuvers, some researchers provide a common metric labeled as Task Complexity. Tasks are defined as physical action units that are undertaken by a robot, and the designation 'complexity' is used to characterize the task that consists of parts in, potentially, intricate arrangements.

In an example of robots which maneuver around obstacles and follow distinct circuits (or mazes), Barker and Ansorge [2] derive Task Complexity as TC = Σ S + time, where Σ S is the number of sections or turns of a maze. Olsen and Goodrich [3] define Task Complexity as TC = TE + IE, where Task Effectiveness (TE) = the number of commands successfully programmed into the robot, and Interaction Effort (IE) = the amount of time required to interact with the robot (to take into account mistakes). Olsen and Goodrich [3] also suggest TC = NT, where NT = Neglect Tolerance which measures autonomy of robot (i.e. measures how robot's effectiveness declines over time when the robot is neglected by the user). In developing the metric, Olsen and Goodrich [3] additionally offer TC = RAD (Robot Attention Demand) which measures the total time user must interact with the robot, where RAD = IE/ (IE + NT). If the tele-operated robot has a small NT and RAD approaches 1, then the user can focus on other things

In the domain of Artificial Intelligence and development of robots, Russel and Norvig [4] determine Task Complexity as TC = PEAS, where P = Performance, E = Environment, A = Actuators, S = Sensors. The metric though is descriptive and does not amount to a numerical number.

The USUS Evaluation Framework for Human-Robot Interaction by Weiss *et al.* [5] focus upon usability, user experience, social acceptance and social impact. Usability is the extent to which a robot can be used by specified users to achieve specific goals with effectiveness, efficiency and satisfaction in a particular context of use. Metrics include effectiveness (i.e. task completion rate), efficiency (i.e. speed at which task is completed), learnability (i.e. how easy the system can be learned by human users), flexibility (i.e. the number of different ways users can communicate with the system), robustness (i.e. the level of support provided, and utility (i.e. the number of tasks the interface is designed to perform). The outcome is descriptive, with no numerical representation.

Murphy and Schreckenghost [6] conducted a meta-analysis of 29 papers that proposed metrics for human-robot interaction. Forty two metrics in total were found. They determined that the metrics were categorized to the object being directly measured; such as the human (N=7), the robot (N=6), or the system (N=29). The systems' metrics were found to be subdivided into productivity, efficiency, reliability, safety and coactivity. They found that the metrics were often not measured directly but most often were inferred through observation. The paper identified proposed metrics but it was recognized that they "have no functional, or generizable, mechanism for measuring that feature."

In summary, there is no consensus regarding a common set of metrics for robot task complexity and its associated human-robot interactions. Although many attempts have been made to develop a taxonomy of metrics, the community has yet to develop a standard framework, and many metrics are task-specific. Therefore, we feel justified in developing our own complexity metric, termed Task Fidelity, specific to the context of our robot-mediated communication discussed in this paper.

II. Robot Tasks in a 3D Virtual World

Our research has been designed to collate data of students collaborating in a 3D virtual world to program a LEGO robot to successfully navigate mazes from start to completion in both the physical world and within a 3D virtual space (see Fig. 1). This is undertaken by (i) designing circuits which necessitate the use of robot maneuvers and sensors; and (ii) experiencing collaboration in a virtual world between students in Japan and UK. These experiences lead to personal strategies for teamwork, planning, organizing, applying, analyzing, creating and reflection. Complex problems are thus presented which necessitate the use of programming skills, collaboration, and cognitive experiences.

In our research we have re-named Task Complexity as 'Robot Task Complexity' (RTC) because the task focuses upon the robot and what the human has to do to manipulate that robot. We call this the 'product' of a robot task. Human-Robot Interaction. (HRI) is the 'interaction' between a human and a robot. The word 'interaction' assumes that the human and the robot are communicating two-way. We call this the 'process' of a robot task. Therefore, we do not consider our process to be Human-Robot Interaction (HRI) but consider our collaboration to be Robot-Mediated Interaction (RMI).

Figure 1. UK and Japan students as avatars collaborating in our 3D virtual world

III. Circuit Task Complexity

In our first iteration (*cf.* Vallance and Martin), in order to quantify each task complexity the programming of the LEGO robot required a determination of an action and a vector [7]. Given the specific purposes of the robot in our research, we utilized the eminent work in robotics by Barker and Ansorge [2] and also Olson and Goodrich [3]; where task complexity is calculated according to the number of sections that make up a given maze. We called this Circuit Task Complexity (CTC) which equals the number of directions (d) + number of maneuvers (m) + number of sensors (s) + number of obstacles (o).

$$CTC = \Sigma\,(d + m + s + o)$$

For example, in Fig. 2 the robot must maneuver within a maze including 2 obstacles in order to reach its target. One interpretation of the problem can be: the number of directions to be programmed is 4, the number of maneuvers is 3, and the number of sensors is 2 (i.e. two touch sensors).

$$CTC = \Sigma\,(d + m + s + o)$$
$$CTC = \Sigma\,(4 + 3 + 2 + 2) = 11$$

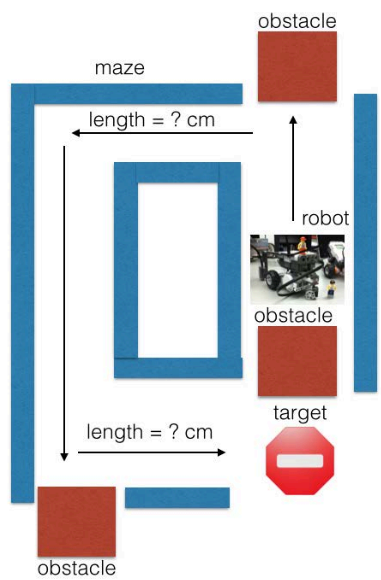

Figure 2. Example of Circuit Task Complexity.

IV. Robot Task Complexity

However, we found that the logic of assigning task complexity to circuits was inadequate. For instance, initially we assigned complexity values to distinct maneuvers such as forward – turn – back. We found over the course of our previous research that as circuits became more challenging, the Mindstorms NXT programming became more complex. This was especially the case when we needed to add sensors to maneuver around and over obstacles. Simply adding the number of obstacles to the Circuit Task Complexity was flawed because the programming required to maneuver over a bridge using touch sensors, for instance, was far more complex than maneuvering around a box using touch sensors. Consequently, we modified our task complexity to be determined by the NXT program solution rather than the circuit to be navigated. We call this Robot Task Complexity (RTC), which is measured as:

$$RTC = \Sigma Mv_1 + \Sigma Sv_2 + \Sigma SW + \Sigma Lv_3$$

where,
M = number of moves (direction and turn)
S = number of sensors
SW = number of switches
L = number of loops

where v = number of decisions required by user for each programmable block
$v_1 = 6$
$v_2 = 5$
$v_3 = 2$

In the NXT Mindstorms software, the Move block controls the direction and turns that the LEGO robot will take. There are six variables that need to be considered: NXT 'brick' port link, direction, steering, power, duration, and next action. In other words, the students have to make six specific decisions about the values which make up the programmable block. Therefore, we assign v_1 a value of 6. There are eight common sensors which are used in our tasks (timer, light, ultrasonic, color, touch, sound, distance, wait) with the sensors' capabilities determined by 5 variables (so we assign $v_2= 5$). Although some sensors have 6 decisions built in and some have 5, the difference is that the extra decision is simply cosmetic as in 'speak an alert' so does not impact on the robot's performance or capability to complete the task. All sensors are tagged as S. A loop has only two variables to consider so we assign $v_3 = 2$.

Given the circuit shown in Fig. 2 above, the robot has to be programmed to move in 4 directions, with 3 turns and 2 touch sensors. A NXT program solution in Fig. 3 can then be used to calculate the Robot Task Complexity.

$$RTC = \Sigma Mv_1 + \Sigma Sv_2 + \Sigma SW + \Sigma Lv_3$$
$$RTC = (8 \times 6) + (3 \times 5) + 0 + 3$$
$$RTC = 66$$

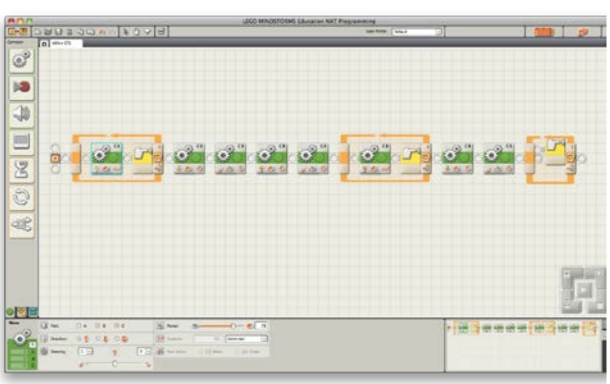

Figure 3. LEGO NXT program solution

We acknowledge that, at present, our modified Robot Task Complexity metric applies only to the LEGO Mindstorms robot, but it does provide a useful indicator in our attempts to analyze the experiential learning during the collaborative tasks. We are currently experimenting with LabView 2011 software with NXT module, and the LEGO EV3 software.

V. TASK FIDELITY

Given that our context is Robot-Mediated Interaction (RMI) in a 3D virtual space, this applied research can also determine how immersed students become within the process of each task. To record 'immersion' (a cognitive phenomena also referred to as 'flow' (*cf.* Csikszentmihalyi & Nakamura [8])), data are collected from the students during and after each task. Questions were chosen based upon research in immersivity by Pearce *et al.* [9]. With optimal challenge – skill relationship, the students become immersed in the Robot-Mediated Interaction (RMI) tasks.

The collated data included total challenge and skill values, Circuit Task Complexity values, and Robot Task Complexity values (*cf.* Vallance *et al.* [10]). In order to compare the data from twenty eight (28) tasks it was necessary to scale all the values between 0 and 1. For instance, for the challenge and skill values, in each task we simply divided the sum scores of the students by the maximum score possible. For the Circuit Task Complexity values we took the maximum CTC value and divided it into each CTC value. Similarly, for the Robot Task Complexity values we took the maximum RTC value and divided it into each RTC value. All values are thus represented between 0 and 1. This allows us to represent the data graphically and thus determine the immersion in the case of challenge and skills, and Task Fidelity (TF) (see below) in the case of Circuit Task Complexity and Robot Task Complexity values.

Consequently, the complexity of the task can now be quantified by a new metric which we term Task Fidelity. For example, from the data discussed in Vallance *et al.* [10], the graph in Fig. 4 of Circuit Task Complexity versus Robot Task Complexity graphically reveals the plotted differences in the researcher's (in the role of instructor or teacher) expected level of complexity (i.e. the Circuit Task Complexity) and the students' achievement (i.e. the Robot Task Complexity). Ideally one would expect the two plotted areas to merge; in other words, the researcher (or teacher) provides a task commensurate with the expected successful outcome developed by the learners. Looking at the graph in Fig. 4 this assumption mostly appears to be the case. However, we can also numerically represent the differences between anticipated task complexity and successful accomplishment. This is called Task Fidelity, and is calculated as:

Task Fidelity = Circuit Task Complexity - Robot Task Complexity

$$TF = CTC – RTC$$

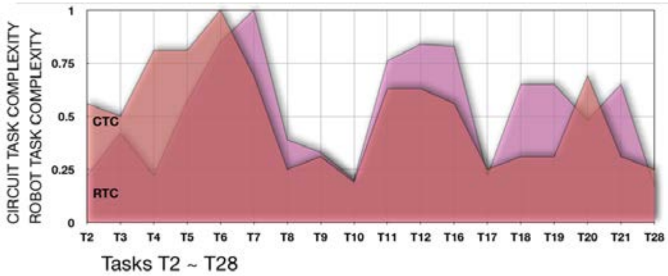

Fig. 4 Graph of Circuit Task Complexity and Robot Task Complexity for Tasks 2 ~ 28.

Figure 5 portrays the results of Task Fidelity plotted against the order of tasks of increasing challenge from the data discussed in Vallance *et al.* [10].

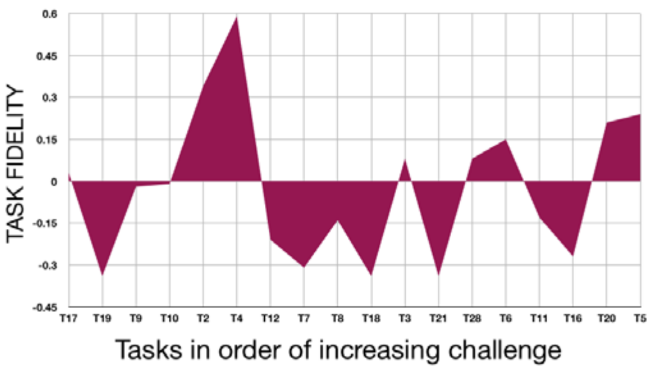

Fig. 5 Task Fidelity, where tasks have been rearranged in order of increasing challenge.

To reiterate, Task Fidelity is an indicator of the complexity of the circuit compared with the complexity of the program to complete the circuit. The zero line indicates ideal Task Fidelity; or ideal task complexity. Data plotted above the zero line indicate that the robot program was more complex than the circuit the robot had to maneuver. Data below the zero line indicate that the circuit was more complex than the robot program required to successfully navigate it. Our data reveal that for most tasks the programming required to complete the circuits was less than the considered complexity of the circuit (i.e. most data points are below the zero line of Task Fidelity). This appears to be the case across the range of challenges indicated by the students. The exceptions are T2, T4, T5 and T20 where TF values are above the zero line of Task Fidelity. These tasks involved sensors. Programming of sensors is indeed more complex for the students and this was also reflected in the immersion data mentioned above; students were most anxious when engaged in tasks requiring sensor programming and were thus less immersed in the challenge. However, as their skills of sensor programming increased, immersivity increased as indicated by Task 28 where Japanese students were taught by UK students within our 3D virtual space to program the robot's use of light and color sensors to initiate specific actions. TF value for T28 was only + 0.08; slightly above the optimal line. The challenge is to seek tasks similar to T28 where immersivity is close to or on the optimal path of immersivity, and task complexity is close to or on the optimal line of Task Fidelity.

VI. Conclusion

In conclusion, the literature reveals that there is no consensus regarding a common set of metrics for robot task complexity and its associated human-robot interactions. Circuit Task Complexity and Robot Task Complexity have thus been calculated alongside immersivity to determine a new metric for measuring tasks involving robots, which we have termed Task Fidelity. We will continue to implement the metric in diverse robot scenarios within our 3D virtual space involving synchronous collaboration between students in Japan and UK.

References

[1] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, Common Metrics for Human-Robot Interaction. *In proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pp. 33-40, 2005.

[2] S. B. Barker and J. Ansorge. "Robotics as means to increase achievement scores in an informal learning environment." *Journal of Research in Technology and Education*, Vol. 39, No. 3, pp. 229-243, 2007.

[3] D. R. Olsen and M. A. Goodrich. (2003). Metrics for evaluating human-robot interactions. [Online]. Available: http://icie.cs.byu.edu/Papers/RAD.pdf

[4] S. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach, 3rd Edition.* USA: Prentice Hall, 2010.

[5] A. Weiss, R. Bernhaupt, M. Lankes, and M. Tscheligi, The USUS evaluation framework for human-robot interaction. *In Proceedings of the Symposium on New Frontiers in Human-Robot Interaction*, pp. 158-165, 2009.

[6] R. Murphy and D. Schreckenghost, "Survey of metrics for human-robot interaction," in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction***,** Tokyo, pp.197-198, 2013.

[7] M. Vallance and S. Martin. "Assessment and Learning in the Virtual World: Tasks, Taxonomies and Teaching For Real." *Journal of Virtual Worlds Research* Vol. 5, No. 2, 2012.

[8] M. Csikszentmihalyi and J. Nakamura, "Effortless attention in everyday life: A systematic phenomenology," in Effortless attention: a new perspective in the cognitive science of attention and action, B. Bruya, Ed. Cambridge MA: MIT Press, pp. 179-189, 2010.

[9] M. Pearce, M. Ainley and S. Howard. "The ebb and flow of online learning." *Computers in Human Behavior*, Vol. 21, pp. 745–771, 2005.

[10] M. Vallance, C. Naamani, M. Thomas, and J. Thomas, Applied Information Science Research in a Virtual World Simulation to Support Robot Mediated Interaction Following the Fukushima Nuclear Disaster. *Communications in Information Science and Management Engineering (CISME).* Vol. 3, no. 5, pp. 222-232, 2013.

Michael Vallance Ed.D, MSc. PGCE, BSc (Hons) is an academic in the Department of Media Architecture, Future University Hakodate, Japan. He has been involved in educational technology design, implementation, research and consultancy for over fifteen years, working closely with Higher Education Institutes, schools and media companies in UK, Singapore, Malaysia and Japan. He was awarded second place in the Distributed Learning category of the United States Army's 2012 Federal Virtual Worlds Challenge for his research in virtual collaboration and robots. His website is at http://www.mvallance.net Email michael@fun.ac.jp
.

Takafumi Yamamoto, Yuta Goto, and **Kenta Ibayashi** are undergraduate students at Future University Hakodate, Japan.