Issues in Assessing Performance of Social Robots

BRIAN R. DUFFY[1], GINA JOUE[2], JOHN BOURKE[3]
[1,3]Media Lab Europe (MLE),
Sugar House Lane, Bellevue, Dublin 8
[2]Department of Computer Science,
University College Dublin, Belfield, Dublin 4
IRELAND
[1]brd@media.mit.edu [2]gina.joue@ucd.ie [3]Johnb@mle.media.mit.edu http://www.medialabeurope.org

Abstract: - How can the performance of a social robot be measured? Are existing metrics and techniques appropriate for physically and socially embodied autonomous mobile robots? With the rapid expansion of robust robotic systems from strong industrial applications to domestic and entertainment domains which naturally involve social interactions, a requirement to gauge a system's abilities either individually or relatively has emerged. This paper discusses the implications that arise in aiming to assess the performance of robots engaged in a degree of social interaction with people.

Key-Words: - social robotics, performance evaluation, anthropomorphism, artificial intelligence, robot metrics

1 Introduction

> *"It is much easier to make measurements than to know exactly what you are measuring." [John William Navin Sullivan (1886 - 1937)]*

The trend in science has always been to try and quantify observations and results. In psychology, where complex factors abound, a form of quantifying in terms of probabilities and statistical significances has been towed in to justify conclusions and theories. Artificial intelligence and robotics, being even more computationally entrenched, do not even consider the possibility of non-quantitative analyses or benchmarks. Performance assessment has always been elusive in robotics as the problem of measuring intelligence has always surfaced other problems such as the definition of intelligence. As the area of social robotics advance [1][2][3], proper benchmarks to assess social robots become even more difficult to design. How do you measure how well a robot integrates into a social setting? Even if the classic Turing test were appropriate, it is still based on qualitative observations and subjective interpretations, which are considered inappropriate evaluation methods in the circles of science and recalls the age-old qualitative-quantitative debate in the social sciences. Besides, does a social robot have to pass a Turing test in order to perform successfully? For example, cartoon characters would never pass a Turing test for being real entities, but they can still be perceived as integrated entities as humans in social settings. Reminders of the Turing test also raise another issue in performance assessment: does the successful social robot have to be socially intelligent, or only appear to be passed off as socially capable?

This paper argues against a basis for the dichotomies of qualitative-quantitative or behavioural-internal methods of evaluation, in particular for social robotics. These methods complement each other and are interrelated. A competent social robot obviously would require a rudimentary set of social knowledge. However, sophistication does not necessarily require comparable complexities in design because roboticists can take advantage of people's tendency to anthropomorphise in their interpretations (see section 4). The question then turns to the validity of this reliance on subjective and qualitative interpretation, though, after all, social function is about behaviour. Our expectations are reasonable only when we expect to create robots that *know* (not robots *who understand*) social values and morals. Analogously, there is a difference between a robot that is programmed to prioritise recharging its batteries when energy levels are low so that it neglects all other programmed goals, and the human who feels the instinct of survival. Yet, are we satisfied enough with a robot that can *behave* with these survival instincts?

2 Robot Experimentation

In order to assess the functionality and success of a control paradigm for real world robotics, one normally seeks to collect a set of empirical data, analyse it, and hopefully arrive at the conclusion that the appropriate final "value" indicates a level of accomplishment. Realistically, this is not possible in robotics. Experimentation is an inherently subjective process, particularly in the field of autonomous real-world mobile robotics where the control paradigms generally apply little long-term learning and are programmed based on a human's perspective of the problem domain. Existing qualitative and quantitative research methodologies in real world robotics should prompt many of us to rethink some basic assumptions about such notions as: the nature and role of robot research, and the theory behind existing approaches to the design and development of control paradigms for robots functioning in real world environments.

The problem has always been the choice of appropriate references for assessment. It is because of this basic issue that there are no general benchmark experimental procedures for autonomous mobile robotics. This issue of assessment is replete in computational and engineering backgrounds. For example, in computational linguistics, the debate remains as to how to assess the performance of natural language processor that performs tasks on which even human experts disagree.

As the anthropomorphic metaphor in socially interacting robot becomes more dominant (i.e. social robots or humanoid robots [1][2][3][4], the field of robotics also inherit the methodology problems in sociology (e.g. whether quantitative methods capture the complexities in the social domain), without having lost those from robotics. The question is whether the measure of success of a robot is in the accomplishment of a task or whether it replicates how a group of humans would behave in completing the task (i.e. emulation of human behaviour). Is assessing by goals enough or is that too limited?

In seeking to assess multi-agent systems techniques, Jennings *et al.* [5] propose the following qualitative criteria where agents have a sufficient degree of social functionality:

- How do socially responsible agents fare in terms of their individual benefit?
- How does the overall collective fare in terms of the global benefit?
- How do socially responsible agents compare on an individual and overall system perspective to other types of multi-agent system?
- How do responsible communities behave when there is an imbalance in terms of the value of individual and/or social actions?

While these represent a more conceptual analysis of a system of socially capable agents, a stronger set of criteria is required which directly embraces both the physical and social embodiment of a collective of robots required to undertake complex social tasks.

The performance assessment of the social robot consequently becomes severely complex when they not only interact with other robots but also interact socially with people.

The following section briefly reviews existing benchmarking techniques and discusses the issues regarding the development of more suitable approaches for social robot experimentation.

2.1.1 Benchmarking in Robotics

Traditionally, once experiments are designed and undertaken to demonstrate the performance of a system, a set of replicable metrics are required to assess and evaluate the experimental results. However, very little work on the qualitative and quantitative assessment of robot control architectures and its principles has been undertaken in autonomous mobile research. Existing literature on benchmark testing to date includes [6] and the work on the Mars Microrover Navigation [7]. No universal systematic benchmark procedures have been developed to assess differing robot control methodologies. A primary factor in the lack of standardised formal testing is the diversity of robot platforms and their software control algorithms. It is only recently that off-the-shelf robots can be easily purchased, i.e. the Nomad Technologies Inc. or the Real World Interface Inc. family of robots. The majority of research to date has been using unique robot prototypes and highly tailored platforms for experimentation. The primary approaches for performance measurement in autonomous mobile robotics can generally be categorised as:

- Theoretical predictions of performance
- Computer simulations
- Real-world experiments

While complexity problems occur with the generation of an artificial representation of the robot-environment interaction in the first two approaches, the use of real world experimentation leads to problems stemming from the collection of useful data and their interpretation and analysis [8].

In general, internal representation-based approaches to modelling real world scenarios (as found in classical AI) pose basic problems in maintaining real-world validity and system robustness [13]. On the other hand external observer-dependent approaches as found in emergent systems are difficult to both replicate and quantify [14][15]. In reviewing the validity of simulations, Gat [9] noted that two broad categories exist in mobile robot research:

- Theoretical work with little experimental verification

- Experimental work with little theoretical verification

Theoretical and simulation perspectives on real world robot performance cannot sufficiently model or approximate such factors as sensor noise, occlusions and resolution limits, friction, surface deformation, to list but a few. While often used as a criticism of robot research, the strong linking of theoretical and experimental approaches in physically embodied systems will continuously plague researchers in both camps until someone conveniently develops a working realistic model of the world (!). For now, where researchers have traditionally resorted to theory and simulation to justify techniques, the real-world puts a spanner in the works.

While observer dependent assessment provide a degree of abstraction from the actual mechanisms employed to generate system behaviour, analysing these mechanisms allows another perspective of the robot. Smithers points out [10] that behaviour-based approaches [11][12] in robotics result in the robot behaviours being an emergent property of agent-environment interactions. This contradicts, in a sense, the principle of many behaviour-based robot control strategies, which aim to develop a robust, coherent and *specific* set of behaviours like `follow_wall`. Assessment of such explicit behaviours is straightforward, as they must be formally specified before implementation. Comparing the implemented behaviour against the definitions show that specifications *should* encompass such notions as behaviour thresholds, tolerances whereby the observable execution of the behaviour can be assessed based on these criteria. To design an explicit behaviour without such dynamic attributes is a gross oversimplification of the complexity of the real world and contradicts the basic issues of embodiment.

Another dimension arises when one seeks to develop more complex behaviours, and the corresponding definition of reference parameters and tolerances for their assessment becomes severely complex. Similarly, emergent behaviours, like those found in [16], do not lend themselves to such analysis. Consequently, it has generally been argued that it is important to augment the “it works” form of evaluation with more quantitative measures of performance given that experimental replication with such emergent behaviours is difficult.

The VR-Workbench [17] seeks to address the synthesis between observational and internal perspectives with a view towards reinforcing the notion of physical embodiment of inherently symbol manipulating computer-based systems through the fusion of both real-world experimentation and a virtual reality manifestation. Through the process of recalibration of a robot’s position in the VR space based on the synchronisation of the robots distance from location cues in both the VR and real spaces, the robot’s localisation error map can be generated and fed back into the control system to develop a more robust robot localisation strategy. Unfortunately such a system is still environmentally constrained with limited transference to alternate physical spaces due primarily to sensor limitations such as the reflectivity issues of the robots sonar sensors and calibration issues of its colour vision system.

Perhaps one of the more successful attempts to tackle the issue of standardisation for multi-robot performance is the RoboCup initiative (www.robocup.org). It provides standard test environments, robot functionality and construction constraints in order to assess alternate control paradigms. To date, the only empirical reference derived from RoboCup games is the scoring and the consequent championship winning team. The audience passively observes particularly obvious traits exhibited by any one system over another, i.e. speed, reliability, adaptability and coordination. This highlights a key point in future benchmarking of embodied systems. In order to assess the performance of a robot, such apparently abstract yet concrete criteria as the final score of a RoboCup competition *does* provide a quantitative assessment of one system over another without distilling behaviours into local disjoint measurements.

3 Social Robotics

Social robotics adds an even more opaque set of complexities to benchmarking. How do you measure successful social interaction? What are the goals in designing a social robot? As robotics stride the social domain hence psychology, and with the difficulties in designing quantitative metrics, understanding performance qualitatively seems to be the most logical first step. This naturally invokes the qualitative-quantitative debate. Moreover, this in turns leads to the “problem” of people’s tendency to project anthropomorphic qualities on other entities and the stigma of anthropomorphism in science.

3.1 Goals of the social robot

First, the goals of a social robot must be delineated in order to know what performance to assess. A social robot must be able:

- *To interact in a given social scenario* (exhibit social initiative); by virtue of the definition of "social", this implies being able to interact. Being able to interact includes interpretation of others' expressions (whether verbal or gestural) as well as portraying expressions, which are interpretable and faithful to robot's representation of what it has expressed.
- *To adapt*. Social intelligence is about adapting and learning behaviours of other participants in the social circle (whether human or machine) and situations

In assessing these goals, is goal or task completion enough? Perhaps if the task is well defined enough. But how can subjectivity be separated from qualitative observation? Is subjectivity all that bad (onwards to the qualitative-quantitative debate)?

With the amazing complexities of sociality, observer-dependency appears to an inherent feature in evaluating the social robot. In this case can we even dissociate qualitative observation from performance assessment?

3.2 Qualitative vs. Quantitative

In 1904, Edward Thorndike stated, "*anything that exists, exists in a certain quantity and can be measured*". In 1927, the motto: "*When you cannot measure, your knowledge is meagre and unsatisfactory*" was prominently and permanently carved onto the face of the University of Chicago's social science research building. Distinguished faculty of the day were alleged to have muttered "*and if you cannot measure, measure it anyhow*" [20]. The qualitative-quantitative dichotomy in fact dates back as early as the 17th century where "quantitativists" were characterized by some as "*vulgar statisticians*" [21].

Quantification is a human and political process of discarding information from the incredibly rich and complex fabric of real world environments. Information is lost and discarded when situations and scenarios are reduced to numbers. In theoretical systems, such control is possible and such quantification appropriate but, as the sciences of chaos and complexity teach us, real world systems are neither simple nor predictable when so oversimplified, especially after bifurcations of key parameters, which drive the behaviour of any otherwise simple system into even greater uncertainty. Emergent intelligence demonstrates this [14][15].

In essence, researchers of those sciences pertaining to the complexity of the real world should be careful not to succumb to a false certainty of quantification, where there exists the false hope of over controlling reality. The inherent flaw lies in the empirical study of natural systems and the consequent abstraction from the fundamental embodiment issues found in autopoietic entities [22][23]. Subjecting an animal to mechanistic qualitative and quantitative analysis forces a digital perspective of an analogue world. It leads to some assumptions, though needed for digitisation, may be unknowingly fundamentally flawed. This *should* provide researchers with an insight into why robot systems consistently fail. When a system becomes embodied in a physical world, the world and its uncompromising complexities are inherent attributes of the overall system to be "measured". Real world robotics is one such domain. This reiterates the fundamental embodiment issues raised in [22][23].

Moreover, so called quantitative methods and data often have qualitative bases. A survey that collects people's opinions of da Vinci's Mono Lisa painting (love, like, dislike, hate) is qualitative. However, when the survey requests the opinions to be on a scale from 0 to 10 is considered quantitative – is this quantised data better or more valid? Or is the Turing test invalid because it is more qualitative in nature than quantitative?

Social robotics is robotics impinging on social science, and thus inherits some of the problems found in both areas. Cognition and human behaviour are more complex – how do you isolate factors for things as complex as social interactions?

In recent times, however, the competing paradigms of quantitative and qualitative research have become almost working partners in educational research.

One of the pitfalls often associated with qualitative methods and evaluation have been anthropomorphising and overly subjective interpretations. They also make comparisons of different research difficult. Qualitative methods do tend to allow for more unreliable conclusions and wider disagreements (hence reduced ability to predict outcomes for same scenarios but different observers) than purely quantitative approaches. However, much of the roots of science have stemmed from qualitative observations, which bootstrap for subsequent mathematical formulation.

3.3 Assessment in Social Robotics

With the debate over qualitative-quantitative approaches, and the unsatisfactory and working definitions of "intelligence" (much more "social intelligence"), a standardised performance metric remains work in progress.

The panel of NIST's 2000 Performance Metrics Workshop [18], with the focus of measuring machine intelligence, have suggested that task completion be the best evaluation scenario rather than direct evaluation (thus analogous to the RoboCup testing environment). One suggestion from the conference discussions proposed systematic observations of a robot designed to be a rich social participant in interaction with humans (say in a natural history museum). Evaluation would include systematic qualitative observations/ evaluations combined with quantitative metrics such as the number of people who followed the robot to look at specific exhibits, length of time people spent with the robot, and how well they performed on a quiz compared with those who did not interact with the robot. They admit, however, to the difficulty of a task to be objectively and quantitatively measured. Indeed, many factors such as the robot appearance and mode of interaction (e.g., voice quality) would need to be controlled in order to eliminate any confounding factors when comparing across different architectures.

This section has attempted to lay out the goals of a social robot and the concerns which arise in trying to assess the success with which it reaches these goals. It notes the inevitable (at least for now) partial on qualitative observations. As with all more subjective observations, this leads to a problem of possible inconsistencies among observers. This section also pointed out the importance of anthropomorphism in social robot assessment, and any further developments in performance metrics should require a better understanding of our tendency to project anthropomorphic qualities.

Despite the increased difficulty in comparing across social robot performances, which is further confounded by anthropomorphism, robot design has become relatively easier. That is, in order to design a sophisticated robot, the roboticist does not need to have a design that is equally as complex. The panel of NIST's 2000 Performance Metrics Workshop also agreed that "social behavio[u]r is fundamental: it compensates for the lack of perfection of the individual intelligent system" [19]. This probably is precisely because if robots can behave above a certain threshold of social intelligence and approachability, it is easier for the interacting human to attribute more anthropomorphic features (such as intelligence) to the robot than otherwise, thus compensating for the actual limitations of the robot. If played right, the social capabilities and sophistication of the social robot can almost be achieved through an illusion.

Our tendency to anthropomorphise may simplify design tasks; however, it also highlights an additional dilemma: the success of a social robot in a social scenario seems to be based strongly on subjective responses. For instance, one possibility is to measure performance by how much people can project anthropomorphic qualities given a particular robot (so the more a robot can elicit anthropomorphic projections, the more successful the robot). With our lack of understanding of anthropomorphism, this method is not the most reliable unless we can normalise against the possibility that some people naturally project anthropomorphic qualities much more readily than others.

There is then the problem of a robot's appearance, which is part of the strength of a social robot (social approachability) and is valid in the assessment process in deciding its essential foci. One should be concerned about the appearance insomuch that it facilitates people's tendencies to anthropomorphise; however, do we need to worry about making a social robot "prettier" in order to attract interactors (dealing with the dynamics influencing human social interaction)? Is there a concern of obscuring some of the issues and goals in designing a social robot, and of course, how do we quantitatively control for aesthetics?

Both quantitative and qualitative approaches are important and complement each other. Until we understand the mechanisms underlying anthropomorphism and other social interactions, perhaps performance evaluation remains qualitative. In seeking to address this issue, more research must be undertaken to explain how people project anthropomorphic qualities while interpreting a robot's actions.

Thus, a quantitative (and qualitative) metric remains work in progress, but we may advance with more insights with further understanding of the mechanisms underlying anthropomorphism. In order to begin the unexplored task of researching these mechanisms, we describe a set of experiments we are currently undertaking.

4 Exploring Anthropomorphism: Emotion Robots

Experiments on projective anthropomorphism using a collective of Khepera robots demonstrate the power of employing human-like behaviours to bootstrap the social interaction scenario between people and robots. Preliminary experiments were run where one or many robots were performing wandering, wall-following, or "dancing" behaviours and subjects were required to relate their observations. An interesting result occurred with a robot which was performing a Braitenburg wandering behaviour but which was also sensitive to lighting conditions. Shining a desk lamp on it resulted in a jittery behaviour that many observers attributed to the robot as being "afraid of the light". These informal preliminary experiments led to our exploration of how simple motions could be interpreted as much more complex intentions and lay the basis for the beginning of our research on anthropomorphism and its role in social interaction and social robots.

4.1 The Seven Dwarves

Seven Khepera robots were programmed to perform motion behaviours intended to illustrate the characters of the seven dwarves in the famous Disney version of the fairy tale Snow White: Happy, Dopey, Bashful, Doc, Grumpy, Sneezy, Sleepy.
As the Khepera robots are limited in their actuator capabilities (two wheels with an optional gripper) the expressive capabilities of each robot was highly constrained.

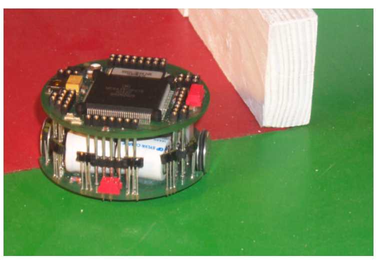

Fig 1. Khepera robot in experimental environment

A web site with video clips of the robots performing is published online with a questionnaire for participants to describe what they perceived the robots to be doing [24]. The online experiment was designed to enable a bigger pool of subjects from more diverse backgrounds to participate in the experiment. The objective was not for people to recognise the motivation for generating such motion behaviours in the robots but rather to express what they believed was the scenario for such motion behaviours.
Preliminary results have already demonstrated people's inclination (whether consciously or not) to overlay a degree of anthropomorphic behaviour in their interpretations of the robots' actions. While relatively few people deduced that the robots were representing the seven dwarves, some participants did define the robots motions in terms of emotional motivations. These experiments are also being extended further to explore other issues (such as the influence of more explicit social associations such as costumes or storytelling through the motions of interacting robots), and shall be all cross-compared with the current baseline experiments.

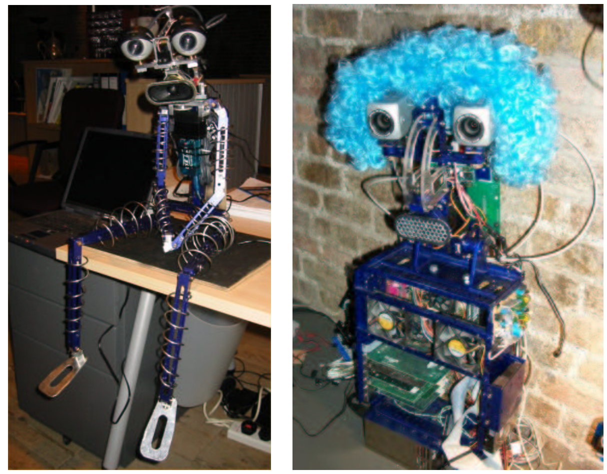

Fig 2. The robots (a) "Joe" and (b) "Anthropos"

Preliminary testing with the robot "Anthropos" and "Joe" have also already demonstrated the influence of anthropomorphic features in one's explanation of a robot's action. Work in progress is aiming to develop these experiments on human-robot social interaction further.

4.2 Discussion

These experiments raise the questions similarly posed for emergent intelligence experiments in robotics, that of repeatability and the ability to define and realise explicit complex behaviours. The development of the *Facial Action Coding System* (see [25]) would suggest that the development of a similar behavioural "alphabet" for motion behaviours for expression, particularly in social scenarios, is imminent. Work such as "Kismet" [3] has already looked to define a set of expression behaviours for robots engaged in social scenarios with people. Such a coding system could sufficiently mask the techniques lying behind the generation of particular behaviours thereby facilitating alternate approaches to behaviour generation. The actual programming strategies employed to realise particular behaviours become arbitrary.
This idea can be extended further to the degree of using "projective intelligence" as the measurement of intelligence in artificial systems rather than basing assessment on the traditionally view defined empirical data. The Turing Test inherently uses projective intelligence in its assessments of being

"intelligent" or not. The fight for "numeric" proof is difficult to validate.

5 Conclusion

If the social sciences reject idea that social sciences can be studied with the same methods as natural or physical sciences, do social robots pose a similar problem? If a social robot's behavior is always bound to the context in which it occurs, behavior must be studied holistically in context rather than being manipulated. If the method is therefore to base assessment on the observer's perspective, this makes qualitative research an intensely personal and subjective style of research for the social robot. But, is it possible to reduce social reality to variables in the same manner as physical reality? Research on embodied robotics and artificial intelligence would disagree [22][23][26].

Exploring projective anthropomorphism could provide key insights into the development of the "illusion of life" and the "illusion of intelligence" through *projective intelligence* in embodied robotic systems. If a robot "looks" intelligent, then the particular computational mechanisms lying behind the realisation of that behaviour are arbitrary. The real issue arises in how to *maintain* such an illusion over time. Does this inherently require that the robot be "intelligent" at a more fundamental level? Do we need to go beyond "pageantry" towards the "truth" if the result is the same? Can pageantry intelligence endure over time? Social robotics provides a powerful framework to finally build a system that *will* pass the Turing Test.

References:

[1] Hara, F. Kobayashi, H., (1995) Use of Face Robot for Human-Computer Communication, *Proceedings of International Conference on System, Man and Cybernetics*, p10.

[2] Duffy, B. The Social Robot. (2000) Ph.D. Thesis, Department of Computer Science, University College Dublin.

[3] Breazeal, C.L., *Designing Sociable Robots*, MIT Press, 2002

[4] Brooks, R., Breazeal, C., Marjanovic, M., Scassellati, B., Williamson, M., The Cog Project: Building a Humanoid Robot**,** *in C. Nehaniv, ed., Computation for Metaphors, Analogy and Agents, Vol. 1562 LNAI, Springer-Verlag, 1998*

[5] Jennings, N.R., Campos J.R., Towards a Social Level Characterisation of Socially Responsible Agents, *IEE Proceedings on Software Engineering, 144 (1), 1997, 11-25*

[6] Colombetti, M., Dorigo, M., Borghi, G. Behavior Analysis and Training: A Methodology for Behavior Engineering, *IEEE Trans. Systems, Man, Cybernetics-Part B, 26, 3, 365-380, 1996*

[7] Matthies, L., Gat, E., Harrison, R., Wilcox, B., Volpe, R., Litwin. T., Mars Microrover Navigation: Performance Evaluation and Enhancement. *IEEE/RSJ Int. Conf. on Robots and Systems (IROS), Pittsburgh PA, August 1995*

[8] Dragone, M., Performance Benchmarks for Robotic Behaviours, *IMPACT Technical Report, Department of Computer Science, University College Dublin, 2000*

[9] Gat, E., Towards principled experimental study of autonomous mobile robots, *Autonomous Robots, Vol.2, P179-189, 1995*

[10] Smithers, T., On quantitative performance measures of robot behaviour, *Journal of Robotics and Autonomous Systems (Special Issue: The Biology & Technology of Intelligent Autonomous Agents), Vol.15, P107-133, 1995*

[11] Brooks, R.A., A Robust Layered Control System for a Mobile Robot, *IEEE Jour. Rob. and Autom., 2(1) 1986*

[12] Steels, L., Building Agents with Autonomous Behaviour Systems, The "artificial life" route to "artificial intelligence". Building situated embodied agents. *Lawrence Erlbaum Associates, New Haven. 1994*

[13] Nilsson N.J., Shakey the robot, *SRI A.I. Center Technical Note 323, April 1984*

[14] Mitsumoto, N., Hattori, T., Idogaki, T., Fukuda, T., Arai, F., Self-organising micro robotic system, *6th International symposium on Micro Machine and Human Science, IEEE, 1995*

[15] Cai, A, Fukuda, T., *et al*. Hierarchical Control Architecture for Cellular Robotic System – Simulation and Experiments, *IEEE Int. Conf. On Robotics and Automation, 1995.*

[16] Lucarini, G., Varoli, M., Cerutti, R., Sandini, G. "Cellular Robotics: Simulation and HW Implementation", *Proc. IEEE Int. Conf. Robotics & Automation, Atlanta,1993, pIII-846-852*.

[17] Duffy, B.R., O'Hare, G. M. P., O'Donoghue, R.P.S., Rooney, C.F.B., Collier, R.W.**,** Reality and virtual reality in mobile robotics, *1st International Workshop on Managing Interactions in Smart Environments MANSE'99, Dublin December 1999*

[18] *Proceedings of the Performance Metrics for Intelligent Systems*, NIST, Gaithersburg, Marlyand, 2000.

[19] Balch, T. et al., Summary of plenary discussions, *Proceedings of the Performance*

Metrics for Intelligent Systems, NIST, Gaithersburg, Marlyand, 2000, Part III.
[20] Custer, R.L., Qualitative research methodologies, *Journal of Industrial Teacher Education, 34(2), 3-6, 1996*
[21] Lazarsfeld, P.F., Notes on the history of quantification in sociology: Trends, sources and problems, *in M G Kendall and R L Plackett (eds.), Studies in the History of Statistics and Probability, II (London, 1977), 213-270*
[22] Sharkey, N., Zeimke, T., Life, mind and robots: The ins and outs of embodied cognition, *Symbolic and Neural Net Hybrids, S. Wermter & R. Sun (eds), MIT Press, 2000*
[23] Duffy, B.R., Joue, G., Intelligent Robots: The Question of Embodiment, *BRAIN-MACHINE'2000, 20-22 December 2000, Ankara, Turkey*
[24] http://www.mle.ie/research/anthropos/questionnaire.html
[25] Ekman, P. and Friesen, W.V. (1977) *Manual for the Facial Action Coding System*, Palo Alto: Consulting Psychologists Press.
[26] Clark, A. *Being There: Putting Brain, Body, and World Together Again.* MIT Press, 1997