Editorial

Introduction to the Special Issue on Cognition, Joint Action and Collective Intentionality

Action editor: Ron Sun

Luca Tummolini *, Cristiano Castelfranchi

Institute of Cognitive Sciences and Technologies (ISTC-CNR) and University of Siena, Via San Martino della Battaglia 44, 00185 Roma, Italy

Many remember an Iron Lady claiming that there is no such thing as society but only individuals. Considerably less, we assume, are also aware that the same Lady granted existence also to families, and hence by extension to groups. If this Lady were a social philosopher, she could have continued her inquiry wondering whether families and groups, like individuals, can have intentional states. Can a family *as* such believe something? Can a family intend to do something *as* a family? More generally, what does it mean to ascribe such intentional states to a collective? What kind of attitudes have the members of a group when they jointly intend to achieve a certain result? What does it take to act as a member of a group? This kind of issues and many others related to the social and institutional reality that surrounds us (*pace* our Lady) fall within the scope of a field of research that goes under the rubric of Collective Intentionality.

Intentionality is a philosophical notion used to refer to the distinct property of mental states of being *about* something; mental states like believing, hoping, fearing, wanting, intending and so on are *about* and are *directed* towards something. Such intentional states are used to understand and structure human actions. When I intend to do a certain action, usually it is because I want to achieve a certain result on the background of what I believe about the world and about my capabilities. It is hardly controversial to claim that such intentional behaviour is not only individual. *We* can intend to do something together (e.g. go to the movies tonight) because *we* want to achieve a certain result on the basis of what *we* believe. Joint intentional action is done on the background of such shared intentional states. The philosophical debate on

these issues is lively and several accounts of how such shared intentional states should be analysed have been proposed (Bratman, 1992; Gilbert, 1989; Searle, 1990, 1995; Tuomela, 1995; Tuomela & Miller, 1988). However the import of this field of research spans well beyond the philosophical arena.

It is often claimed in fact that we, as humans, are a cooperative species. Our ability to act together with our conspecifics vastly surmounts that of other animals (including our closest primate relatives) both in its scale and its temporal extension. Only humans are able to engage in complex collaborative activities that are learned from the others (cultural learning). Only humans are able to create complex tools, structured symbol systems (i.e. language) and social institutions (i.e. government and marriage) as means to facilitate such coordination and cooperation. It has been recently proposed that underlying this uniquely human capability for joint action and cooperation there is in fact this capacity to *share* intentional states (Tomasello, Carpenter, Call, Behne, & Moll, 2005).

This special issue has precisely the aim to bring together researchers from different disciplines (philosophy of mind, social philosophy, developmental psychology, evolutionary anthropology, artificial intelligence and ontology) to further the understanding of the role of cognition in collective intentional activities from the micro-level (the cognitive mechanisms enabling joint activities) to the macro-level (how large scale cooperation is enabled and sustained). The conceptual apparatus originally developed in philosophy has in fact the potential to open new exciting directions of research in the cognitive, computational and social sciences.

Exploring the lower bound of the relation between cognition and collective intentionality, *Pacherie* and *Dokic* provide a clarifying analysis of the function of mirror

* Tel.: +39 06 44595 248; fax: +39 06 44595 243.
E-mail address: luca.tummolini@istc.cnr.it (L. Tummolini).

neurons in joint action. Arguing against maximalist interpretations of mirror neurons, they explore their role in the representation of a joint action sequence and in the online mutual adjustment of participants' actions. Mirror neurons are in fact a mechanism we humans have in common with other species and can account for basic joint actions based on mutual motor action understanding. However more complex forms of cooperation need higher forms of action representation and complex mental attitudes that are unique to our species as it is also shown by the other contributions of this special issue.

Humans, in fact, are able to engage in such cooperative activities even very early in their ontogeny. Exploring this human uniqueness in his contribution *Rakoczy* shows that pretend play can be considered as one of the first arenas in which participants exhibit quite sophisticated abilities of collective intentionality. Another early cooperative activity is represented by communication. While the complex mind-reading abilities necessary for intentional communication are not available until the end of the first year of life, *Tirassa*, *Bosco* and *Colle* argue that infants are notwithstanding able to communicate exploiting an innate capability to share their own mental states with their caregivers. If the capacity of sophisticated forms of cooperation is uniquely human, a developmental approach is needed to understand how such unique ability gradually appears in our species.

Differently, *Tollefsen*'s contribution lies at the upper bound of the relation between cognition and collective intentionality. Beyond traditional approaches that analyse collective intentions and collective beliefs as individual mental states owned by group's members, Tollefsen argues that groups can literally be ascribed minds of their own. This challenging argument is defended drawing on an emerging theoretical framework that considers cognitive processes as emerging from the interaction of the agent with its physical and social environment. On a similar background *Bosse*, *Jonker*, *Schut* and *Treur* explore, from a computational and modelling perspective, how tightly interacting group of agents can share external mental states acquiring a representational content which is collective in the sense that it represents something for the group and not necessarily for its members. Even if the issue of a "collective mind" is still controversial, these contributions testify that the cross-fertilization between themes and approaches from cognitive science can help in recasting questions and issues too quickly dismissed by traditional frameworks.

The formal approach is undertaken also by *Hulstijn* and *Maudet* in their contribution. To proceed smoothly a joint action needs a coordination mechanism to signal when each agent should do his share or which action should select. Focussing on conventional situations, Hulstijn and Maudet extend the use of a mechanism first introduced to characterize natural language dialogue. They show how the uptake mechanism can be used as a means for an implicit negotiation of how to engage in the joint action and how joint intentions and obligations of mutual support can unfold during the activity.

As we have seen in our opening paragraph, the research on joint action and collective intentionality always implies some ontological assumptions on which entities exist in our social reality. A foundational problem for example is to identify the minimal conditions under which a set of agents can be regarded as a collective or group acting together for a common goal. The contribution of *Bottazzi*, *Catenacci*, *Gangemi* and *Lehmann* tackles a similar problem from a more general and broad perspective adopting the tools of applied formal ontology to define the concepts used to described social reality. Differently from traditional approaches, the notion of collective is seen as a special kind of the more general notion of "collection", a notion emerging when the "containment" cognitive schema is applied. Agents in a collection become a collective when they are described from the perspective of a unifying plan.

Once the foundational notion of a collective is clarified, one can start the investigation of the properties that a collective can manifest. In their contribution *Conte* and *Turrini* explore precisely one of these properties. Adopting a formal theoretical approach, the notion of collective autonomy is investigated and the conditions under which a collective can become autonomous even from its very members are explored. The cognitive perspective is particularly apt for this task due to the fact that the interrelationships between individual and collective autonomy are mediated by the mental states of the participants. However, at the same time, such an analysis is of interest also from a political perspective enabling a classification of different kinds of collective systems in relation to their level of autonomy and offering a sound framework to develop testable hypothesis about of the evolution of such systems when their level of autonomy shifts.

Cooperative behaviours however cannot be given for granted. There is in fact a class of situations in which acting together is problematic also from a motivational point of view. It is a longstanding research issue in the behavioural sciences (psychology, economics, anthropology, etc.) to provide an explanation of when and why agents choose to incur in personal costs in order to 'help' or benefit another agent or a group. *Henrich* and *Henrich* address the 'puzzle' of cooperation from the perspective of a culture-gene coevolutionary approach and provide a unitary framework able to derive the most accredited theories of cooperation (Kinship, Reciprocity, Reputation, Social Norms and Punishment and Ethnicity). However beyond the ultimate evolutionary level this contribution also derives from such theories a set of proximate cognitive mechanisms causing the agents to act cooperatively. The scale of cooperation we find in humans cannot be understood without appealing to the unique adaptation we have for culture and for cultural learning.

The importance of cultural and social dynamics to explain human behaviour is underlined also by the next contribution. *Ross* advances a model of the dynamics of self-formation under social pressure. Selves are seen as stabilizing devices for social dynamics facilitating the predict-

ability of human behaviour and are at the same time stabilized by those very dynamics. The modelling tools of the mainstream neoclassical economist are used in an original way and are able to further the integration of microeconomic theory with the conceptual and empirical results offered by the cognitive and behavioural sciences. In sharp contrast with what is often stated, neoclassical economic formalism and evolutionary game theory do not incorporate individualistic assumptions and can be adopted to rigorously model human behaviour even from a non-individualist perspective.

If economic theory can be adequately interpreted to incorporate concepts and approaches coming from the collective intentionality debate, *Grosz* and *Hunsberger* in their contribution show how fruitful can be also another kind of cross-fertilization. From the perspective of Artificial Intelligence they attack one of the foundational problems. By offering an operational model and an agent architecture of how a group decides what to do and how updates (i.e. by refining them) previously chosen options (intention updating), Grosz and Hunsberger provide a fresh approach to the issue of whether collective intentions are reducible to individual intentions or not. To capture the dynamics of intention in joint action, two constraints on the reasoning of the agents are specified (the common content and the coordinated cultivation requirements) clarifying the specific role of obligations in mediating group decision-making and intention update.

The role of obligations stemming from joint decisions, promises and agreements is unravelled also by *Miller* in the next contribution. Miller discusses how obligations as reasons for action arise from two different sources: the collective acceptance of the group that a joint decision, promise or agreement has taken place and the collective acknowledgement of an underlying social norm. In doing so Miller succeeds in providing an analysis of the characteristic compelling force of obligations and of their role in the practical reasoning leading to the choice of one's share in a joint action.

The importance of the notion of collective acceptance is also stressed by *Hakli* in his analysis of the notion of beliefs at the group level. When a group of agents decides to take something as the view of their group, it is commonly stated that the group believes something. Given that however such group exerts a voluntary control on the output of this epistemic decision, group beliefs should be classified more correctly as cases of acceptances, or so Hakli argues. Beyond providing an insightful analysis of a core issue in the collective intentionality literature, this thoughtful comparison between two distinct mental states (belief vs. acceptance) is interesting also from a cognitive perspective. Acceptance in fact is a mental state that is needed for a better account of how we act when we act as group members but it is still neglected in many accounts.

A group attitude can also have "constitutive" consequences. There is a class of situations in which if a group of agents believes or accept that something is the case then something starts to exist in our social reality. Institutional facts and actions are precisely this kind of situations where a fact or an action depends for its existence on the shared attitudes of the participants in question (i.e. money, property and marriages). The last two contributions are precisely devoted to the analysis of situations of this kind.

Lagerspetz deals with a puzzling issue related to emergence of new institutional facts. It would seem that if institutions depend on the beliefs of their participants to exist then for a new institutional fact to emerge, the agents *ought* to share false beliefs about its existence in the first place. Requiring that an agent ought to believe something that is false at that point in time is contradictory with the constitutive nature of belief that, at least ideally, ought to aim at truth. To solve this puzzle, Lagerspetz proposes to analyse the mental attitude the agents share in these situations as a case of *recognition* instead of a believing.

The focus on the cognitive process of recognition is also shared by the last contribution. While institutional actions are intuitively performed as usual individual practical actions, *Tummolini* and *Castelfranchi* argue that they are in fact actions done by a collective of agents. However the fact the attendees 'recognize' it, see and treat the action *as* the institutional action, empowers the first agent to achieve a coordinating effect that is actually a co-power of the whole collectivity.

Acknowledgements

References

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review, 101*(2), 327–341.

Gilbert, M. (1989). *On social facts*. New York: Routledge.

Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication*. Cambridge, MA: MIT Press.

Searle, J. R. (1995). *The construction of social reality*. New York: The Free Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–735.

Tuomela, R. (1995). *The importance of us*. Standford: Standford University Press.

Tuomela, R., & Miller, K. (1988). We-intentions. *Philosophical Studies, 53*, 367–389.