Description

[0001] The present invention relates to an architecture, embodied either in hardware or a software agent or in a combination of hardware and software, adapted to implement self-biased conditioning, that is, an architecture adapted to modify or specialise its responses to particular inputs, dependent upon the circumstances, in the case where the architecture determines that the existing response is inappropriate.

[0002] Such an architecture is of interest for a number of different reasons. Firstly, by being able to modify existing responses to particular inputs, the architecture can adapt itself to changing circumstances and, thus, render its functions better adapted to the achievement of an operational goal. In other words, the architecture can form the basis of an autonomous agent having the ability to perform unsupervised learning. Secondly, an architecture implementing self-biased conditioning appears to a human user to be exhibiting "social intelligence". (In this document, "social intelligence" means the set of skills or responses which enable an agent to interact appropriately with other agents for purposes including but not limited to the achievement of an operational goal.) Thus, the user will find that he can interact in a more intuitive fashion with the machine or programme embodying the architecture and this interaction will be experienced as more congenial.

[0003] Moreover, an architecture implementing self-biased conditioning is particularly well adapted to serve as the basis for new products such as "intelligent interfaces" and devices providing "situated personal assistance" (the personal assistance being "situated" in the sense that this assistance is provided in an appropriate context).

[0004] SPA (situated personal assistance) devices can take a wide variety of forms. Consider, for example, the needs of a tourist visiting a museum such as the Louvre and having only a small amount of time. It would be useful for such a tourist to have at his disposal a device, personalised to know his tastes, capable of indicating items in the museum which are worth his attention. In theory such a device could be created by providing a computer device having access to a database storing details of the floor plan of the Louvre and its exhibits (as well as other museums and the like which the tourist may visit) and pre-programming the computer with rules indicating the user's interests. However, such an approach requires a considerable amount of pre-programming and information gathering, both by the device's designer and by the user. An SPA device having an initial criterion for identifying objects of potential interest to the user, and capable of interacting with the user to learn the user's preferences, is much more interesting. The present invention enables such devices to be made.

[0005] The present invention is based on an analysis of social intelligence in animals and humans which leads to the identification of a set of rules or requirements which should be met by an architecture which seeks to give the appearance of social intelligence. However, it is not suggested that these rules or the architecture proposed according to the invention directly correspond to any particular structure or function of the human or animal brain.

[0006] Traditionally in the field of artificial intelligence (AI), the attempts that have been made to create structures emulating human cognitive processes have ignored the "social" aspect of much of human behaviour. The typical attitude is reflected by the statements *"systems are social if their learning process is connected with its social surrounding"* and *"for many purposes of cognitive simulation, it is of no special significance that thought is social"* made in the article "Situated Action: A Symbolic Interpretation" by A.H. Vera and H.A. Simon, in the journal Cognitive Science, 17, (1993), pages 7-48. However, as Vera and Simon also recognise *"All human behaviour is social. First and foremost, it is social because almost all the contents of memory, which provide half of the context of behaviour, are acquired through social processes - processes of learning through instruction and social interaction."* Moreover, the importance of the ability to interact appropriately with external agents is seen time and time again in the human and animal kingdom in such activities as foraging, mating, imitation, and expressions and experiences of emotion and sympathy.

[0007] Recently, some work in this sphere has been attempted, principally with regard to organisational strategy ("Representing and Using Organizational Knowledge in DAI Systems" by L. Gasser, in "Distributed Artificial Intelligence II" pages 55-78, ed. L.Gasser and M.N. Huhns, 1989, pub. Morgan Kaufmann) or study of group behaviour in the field of behaviour-based AI ("Behaviour-based Artificial Intelligence" by P. Maes, in "From animals to animats" 2, pages 2-10, ed. J-A Meyer, H.L. Roitblat and S.W. Wilson, 1993, the MIT Press).

[0008] Considering collective or "social" behaviour in the human and animal kingdom it can be postulated that such behaviour is based, at least in part, on a set of innate responses or behaviour patterns which are present from birth. For example, the mating behaviour of the fly Drosophila melanogaster appears to be genetically determined: the mutation of a gene controlling mating behaviour can cause reproductive isolation ("Isolation of mating behaviour mutations in Drosophila melanogaster" by E. Nitasaka in "Proceedings of the Third International Meeting of the Society for Molecular Biology and Evolution", 1995, pages 63-64). Moreover, in humans it has been found that damage to a certain portion of the brain leads to defective social behaviour ("Descartes Error" by A.R. Damasio, 1994, Avon Books).

[0009] It can be hypothesised on this basis that social behaviour has two innate requirements:

(1) the ability for agents to observe their own actions and those of other agents, and
(2) the possession of a set of primitive responsive behaviour routines.