# Participation in Smart Systems

Sabine Thürmel

Munich Center of Technology in Society, Technische Universität München (TUM), Munich, Germany
sabine@thuermel.de

**Abstract.** Both humans and nonhumans can commit to participate in distributed problem solving in smart systems. Therefore the state of the art in collaborative coordination in agent-based smart systems, commitment to joint action, and the potential dysfunctional cooperative behaviour in such social computing systems is described.

**Keywords:** participation, social computing, multiagent systems, smart systems

## 1 The Participatory Turn

Software agents are no longer mere tools, but have become interactions partners. The degrees of freedom built into computational artefacts can materialize in individual acts, mandated actions or collaborative interaction. New capabilities may emerge over time on the individual level. Self-organisation and coalition forming on the group level can occur. New cultural practices and novel institutional policies may emerge. Due to these developments we may speak of a participatory turn when assessing the current division of labour between humans and nonhumans.

Participation of human (and nonhuman) actors in computer-based environments requires the communicative involvement within a computer-mediated and (frequently) open organisational structure where a predefined goal is pursued.

Purely human online participation is explored in a wide variety of research projects e.g. at the Alexander von Humboldt Institute for Internet and Society [1]. The study of the motivation for the participation in e-petitions [2] is one concrete example of such investigations [3].

Participation of nonhumans (and humans) can be found in multiagent systems (MAS). MAS focus on the simulation of complex interactions and relationships of individual human and/or nonhuman agents. They represent a variant of social computing systems. Examples range from swarm intelligence systems to the simulation of sophisticated organisational structures. Social computing systems and especially MAS may be deployed in experimental environments as well as outside the laboratory. In testbed environments they are composed exclusively of software agents. From a computer scientist's perspective they are best suited to offer heuristics for NP-complete problems in planning, optimization and all kinds of knowledge acquisition in open environments where knowledge is local and distributed. They represent a variant of crowd-based socio-cognitive systems (CBSC). As Pablo Noriega rightly remarked after the workshop CBSC may also enable interactions to accomplish activities that need not (may not) be conceived as problems and even when you design one such system to solve one particular problem there needs not be an epistemic challenge. While this is also true for MAS, it must be noted that they are currently mainly used in computational sciences projects - may it be in computational science and engineering, computational sociology or even in legal engineering: "Crowd simulation" systems are useful if evacuation plans have to be developed. Demonstrators for the coordination of emergency response services in disaster management systems, based on electronic market mechanisms, have been built [4]. The Agile project (Advanced Governance of Information services through Legal Engineering) even searched for a Ph.D candidate to develop new policies in tax evasion scenarios based on ABMs [5]. The novel technical options of "social computing" do not only offer to explain social behaviour but they may also suggest ways how to change it.

Moreover, MAS provide a basis to cyberphysical systems. Whereas "classical computer systems separate physical and virtual worlds, cyberphysical systems (CPS) observe their physical environment by sensors, process their information and influence their environment with actuators according to communication devices" [6]. Agent-based cyberphysical systems may be found in smart energy grids [7] or distributed health monitoring systems [8]. These systems are first simulated and then deployed to control processes in the material word. In the latter case humans may be integrated for clarifying and/or deciding non-formalized conflicts in an ad-hoc manner.

Automatic collaborative routines or new practises for ad-hoc coordination and collaboration are established. Novel purely virtual or hybrid contexts realizing collective and distributed agency materialize. Therefore it becomes vital to understand collective coordination in such smart systems.

## 2 Collective Coordination in Current Smart Systems

The individual elements of smart systems may be defined as "miniaturized devices that incorporate functions of sensing, actuation and control. They are capable of describing and analyzing a situation, and taking decisions based on the available data in a predictive or adaptive manner, thereby performing smart actions. In most cases, the "smartness" of the system can be attributed to autonomous operations based on closed loop control, energy efficiency, and networking capabilities" [9]. Examples include the internet of things and the above mentioned cyberphysical systems.

These systems form part of the intelligent infrastructure of today's world. Smart systems have a huge impact on our socio-cognitive environment since „machines don't just replace what we do, they change the nature of what we do: by extending our capabilities, they set new expectations for what's possible and create new performance standards and needs. …Our tools change us" [10, p.5]. Moreover it can be stated that in systems where humans and nonhumans collaborate "we'll outsource some decisions to machines completely, while also assimilating computational rationality into our own decision processes" [10, p.2]. To put it more precisely: "the delegation of control functions to autonomous machines limits the options for human actions and decisions thus increasingly forcing humans into adaptive behaviour" [11, p.28]. Even such adaptive behaviour is a nontrivial task since these systems may be able to adapt to changes in the environment themselves. One option for potentially successful interaction and coordination of humans and nonhumans is offered by the above mentioned multiagent systems.

Current agent-based software systems range from swarm intelligence systems, based on a bionic metaphor for distributed problem solving, to sophisticated e-negociation systems [12]. The software agents demonstrate instrumental rationality, distributed control and division of labour. The commitment of the software agents to pursue a goal is "hard-wired" in most current applications.

## 3 Commitments in Joint Action

Higher degrees of freedom are provided if the commitment to a specific task or even to distributed problem solving is not fixed during execution but may change. In the human case "commitments and predictability in joint action" are a research field in its own right (u. a. [13,14,15,16]). Commitments to joint action may not be taken for granted even in systems characterized by division of labour, distributed control and instrumental rationality. Pacherie distinguishes two variants: "interdependent individual commitments powered by practical rationality" and "joint commitments powered by social normativity: obligations & entitlements" [16]. Humans may display both whereas current technical agents may exhibit the former but not necessarily the latter. It is currently an open question whether synthetic social norms should count as obligations and provide a basis for entitlements outside virtual environments.

However, the fact that current technical agents "lack humans' consciousness, intentionality and free will" (Moor 2006, p. 20) does not mean that they do not possess a degree of "social autonomy in a collaborative relationship". This form of goal-autonomy was defined by Falcone and Castelfranchi as having to two components: "a) meta level autonomy that denotes how much the agent is able and in condition of negotiating about the delegation or of changing it; b) a realization autonomy that means that the agent has some discretion in finding a solution to an assigned problem, or a plan for an assigned goal" [17, p. 407]. Even certain current software agents may possess this kind of social autonomy thus displaying a certain proto-social behaviour. Such software agents need not necessarily be based in a Belief-Desire-Intention (BDI)-model [17, p. 416]. However, if one intends to base a computational model of trust on BDI-agents, an elaborate approach is to be found in [18]. As an aside, it should be mentioned, that one cannot only model trust, but also implement "mischievous" software agents, agents who aim at spreading false information, if suits them. Incidentally, in the biological world this is an exclusively human behaviour [19].

This paper cannot expand on the similarities and differences of current human and technical agents. It must suffice to state that human capabilities and those of technical agents may differ widely. Their acts are based on different cognitive systems, different degrees of freedom and only partially overlapping spheres of experience.

## 4 Dysfunctional Cooperative Behaviour

Even criminal behaviour, deliberate misinterpretations of norms or negligence can be studied in MAS if it is based on bounded rationality. Investigations into machine ethics and the treatment of artificial agents as legal subjects are very instructive when searching for commonalities and fundamental differences in unethical or illegal behaviour between humans and nonhumans. Books as “the law of robots” [20] and “a legal theory for autonomous artificial agents” [21] demonstrate this.

Chopra and White are convinced that “in principle artificial agents should be able to qualify for independent legal personality, since this is the closest legal analogue to the philosophical conception of a person” [21, p. 182]. In their view “artificial agents are more likely to be law-abiding than humans because of their superior capacity to recognize and remember legal rules” [21, p. 166]. If they do not abide by the laws “a realistic threat of punishment can be palpably weighed in the most mechanical of cost-benefit calculations” [21, p. 168].

Pagallo perceives the legal personhood of robots and their constitutional rights as an option only being relevant in the long term [20, pp. 147]. However he discusses at length both human greediness using robots as criminal accomplices and artificial greediness. He states that “in certain fields of social interaction, “intelligence” emerges from the rule of the game rather than individual choices” [20, p.96]. Thus such social and asocial intelligence might be acquired by (rational) nonhuman agents, too.

Moreover investigations into the potential ethical status of software agents have been undertaken (e.g. [22]) and propositions to teach “moral machines” to distinguish right from wrong have been developed (e.g. [23]).

In order to clarify the state of the art in software agents’ ethics Moor’s distinctions between ethical-impact agents, implicit ethical agents, explicit ethical agents and full ethical agents may be used [22]. In social computing the three classes of lesser ethical agents may be found: software agents used as mere tools may have an ethical impact; electronic auctioning systems may judged implicit ethical agents, if “its internal functions implicitly promote ethical behaviour—or at least avoid unethical behaviour” [22, p. 19]; disaster management systems based on MAS systems [4] may be exemplary explicit ethical agents if they “represent ethics explicitly, and then operate effectively on the basis of this knowledge” [22, p. 20]. It is open to discussion whether any software agent will ever be a full ethical agent which “can make explicit ethical judgments generally is competent to reasonably justify them” [22, p. 20]. But the first variants of ethical (machine) behaviour, i.e. proto-ethical systems, are already in place.

Analogous to this classification of ethical behaviour displayed by software agents a wide variety of amoral agents could be implemented. They could range from of unethical impact agents, implicit unethical agents to explicit unethical agents e.g. based on virtue ethics. They could be modelled for use in online games. Such games could provide sheer entertainment, edutainment or form part of the currently so popular

serious games. The latter “have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement” [24, p.5].
Agent-based models allow to model a wide variety of social and asocial behaviour. Yet when transferring the insights gained in the laboratory to real world scenarios, one must proceed with great care. Humans, even if they do not always “follow the rules of the game” are able to perceive others not only as social tools but as valuable peers and act accordingly.

## References


1. Alexander von Humboldt Institute for Internet and Society, http://www.hiig.de/en/, accessed November 24th, 2013
2. E-petitionen for the Bundestag https://epetitionen.bundestag.de/ (2013), accessed November 24th, 2013
3. Alexander von Humboldt Institute for Internet and Society: Motiviation for partizipation in e-petions. http://www.hiig.de/en/project/motivation-for-participation-epetitions/, (2013), accessed November 24th, 2013
4. Jennings, N.: ALADDIN End of Project Report, www.aladdinproject.org, (2010), accessed October 24th, 2013
5. Leibnizcenter for Law: Multi-agent PhD position available. www.leibnitz.org/wp-content/uploads/2011/02/wervering.pdf, (2011), accessed January 24th, 2012
6. Mainzer, K.: Information and Computation of Complex Dynamical Systems. Workshop “Philosophy of the Information and Computing Sciences”, Lorenz Center, http://www.lorentzcenter.nl/lc/web/2010/374/presentations/Mainzer.pdf, (2010), accessed November 24th, 2013
7. Wedde H., Lehnhoff,. S., Rehtanz Chr., Krause, O.: Bottom-Up Self-Organization of Unpredictable Demand and Supply under Decentralized Power Management. In: Proceedings of the 2nd IEEE International Conference on Self-Adaptation and Self-Organization (SASO'08), pp.10-20, IEEE Press, Venice (2008)
8. Nealon J., Moreno A.: Agent-Based Health Care Systems. In Nealon J., Moreno A. (eds.) Applications of Software Agent Technology in the Health Care Domain, Whitestein Series in Software Agent Technologies, pp. 3-18, Birkhäuser Verlag, Basel (2003)
9. Wikipedia: Smart Systems. http://en.wikipedia.org/wiki/Smart_system, (2013), accessed November 24th, 2013
10. Gorbis, M.: Human plus machine. http://www.iftf.org/node/3960, (2011), accessed November 24th, 2013
11. Weyer, J. : Die Kooperation menschlicher Akteure und nicht-menschlicher Agenten. Ansatzpunkte einer Soziologie hybrider Systeme, Arbeitspapier Nr. 16, Soziologische Arbeitspapiere der Universität Dortmund (2006)
12. Woolridge, M.: An Introduction to MultiAgent Systems, 2nd ed., John Wiley & Sons: New York (2009)
13. Bratman, M. E.: Shared agency. In Mantzavinos, C. (ed.) Philosophy of the social sciences: Philosophical theory and scientific practice, pp. 41–59, Cambridge University Press, Cambridge (2009)
14. Gilbert, M.: Walking together: A paradigmatic social phenomenon. Midwest Studies in Philosophy, vol. 15(1), pp. 1–14 (1990)

15. Pacherie, E.: The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency. In Seemann A. (ed.) Joint Attention: New Developments, pp. 343-389. MIT Press, Cambridge (2012)
16. Pacherie, E.: Commitments and predictability in joint action. Presentation at the ENSO III Conference, Helsinki, unpublished manuscript, (2013)
17. Falcone, R., Castelfranchi, C. : The human in the Loop of a Delegated Agent: The Theory of Adjustable Social Autonomy, *IEEE Transaction on Systems, Man, and Cybernetics – Part A: Systems and Humans*, Vol. 31, No 5, pp. 406-418 (2001)
18. Castelfranchi, C., Falcone, R.: *Trust Theory: A Socio-Cognitive and Computational Model*, Wiley Series in Agent Technology, John Wiley and Sons: Chichester: UK (2010)
19. Buttelmann, D., Carpenter, M., Tomasello, M.: Eighteen-month-old infants show false belief understanding in an active helping paradigm, *Cognition*, Volume 112, Issue 2, pp. 337-342 (2009)
20. Pagallo U.: The Law of Robots: Crimes, Contracts, and Torts, Springer, Dordrecht (2013)
21. Chopra,S., White, L.: A Legal Theory for Autonomous Artificial Agents. Ann Arbor, The University of Michigan Press (2011)
22. Moor, J. The Nature, Importance, and Difficulty of Machine Ethics. IEEE INTELLIGENT SYSTEMS, pp. 18-21, IEEE Press, New York (2006)
23. Wallach, W., Allen, C. Moral Machines: Teaching Robots Right from Wrong, University of Oxford Press: New York (2009)
24. Abt, C.: Serious Games. Viking Press, New York (1970)