Encyclopedia of Cognitive Science, Nature Publishing Group, Macmillan Reference Limited, Nov 2002

Level 2

Cognitive Processing through the Interaction of Many Agents

Chris Jones,
Maja J Matarić,
Barry Werger

CONTENTS

A collection of interacting autonomous entities, called 'agents', may be capable of creating complex cognitive processes and behaviors, which could not be achieved by a single agent, without the need for outside centralized coordination or control.

INTRODUCTION

Several theories of cognition, most notably Minskys's Society of Mind, posit that intelligent behavior can be seen as the result of interaction of simple processes. Minsky states: 'Very few of our actions and decision come to depend on any single mechanism. Instead, they emerge from conflicts and negotiations among societies of processes that constantly challege one another.' (Minsky, 1986). The central tenet of such theories of cognition and behaviour is that complex system-level behavior can emerger from the interaction of multiple, possibly numerous, components.

A canonical example of such emergence is the function of he human brain. The brain itself is made up of billions of simple neurons organized into a massively connected network (Nicholls *et al.,* 2001). In general, an individual neuron acts as a comparatively simple processing unit that receives signals from a set of neighboring input neurons, and under appropriate conditions transmits signals to a set of its neighboring output neurons. From this network of interacting neurons emerges the complexity of human cognition and behavior. No single neuron or subset of neurons is responsible for this complexity; rather it is the result of their interactions.

Several disparate research fields are actively involved in investigating the principles of interaction among a collection of components. These include cognitive science, computer networks, distributed systems, artificial life, collective robotics, multi-agent systems, as well as others. The rest of this article aims to explain conceptually, and show through examples, how complex cognition and behavior can emerge from the interaction of individual components and how those emergent behaviors can be used in a variety of ways.

AGENTS

The term 'agent' has become a popular choice for a nontrivial component of a system with many inter-acting component that result in emergent behavior. Precisely defining an agent remains difficult, as agents come in many guises. An agent could be a piece of software, an specific computer on the Internet, a mobile robot, or even a person. In general, an agent is an autonomous entity, situated in an environment, and equipped with some degree of intelligence. Being 'situated' places a number of constraints on how the agent can operate (Brooks, 1991; Maes, 1990). It implies that the agent has some means of sensing its environment, but the sensing may be limited and inaccurate. For example, a mobile robot may be situated in an office environment and have a means of sensing the distance to nearby objects; thus its sensing capabilities are both limited and prone to error and noise. Situatedness also implies an interaction with the environment. Thus, the same robot may be able to drive around, thus affecting the environment with its placement, and perhaps move objects, intentionally or otherwise. Conversely, the actions of a situated agent are influenced by the environment. The objects the robot encounters affect what it senses and how it behaves. Another implication of being situated is that an agent is constrained by environmental characteristics. For example, a mobile robot cannot drive through walls nor can it avoid falling if it drives down a set of stairs. Finally, the agent's characteristics, i.e., its computational, sensory, and actuation capabilities, influence how it interacts with its environment, which includes other agents.

SOCIETIES OF AGENTS

A collection of interacting agents is referred to as a 'society of agents'. Using this metaphor, a brain is a society of agents, as is a team of mobile robots cooperating on a task. Such agent societies are interesting for a number of reasons. First, for certain tasks and/or environments, a society of agents is the only viable or efficient solution. Second, even for tasks that can be handled by an individual agent, there may be more efficient, adaptive, and robust solutions performed by a society of agents.

A society of agents may consist of a homogeneous or heterogeneous collection of agents. In a homogeneous society, all agents are identical, while in a heterogeneous society, agents may have different characteristics. The variations in capability may result in hierarchies, specializations, or various other forms of social organization. Consequently, heterogeneous societies are generally more complex to control but are typically capable of larger set of tasks.

The humane immune system (Segel and Cohen, 2001) is an excellent example of society of heterogeneous yet simple agents. It is capable of protecting the body against infection and invarsion by foreign substances of all kinds, whether bacteria, virus, parasite, etc., which can be viewed as a very large set of different 'tasks' to be accomplished. Each agent is this society is very specific, but the large number of agents of each kind and in total, combined with the ability to generate additional agents when needed, produces an unprecedented defensive functionality.

AGENT INTERACTION

Agents situated in a shared environment have ample opportunity to interact, by directly sensing each other, communicating, coordinating actions, and even competing. The shared environment in which the interactions take place may be an abstract data space, the physical world, or anything in between. A society of software agents may interact through a personal computer or even the entire Internet. Likewise, a society of mobile robots may interact in an office environment or on the surface of Mars. Furthermore, multiple environment types maybe be spanned in a single multi-agent system. For example, in multi-robot systems using behavior-based control (Matarić, 1994) individual robots are controlled by a collection of internal agents that interact through a computational environment, while the society of physical robots that contains them interacts interact in and through the physical world. The nature of their interactions in a society of agents depends upon such factors as agent capabilities, environmental constraints, and desired local and global behavior.

The spectrum of agent interaction is broad. At one end are methods that employ larger number of simple, identical agents, connected together in patterns which lead to useful computation as a result of data flow through the system. At the other end are systems of complex, specialized agents which explicitly negotiate for task assignments and resources. Mechanisms for agent interaction can be broadly classified as fitting into the following, often overlapping categories: interaction through the environment, interaction through sensing, and interaction through communication. Each is described in turn.

Interaction Through the Environment

The first mechanism for interaction among agents is through their shared environment. This form of interaction is indirect in that it consists of no explicit communication or physical interaction between agents. Instead, the environment itself is used as a medium of indirect communication. This form of interaction is indirect in that it consists of no explicit communication or physical interaction between agents. Instead, the environment itself is used as a medium of indirect communication. This is a powerful method of interaction that can be utilized by very simple agents with no capability for complex reasoning or for direct communication.

Stigmergy is an example of interaction through the environment employed in a variety of natural insect societies. Originally introduced in the biological sciences to explain some aspects of social insect nest-building behavior, Stigmergy is defined as 'the process by which the coordination of tasks and the regulation of construction does not depend directly on the workers, but on the constructions themselves' (McFarland, 1985; Holland and Melhuish, 1999). The notion was originally used to describe the nest-building behavior of termites and ants (Franks and Deneubourg, 1997). It was shown that coordination of building activity in a termite colony was not inherent in the termites themselvers. Instead, the coordination mechanisms were found to be regulated by the task environment, in this case the growing nest structure. A location on the growing nest stimulates a termite's building behavior, thereby transforming the local nest structure, which in turn stimulates additional building behavior of the same or another termite.

Examples of artificial systems in which agents interact through the environment include distributed construction (Bonabeau *et al.,* 1994), sorting (Deneubourg *et al.,* 1990), clustering (Beckers *et al.,* 1999), object manipulation (Donald *et al.,* 1993) analysis of network congestion (Huberman and Lukose, 1997) and phenomena such as the spread of computer viruses (Minar *et al,* 1998).

Interaction Through Sensing

The second mechanism for interaction among agents is through sensing. As described by Cao *et al.* (1997), interaction through sensing 'refers to local interaction that occur between agents as a result of sensing one another, but without explicit communication.' This form of interaction is also indirect, as there is no explicit communication between agents; however, it requires each agent to be able to distinguish other agents from miscellaneous objects in the environment. Interaction through sensing can be used by an agent to model the behavior of another agent or to determine what another agent is doing in order to make decisions and respond appropriately. For example, flocking birds use sensing to monitor the actions of other birds in their vicinity in order to make local corrections to their own motion. It has been shown that effective flocking results from quite simple local rules followed by each of the birds in the society (i.e., flock), responding to the direction and velocity of the local neighbors (Reynolds, 1987). Such methods of interaction through sensing can be found in use in mobile robot flocking, following, and foraging (Matarić 1995), robot soccer (Werger, 1999), robot formations (Fredslund and Matarić, 2002), and simulations of behaviorally realistic animations of fish schooling (Tu and Terzopoulus, 1994). Other applications of interaction through sensing include human-like physical or visual interaction between physical agents (Murciano and del R. Milian, 1996, Michaud and Vu, 1999; Nicolescu and Matarić, 2000), including the ability to understand and influence the motives of other physical agents (Breazeal and Scassellati, 1999; Ogata *et al.,* 2000).

Interaction Through Communication

The third mechanism for interaction among agents is through direct communication. Unlike the first two forms of interaction described above, which were indirect, in interaction through communication agents may address other agents directly, either in a system-specific manner or through a standard agent communication protocol such as KQML (Finin *et al.,* 1996) or CORBA (Vinoski, 1997). Such agent-directed communication can be used to request information or action from other or to respond to request received from others. Communications may be task-related rather than agent-directed, in which case it is made available to all (or a subset) of the agents in the environment. Two common task-related communication schemes are blackboard architectures (Schwartz, 1995; Gelernter, 1991) and publish/subscribe messaging (Arvola, 1998). In blackboard architectures, agents examine and modify a central data repository; in publish/suscribe messaging, subscribing agents request to receive certain categories of messages, and publishing agents supply messages to all appropriate subsctibers. In some domains, such as the Internet, communication is reliable and of unlimited range, while in others, such as physical robot interaction, communication range and reliability are important factors in system design (Arkin, 1998; Gerkey and Matarić, 2001).

FROM AGENT INTERACTION TO COGNITION AND BEHAVIOR

Given a society of interacting agents, how is complex system-level behavior achieved? Interaction among the society members is not sufficient in itself to produce an interesting or useful global result. In order for the interacting agents to produce coherent global behavior, there must be some overarching coordination mechanism that appropriately organizes the interactions in both space and time.

There are many coordination mechanisms by which to organize the various interactions among agents in order to produce coherent system-level behavior. Self-organization techniques are based on a set of dynamical mechanisms whereby structures appear at the global level of system from interactions among its lower-level components. The rules specifying the interactions among the system's constituent units are executed on the basis of purely local information, without reference to the global pattern, which is an emergent property of the system rather than a property imposed upon the system by an external ordering influence'. (Bonabeau *et al.,* 1997). Methods such as genetic algorithms (Holland, 1975), machine learning techniques such as reinforcement learning (Sutton and Barto, 1998), and distributed constrains satisfaction (Clearwater *et al.* 1991) can all be used to design agents and their interactions such that the resulting behavior meets desired system-level goals. Agents may also explicitly negotiate with each other for resources and task assignments in order to coordinate their behavior. One such approach, employed in human as well as synthetic agent societies, is 'market-based' coordination, where individual agents competitively bid for tasks, which they must either completer or report as broken contracts (Gerkey and Matarić, 2002). Auchtions are a common coordination method in market-based techniques. In auctions, the most appropriate agents are continuously selected and (re)-assigned to various non-terminating to roles (Tambe and Jung, 1999; Werger and Matarić, 2000). In contrast, more symbolic negotiation protocols based on distributed planning involve multiplestages, in which agents first share their plans, then criticize them, and finally update them accordingly (Bussmann and Muller, 1992; Kreifelt and von Martial, 1991; Lesser and Corkill, 1981). Game-theoretic approaches to negotiations have proven effective in situations where agents may be deceptive in their communication (Rosenchein and Zlotkin, 1994). In most complex models of negotiation-based coordination, agents reason about the beliefs, desires, and intentions of other agents, and influence those using specialized techniques (Brandt *et al.,* 2000).

SUMMARY

As was stated earlier, in systems where complex global behavior emerges from the interactions of a society of simple agents, as the complex function of a society of simple agents, as the complex function of the brain emerges from the interactions of a larger society of neurons, the resulting complexity cannot be attributed to any single agent but instead to the interaction of all agents. Agents and their, often local, interactions with each other and with the environment generate the resulting global behavior of the system – no additional external coordination mechanism is needed. Interaction in the agent society can take place through several mechanisms, including interaction through the environment, through sensing, and through communication. In lieu of a central coordinator, a society of agents coordinates its interactions to produce desired system-level behavior through such mechanisms as self-organization, machine learning techniques, or more complex negotiation mechanisms.

The notion that complex behavior can arise from the interaction of simple agents has powerful and far-reaching implications. Many fields, ranging from biology, to artificial intelligence, computer networking, and business management, all find inspiration and motivation from these principles.

References

Arkin R (1998) *Behavior-Based Robotics*. Cambridge, MA: MIT Press.

Arvola C (1998) Transactional Publish / Subscribe: The Proactive Multicast of Database Changes. Proceedings, *ACM SIGMOD International Conference on Management of Data*, Seattle, Washington, 521.

Beckers R, Holland OE and Deneubourg JL. (1994) Proceedings, *Artificial Life IV*, Brooks R and Maes P, eds., MIT Press, 181-189.

Bonabeau E, Theraulaz G, Arpin E and Sardet E (1994) The building behavior of lattice swarms. In: Brooks R and Maes P (eds) *Artificial Life* 4:307-312.

Bonabeau E, Theraulaz G, Deneubourg J-L and Camazine S (1997) Self-organization in social insects. *Trends in Ecology and Evolution* 12(5): 188-193

Brandt F, Brauer W and Weiss G (2000) Task Assignment in Multiagent Systems based on Vickrey-type Auctioning and leveled Commitment Contracting. Proceedings, *Fourth International Workshop on Cooperative Information Agents,* 95-106.

Breazeal C and Scassellati B (1999) How to Build Robots that Make Friends and Influence People. Proceedings, *IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS-99), Kyongju, Korea.

Brooks RA (1991) Intelligence without Reason. Proceedings, *International Joint Conference on Artificial Intelligence*, Sydney, Australia, 569-595.

Bussmann S and Muller JA (1992) Negotiation Framework for Cooperating Agents. In: Deen SM (ed.) Proceedings, *Special Interest Group on Cooperating Knowledge Based Systems*, Dake Centre, University of Keele.

Cao Y, Fukunaga A and Kahng A (1997) Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots* 4: 7-27

Clearwater HS, Huberman, BA, Hogg T, Cooperative Solution of Constraint Satisfaction Problems. *Science*, 254, November, 1181-1183.

Deneubourg JL, Goss S, Franks, Sendova-Franks, A., Detrain C and Chretien L (1990) The Dynamics of collective Sorting: Robot-Like Ants and Ant-Like Robots. Proceedings, *Simulation of Adaptive Behavior*, MIT Press, 356-363.

Donald BR, Jennings J and Rus D (1993)., Towards a Theory of Information Invariants for Cooperating Autonomous Mobile Robots. Proceedings*, International Symposium on Robotics Research*, Hidden Valley, PA, October.

Dorigo M, Di Caro G and Gambardella LM (1999) Ant algoritms for discrete optimiztion. *Artificial Life* 5(3): 137-172

Finin T, Labrou Y and Mayfield J (1996) KQML as an agent communication language. In: Bradshaw JM(ed.) *Software Agents.* Cambridge*,* MA: AAAI/MIT Press.

Franks NR and Deneubourg J-L (1997) Self-organising nest construction in ants: Individual worker behavior and the nest's dynamics. *Animal Behaviour* 54: 779-796.

Fredslund J and Matarić M (2002) A General, Local Algorithm for Robot Formations. *IEEE Transactions on Robotics and Automation*, in press.

Gelernter D (1991) *Mirror Worlds.* New York, NY: Oxford University Press.

Gerkey and Matarić MJ (2001) Principled Communication for Dynamic Multi-Robot Task Allocation. Proceedings, *International Symposium on Experimental Robotics 2000*, Waikiki, Hawaii, December 10-13, 341-352.

Gerkey and Matarić MJ (2002) Sold!: Auction methods for multi-robot coordination. To appear in *IEEE Transactions on Robotics and Autonmation, special issue on Multi Robot Systems*, in press.

Holland JH (1975) *Adaptation in Natural and Artificial Systems.* Ann Arbor, MI: The University of Michigan Press.

Holland OE and Melhuish C (1999) Stigmergy, self-organization, and sorting in collective robotics. *Artificial Life* 5(2): 173-202.

Huberman A and Lukose RM (1997) Social dilemmas and Internet congestion. *Science*, 277:535—537.

Kandel ER, Schwartz JH and Jessell TM (1995) *Essentials of Neural Science and Behavior.* Appleton and Lange, Norwalk, CT.

Kreifelt T and von Martial FA (1991) A negotiation framework for autonomous agents. In: Demazeau Y and Muller JP (eds) *Decentralized A.I.*. Oxford, UK: Elsevier Science.

Lesser V and Corkill D (1981) Functionally accurate, cooperative distributed systems. *IEEE Transactions on Systems, Man and Cybernetics* 11(1): 81-96

Maes P (1990) Situated agents can have goals. *Robotics and Autonomous Systems* 6: 49-70.

Matarić MJ (1994) *Interaction and Intelligent Behavior*. MIT EECS PhD Thesis, MIT AI Lab Tech Report AITR-1495.

Matarić MJ (1995) Designing and understanding adaptive group behavior. *Adaptive Behavior* 4(1): 51-80.

Michaud F and Vu MT (1999) Managing robot autonomy and interactivity using motives and visual communication. Proceedings, *Conference on Autonomous Agents*, Seattle, Washington, 160-167.

Minar N, Hultman K and Maes P (1999) Cooperating mobile agents for mapping networks. Proceedings, *First Hungarian National Conference on Agent Based Computation.*

Minsky M (1983) *The Society of Mind.* New York, NY: Simon and Schuster.

Murciano A and del R Milián J (1996) Learning signaling behaviors and specialization in cooperative agents. *Adaptive Behavior* 5(1).

Nicolescu M and Matarić M (2000) Learning Cooperation From Human-Robot Interaction. Proceedings, *Fifth International Symposium on Distributed Autonomous Robotic Systems* (DARS-00), 4-6 Oct, Knoxville, TN.

Ogata T, Matsuyama Y, Komiya T *et al.* (2000) Development of Emotional Communication Robot: WAMOEBA-2R – Experimental Evaluation of the Emotional Communication between Robots and Humans. Proceedings, *IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS-00), pp. 175-180, Takamatsu, Japan, November.

Reynolds C (1987) Flocks, herds, and schools: a distributed behavior model. *Computer Graphics* 21(4). 25-34.

Rosenchein JS and Zlotkin G (1994) *Rules of Encounter: Designing Conventions of Automated Negotiation among Coinputers.* Cambridge, MA: MIT Press.

Schwartz DG (1995) *Cooperating Heterogeneous Systems.* Dordrecht: Kluwer Academic Publishers.

Segek LA and Cohen IR (2001) *Design Principles for the Immune System and Other Distributed Autonomous Systems.* Oxford, UK: Oxford University Press.

Sutton and Barto (1998) *Reinforcement Learning: An Introduction,* Cambridge, MA: MIT Press

Tambe M an Jung H (1999) The benefits of arguing in a team. *AI Magazine* 20(4).

Tu X and Terzopoulos D (1994) Artificial Fishes: Physics, Locomotion, Perception, Behavior. *Computer Graphics, SIGGRAPH 94 Conference Proceedings*, pp. 43-50, July.

Vinoski S (1997) CORBA: Integrating Diverse Applications Within Distributed Heterogeneous Environments. *IEEE Communications Magazine*, February.

Werger B (1999) Cooperation without deliberation: a minimal behavior-bases approach to multi-robot teams. *Artificial Intelligence* 110: 293-320

Werger B and Matarić MJ (2000) From Insect to Internet: Situated Control for Networked Robot Teams. *Annals of Mathematics and Artificial Intelligence* 31(1-4):173-198.

Further Reading

Bonabeau E, Dorigo M and Theraulaz G (1999) *Swarm Intelligence: From Natural to Artificial Systems.* Oxford, UK: Oxford University Press.

Brooks R (1999) *Cambrian Intelligence.* Cambridge, MA: MIT Press.

Minsky M (1986) *The Society of Mind.* New York, NY: Simon and Schuster.

Russell S and Norvig P (1995) *Artificial Intelligence:: A mordern Approach.* Englewood Cliffs, NJ: Prentice Hall

Glossary

Agnet An autonomous entity situated in an environment with some degree of intelligence, sensing, communication, and actuation capabilities.

Agent interaction Methods by which agents may exchange information, coordinate action, etc.

Interaction through communication A method of interaction among agents through direct communication channels (i.e., radio, Internet, etc.).

Interaction through the environment A method of indirect interaction among agents using the environment as the medium of communication.

Interaction through sensing A method of indirect agent interaction through passive sensing of other agents and their actions.

Society of agents A collection of interacting agents situated in a shared environment.

Keywords: Multi-agent systems, emergent behavior, self-organization, distributed systems, society of agents.