Social Empowerment Of Autonomous Mobile Robotics

Brian R. Duffy

PRISM Research Group, Department of Computer Science,
University College Dublin, Belfield, Dublin 4, Ireland

E-mail: Brian.Duffy@ucd.ie

Abstract: This work aims at demonstrating the necessity of embracing a strong notion of social embodiment in designing a real-world robot with coherent "intelligent" social behaviour. It develops the current conception of embodiment beyond mere physical aspects by demonstrating the inherent importance of social embodiment. It also develops a corresponding social framework to maximise the fundamental social attributes found when more than one robot co-inhabit a physical space.

1 Introduction

To date, a fundamental facet of embodiment has, on the whole, been neglected in autonomous mobile research and in artificial intelligence as a whole, that of *social embodiment* [1] [2]. Embodiment has been interpreted as being the physical existence of an entity, i.e. a robot, in a physical environment (by robot it is understood to represent a physical body with actuator and perceptor functionality). By virtue of it physical presence, whether static or dynamic, there is interaction between the robot and the environment. At a fundamental level, this can be the physical space occupied by the robot and extending to the robot's ability to move, change, and perceive the environment. When a second robot is added, this introduces a definite element of social interaction, even without any direct inter-robot communication. The perceptions of another robots motions, whether abstract notions of a moving obstacle or its clear distinction as another individual robot, influences the observing robots behaviour. The social implications of two robots coexisting in an environment add another dimension to the complexity of each robot's environment, which cannot be ignored.

This work seeks to develop *explicit* social functionality between a collective of autonomous mobile robots as distinct from work on "emergent intelligence" [3].

2 Social Power

The current state of the art in the realisation of high-level social behaviour has not extended far from a conceptual interpretation and understanding with Newell's *Social Level* [4], and Jennings *et al.*'s *Principle of Social Rationality* [5]. A wealth of anthropomorphic social analogies in the pursuit of the intelligent robot has therefore not been exploited.

In a paper entitled "Social Power: A point missed in Multi-Agent, DAI and HCI" [6] (DAI – Distributed Artificial Intelligence, HCI – Human Computer Interface), Castelfranchi suggests that there has been a "*serious lack of realism in Multi-Agent and interaction studies*" where "*sociality or the agents is merely postulated rather than explained in terms of their* *dependence*". In addressing this issue, Castelfranchi proposes a distinction between Distributed Artificial Intelligence and what he terms the "Social Simulation Approach", i.e. the difference between the research concerned with intelligence, problem solving and system architecture where society is "*used as a metaphor*", and, on the other hand, work that deals specifically with social interaction.

This paper proposes that in order to develop an artificially intelligent physically situated robotic entity, social embodiment is a necessary criterion (see also [1] [2]). A robot must have both the capabilities to be social in conjunction with its abilities to solve social problems. While this is not a new notion, it has not been developed in the context of robotics and artificial intelligence as an all-encompassing coherent approach. The next section therefore aims at taking intentional multi-robot systems a stage further by developing the "social robot".

3 The Social Robot

In order to realise what could be termed a "Social Robot" [1] [7], the following sections seek to realise the social attributes pertinent to the social interaction of a collective of real-world autonomous mobile robots.

Kinny *et al.* [8] propose the notion of the "internal" and "external" aspects of an agent where the *internal* features of the agent comprise its *beliefs*, *desires*, and *intentions*, while the *external* features relate to features of the social group, i.e. the roles and relationships within the system. In this work, the notion of *internal* and *external* discussed in the following sections relates to the attributes of a single embodied agent or robot analogous to Shoham's notion of *capabilities* in multi-agent systems [9]. The *internal attributes* of the robot are analogous to Kinny *et al.*'s [8] *internal* features. While Kinny *et al.*'s *external* aspect relates to the social interaction, i.e. the services an agent provides, its interactions, and the syntax and semantics for the communication between agents, this notion is here developed further with the addition of *stereotypes*, *roles* and *characters* as the *social* features.

In order to encompass the added complexity of dealing with *embodied* agents (both physically and socially), the original *internal* and *external* features of an agent proposed by Kinny *et al.* [8] are insufficient, as they do not address the added complexity of the physical environment. A discussion of strong physical and social embodiment can be found in [1] [2]. It follows that in order to address the issue of embodiment, there are two distinct robot attributes that are local and particular to each robot within the social system:

- *Internal Attributes:* Beliefs, desires, intentions, i.e. name, character, mental capabilities, the robot's knowledge of self, experiences, a priori and learned knowledge, processing capabilities (i.e. algorithms)
- *External Attributes:* the physical presence of the agent in an environment; its actuator and preceptor capabilities (i.e. a robot equipped with extra sensory equipment compared to another), the physical features of the robot, i.e. physical dimensions.

And one global system attribute which subsumes the social functionality of the collective of robots:

- *Social Attributes:* Identity, character, stereotype, roles.

The *Social Attributes* are more abstract social features that exist to facilitate the interaction between robots. While some pertain to the robot itself, they nevertheless constitute attributes existing in the social system that are necessary for the social functionality of the system. These attributes are developed in greater detail in the following sections.

3.1 Identity

When social interaction exists, each element of the social group must be able to be differentiated from others. The robots require a sense of themselves as distinct and autonomous individuals obliged to interact with others in a social environment, i.e. they require an *identity*. Identity refers to the property appropriate or peculiar to an individual, that quality which distinguishes one person or thing from another whether it refers to a collective set of attributes or a separate entity in itself. Identity and the "self" have emerged as a central focus of theory and research in many domains of social and behavioural science: "*The large volume of empirical research on the self has convinced most social and behavioural scientists that the self is real, and that no science of the human experience is complete without accounting for it*" [10].

Suppose robot R_i is a part of a social group S. A social group must involve more than one robot. That is,

$$\exists\, R_i , R_j \in S \mid R_i \neq R_j \wedge S \subseteq \boldsymbol{R}$$

where $\boldsymbol{R}$ is the set of all possible robots

Smithers [11] discusses *identity* from the perspective of an observer where an entity (robot) is "*required to have the means to introduce additional dynamical properties in the interactions with its environment*". While this perspective is fundamentally based on classical approaches to robotics where the controller is simply placed in its environment, the social implications of multiple robots force a stronger notion of identity. This stronger notion will therefore differentiate a particular robot from its environment and is primarily independent of any particular behaviour that it may or may not exhibit.

Following from this, a definition of *robot identity* can be proposed:

> ***Robot identity:*** *that which differentiates a system from any other and preserves its distinction from its environment regardless if it has any observable influence on or actuation with its environment*

The identity of a robot is made up of its internal and external attributes. Examples of *internal-attributes* include: language, processing capabilities (i.e. vision, filtering, or smoothing algorithms), social knowledge, physical environmental knowledge, etc. Examples of *external-attributes* include: camera, sonar, wheels and motors, bumper, i.e. the physical features of the robot. The *identity* of R_α or $I(R_\alpha)$ is therefore founded on the union of both its particular *external attributes* EA (i.e. sensors, actuators, dimensions) and *internal attributes* IA (i.e. knowledge of its social and physical environment, processing capabilities, plan library):

$$I(R_\alpha) = EA(R_\alpha) \cup IA(R_\alpha),$$

where

$$\exists\, ea_i , ea_j \in EA(R_\alpha) \mid ea_i \neq ea_j \text{ and}$$
$$\exists\, ia_i , ia_j \in IA(R_\alpha) \mid ia_i \neq ia_j$$

Thus

$$EA = \{ ea_1, ea_2, ea_3, \ldots ea_n \} \text{ and}$$
$$IA = \{ ia_1, ia_2, ia_3, \ldots ia_m \}$$

implies:

$$I(R_\alpha) = \{ ea_i, \ldots , ea_n , ia_1 , \ldots , ia_m \}$$

In [9], Shoham assumes that capabilities are fixed. Intuitively, the use of internal and external attributes dictate that some are static and others are dynamic. Here it is understood that the external features of the robot (i.e. sensors, actuators) are static, while the internal features are indicative of the knowledge, experience, and processing capabilities (i.e. smoothing or filtering algorithms) that a robot has either obtained over time or has been initially provided with, and are therefore dynamic. As the *identity* of each robot is governed by a combination of its internal and external attributes, any changes in its internal attributes will therefore have an effect on its identity. The *identity* of the robot can therefore develop over time.

3.2 Character

The use of the terms "identity" and "character" in mobile robotics, or more generally embedded systems, is so far limited to non-existent (see [1] for a review).

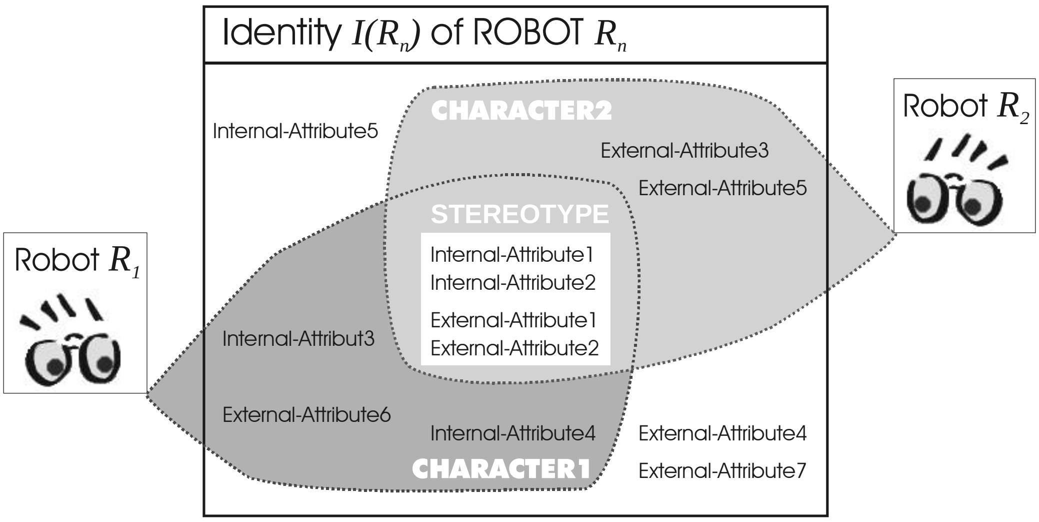

Figure 1: A defined subset of internal and external attributes constitutes the stereotype that robot R_n is associated with.

In this work, *character* is the *perceived* identity of one robot by another. Character is based on initial stereotype bootstrapping (see section 3.3), communicated knowledge, and/or experience in working with that robot over time. It corresponds to another robot's "interpretation" of one robot's identity based on the knowledge it has about that robot.

> ***Character:*** *The combination of perceived features or qualities that distinguishes one entity from another in that entity's social envelope.*

Character deals with the fundamental attributes an agent or robot is *perceived* by others in its social environment to have, its capabilities albeit physical or mental. *Character* represents a subset of the list of internal and external attributes of each robot. The character of robot R_2

has of the robot R_1 at any time t may or may not be the same as how robot R_3 has characterised robot R_1 (as shown in figure 1).

The character of robot R_j as perceived by robot R_i at time t_n is a subset of the total set of internal and external attributes that comprise identity:

$$C(R_j, R_i\ t_n) \subseteq I(R_j)$$

While a robot's *character* is transient over time, this *perceived* identity of another should be strongly founded on some fundamental set of internal and external attributes that describe that robot. This is achieved by the use of stereotypical representations, or *stereotypes*.

3.3 Stereotypical Representations

While character representations are initially independent of a global task or set of subtasks to be undertaken, the notion of stereotypical representations or *stereotypes* is proposed in order to bootstrap the initial stages of social interactions and reduce the complexity of maintaining such internal representations. This ensures that the core representations of another robot's identity remain valid for different robots. A fixed subset of internal and external attributes therefore comprises the *stereotype* with which each robot is associated. The perceived character of a robot is fundamentally based on a set of stereotypical representations available to each robot and developed through communication, collaboration and experience. At time t_0 when robot R_1 meets robot R_2, such stereotypical representations are used to facilitate the "introduction" procedures between the two robots, i.e. the initial communication to ascertain who and what the other robot is. Each robot is equipped with a finite set of stereotypical representations. These stereotypes "moderate" the character representations that one robot builds of another.

Each robot has one stereotype associated with it. The stereotype of the robot R_i is given by:

$$St(R_i)$$

There are a fixed number of stereotypes in the social environment:

$$Stereotype\ List = \{St_1, St_2, \ldots St_y\}$$

A stereotype comprises a defined finite list of *internal attributes* and *external attributes*:

$$St_x = \{ ea_1, \ldots , ea_\alpha , ia_1 , \ldots , ia_\beta \}$$

The stereotype of robot R_i is a subset of its *character* as perceived by robot R_j at time t, which in turn is a subset of its identity:

$$St(R_i) \subset C(R_i, R_j\ t) \subset I(R_i)$$

Each robot in the social group has knowledge of the possible stereotypes that may exist in its social environment and all details pertaining to each stereotype in that list (new stereotype information can be added to the system at any time). A robot therefore, knowing the stereotype associated to a particular robot it encounters, will also know all the internal and external attributes associated with the stereotype of that robot.

The use of stereotypes is indicative of exactly *how* another robot can help in completing a required global social task. This facilitates the breakdown of complex collaborative tasks into individual robot allocated subtasks by reinforcing the communication of elements of one robot's *identity* to another. The stereotype list of robots in the collective therefore provides a basis for subtask allocation when a global social task is undertaken by a robot collective and is discussed in the following section.

3.4 The Global Task

In order to have a social group of collaborating robots perform a global social task, the division of a social task into suitable subtasks to be performed by the co-operating collective of robots is required. These subtasks must be allocated relative to each robot's capabilities thereby dictating that all robots should be altruistic in nature. Commitments are then required by all robots to collectively work towards the global objective or goal.

When the global task has been decomposed into a set of subtasks, the issue is how to allocate these subtasks to appropriate robots within the social group. The notion of "role" is introduced to facilitate this subtask allocation.

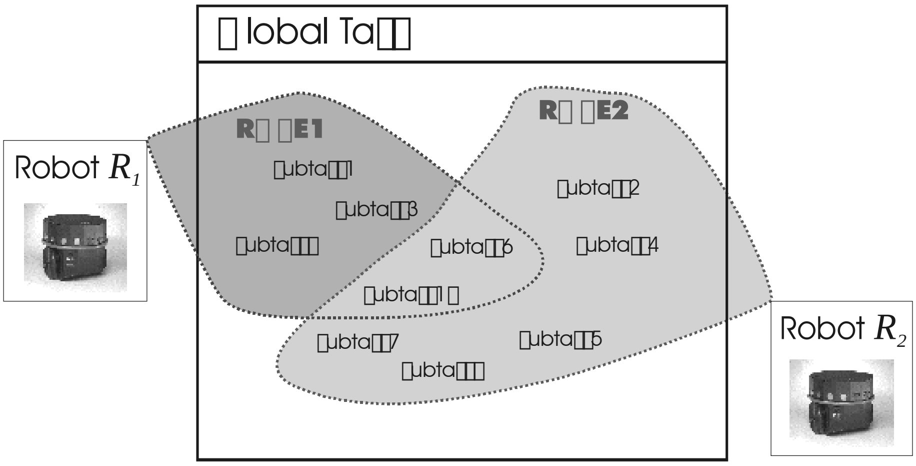

Figure 2: The roles that robot R_1 and robot R_2 are required to undertake in the realisation of the global task.

A *role* therefore corresponds to a subset of the social plan of tasks to be performed by an individual robot, and the corresponding behaviours to achieve these tasks. A task constitutes undertaking to perform a plan of behaviours. The use of stereotypes is based on Haddadi's [12] formal framework that develops on the BDI model of Rao and Georgeff [13]. The use of *roles* corresponds to sub-plans with the global task being the main plan. The formal specification of roles can be found in [1].

A robot may be assigned any number of *roles*. In such cases, there is temporal ordering of such *roles*. A *role* may have any number of subtasks, in any order, with possible repetitions of subtasks. If one attempts to assign a global task to a collective of robots that do not have the capabilities to perform what is required of them, this will become apparent via the stereotypes within the collective. The stereotypes will therefore constrain the incorrect allocation and subsequent robot functionality-based failure of the task based on their capabilities.

As physical attributes (external) of a robot are assumed static and absolute (i.e. the hardware configuration) and the internal attributes are more dynamic in nature (i.e. environmental knowledge), the *robot□ role* allocation is primarily based on a robot's physical attributes and its associated internal attributes (i.e. a camera and vision algorithms) with the dynamic internal aspects acting as more influential when indecision exists as to whether robot R_1 or robot R_2 can be assigned a role r_j.

A subtask constitutes a plan of elemental behaviours that can be executed by the appropriate robot. Each subtask may have any number of behaviours (i.e. `follow_wall`, `avoid_obstacle`, `take_snapshot`), in any order, with possible repetitions of behaviours (see [1]). These behaviours are dynamic in nature and may be over-ridden by reflex behaviours in emergency situations.

Each subtask t_i is comprised of a subset of all possible behaviours B:

$$t_i \subseteq B = \{ b_1 , b_2 ,\dots b_n \}$$

The behaviours are temporally ordered, ensuring that some behaviours are only initiated when appropriate others have been completed.

3.5 The Whole Picture

The concepts of *identity*, *character*, *stereotypes*, and *roles* proposed have been developed as a complete and integrated framework to facilitate the development of a social community of robots with suitable functionality to complete required social tasks.

In returning to the minimum set of concepts necessary for social behaviour and a society as defined by Jenning *et al.*'s [5] *Principle of Social Rationality*, the following objectives have been undertaken:

- Defining an independent primitive for each agent, i.e. its *identity*.
- Defining an independent primitive for one agent's model of another agent, i.e. its *character*.
- Description of task decomposition notions such as *roles* analogous to [14] [8].
- Assessing what "capabilities" exist in the co-operative, i.e. a global list of the attributes available to achieve the global task, i.e. *stereotype* listing.
- A mapping between the robots and the tasks to be undertaken (i.e. which robot undertakes which *role* initially based on its *stereotype* association).

The system objectives address the notions of *acquaintance*, *influence*, and *rights and duties* in a social collective of embodied robotic entities. The first two are new approaches to multi-robot control methodologies. This paper proposes that the definition and development of these primitives facilitates the development of a complex explicit social structure between a collective of socially capable robotic entities functioning in real-world environments. This not only addresses the issue of developing from the conceptual notion of the *Principle of Social Rationality*, but also seeks to apply this to the real world domain.

4 Conclusions

The previous sections have addressed the basic issues in embracing both physical and social embodiment within a collective of autonomous mobile robots undertaking explicit social tasks. The conceptual notions discussed in section 3 in the form of a socially empowered autonomous mobile robot, the *Social Robot*, have been realised [1] [7].

Experimentation has demonstrated that the concept of a Social Robot is possible through the grounding of the conceptual notions of beliefs, desires and intentions within a physically and socially embodied autonomous mobile robot [1] [15]. This supports work on realising stronger anthropomorphic mobile robot control paradigms at an architectural level. By embracing a stronger notion of embodiment (physical and social), such paradigms become less conceptual, more practicable and reproducible. The social framework maximises the fundamental social attributes found when more than one robot co-inhabits a physical space. The Social Robot therefore constitutes a next stage in the evolution of explicit autonomous mobile robot control by extending research in the multi-robot task domain. The Social Robot framework is not a localised solution to a specific problem but rather a generic approach for physically and socially embodied robots.

References

[1] **B.R. Duffy,** *The Social Robot*, PhD Thesis (submitted), Department of Computer Science, University College Dublin, Ireland, November 2000.

[2] **B.R. Duffy, G. Joue,** "Intelligent Robots: The Question of Embodiment", *BRAIN-MACHINE'2000*, Ankara, Turkey*,* 20-22 December 2000.

[3] **Y.U. Cao, A.S. Fukunaga, A.B. Kahng,** "Cooperative Mobile Robotics: Antecedents and Directions", *Autonomous Robots **4**,* pp.1-23, 1997.

[4] **Newell,** "The Knowledge Level", *Artificial Intelligence* **18**, pp.87-127, 1982*.*

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

[6] **C. Castelfranchi,** "Social power. A point missed in multi agent, DAI and HCI". *In Y. Demazeau & J.P. Mueller (Eds.) Decentralized AI*, Amsterdam, Elsevier, 1990*.*

[7] **B.R. Duffy, C.F.B. Rooney, G.M.P. O'Hare, R.P.S. O'Donoghue,** "What is a Social Robot?" *10th Irish Conference on Artificial Intelligence & Cognitive Science,* University College Cork, Ireland 1-3 Sept., 1999.

[8] **D. Kinny, M. Georgeff, A. Rao,** "A methodology and modelling technique for systems of BDI agents", *Proc. Of 7th European Workshop on Modelling Autonomous Agents in Multi-Agent Worlds (MAAMAW'96),* LNAI 1038, Springer-Verlag, pp.56-71, 1996.

[9] **Y. Shoham,** "Agent-oriented programming", *Artificial Intelligence* **60**(1), pp.51-92, 1993.

[10] **V.G. Hardcastle** "A Self Divided A Review of Self and Consciousness: Multiple Perspectives" Frank S. Kessel, Pamela M. Cole, and Dale L. Johnson (Eds.), *Psyche* **2**(1), April 1995.

[11] **T. Smithers,** "On quantitative performance measures of robot behaviour*", Journal of Robotics and Autonomous Systems (Special Issue: The Biology & Technology of Intelligent Autonomous Agents)* **15**, pp.107-133, 1995.

[12] **Haddadi,** "Reasoning about cooperation in agent systems: A pragmatic Theory", *Ph.D. Thesis,* University of Manchester Institute of Science and Technology (UMIST), March 1995.

[13] **A.S. Rao, M.P. Georgeff,** "Modelling Rational Agents within a BDI Architecture", *Principles Of Knowledge Representation & Reasoning,* San Mateo, CA, 1991.

[14] **P. Panzarasa, T.J. Norman, N.R. Jennings,** "Modeling sociality in the BDI framework", *submitted to World Scientific,* 1999.

[15] **B.R. Duffy, R.W. Collier, G.M.P. O'Hare, C.F.B. Rooney, R.P.S. O'Donoghue,** SOCIAL ROBOTICS: Reality and Virtuality in Agent-Based Robotics", *in Bar-Ilan Symposium on the Foundations of Artificial Intelligence: Bridging Theory and Practice (BISFAI),* Ramat Gan, Israel, 1999.