Sintelnet WG5 Workshop on Crowd Intelligence: Foundations, Methods and Practices

Proceedings of the Sintelnet WG5 Workshop on Crowd Intelligence: Foundations, Methods and Practices

Marta Poblet, Pablo Noriega and Enric Plaza (eds.)

Barcelona, Catalonia
January, 8-9 2014

Sponsored by:

European Network for Social Intelligence (Sintelnet)

With the support of:

RMIT University
UAB Institute of Law and Technology
Artificial Intelligence Research Institute (IIIA-CSIC)
Institut d'Estudis Catalans - Societat Catalana de Filosofia

Foreword

The SINTELNET WG5 Workshop on Crowd Intelligence was held in Barcelona on 8-9 January 2014. The European Network for Social Intelligence (SINTELNET) regularly organizes Interdisciplinary Workshops to explore and discuss the interplay of humanities, philosophy, social science and information technologies around key social intelligence notions. The goal of this WG5 Workshop was to provide an open forum for discussion on the theoretical foundations, methods and tools, and practices of crowd intelligence. The workshop paid attention to the interplay between humans and computers in different domains and discussed the complexities linked to coordinating crowds. Participants also engaged in conversations on appropriate methods, lessons drawn from case studies, and identified areas for further research.

The Workshop brought together twelve participants mainly from universities and research centers in Catalonia, UK, Ireland, Germany, and Morocco. The workshop received 12 original submissions, covering a number of different domains within the crowd intelligence topic (theoretical foundations, education, prediction markets, arbitration, constitution making, and disaster management). All submitted versions were reviewed by at least two members of the Program Committee. These proceedings finally include ten of these papers.

We sincerely thank the Program Committee members for reviewing all submitted papers and providing candid feedback to improve their revised versions. We are also grateful to the Institute of Catalan Studies (IEC) and its staff for providing the venue and the technical support throughout the sessions. Last but not least, we would like to thank the participants who submitted their papers, provided valuable input in the discussions during the workshop, and afterwards issued the revised versions that are now composing these proceedings.

Marta Poblet, Pablo Noriega, and Enric Plaza

Workshop Organizers

Program Committee

- Pompeu Casanovas (Autonomous University of Barcelona, Catalonia)
- Mark d'Inverno (Goldsmiths, University of London, UK)
- Corinna Elsenbroch (University of Surrey, UK)
- Gert Jan Hofstede (Wageningen University, The Netherlands)
- Mark Klein (MIT, USA)
- Peter McBurney (King's College, London, UK)
- Lin Padgham (RMIT University, Australia)
- Judith Simon (University of Vienna, Austria)
- Harko Verhagen (Stockholm University, Sweden)
- Daniel Villatoro (Barcelona Digital; University of Barcelona, Catalonia)
- John Zeleznikow (Victoria University, Australia)

Table of Contents

Crowd-Based Socio-Cognitive Systems

Pablo Noriega[1] and Mark d'Inverno[2]

[1] IIIA-CSIC, Barcelona, Spain pablo@iiia.csic.es
[2] Goldsmiths College, London, UK
dinverno@gold.ac.uk

Abstract. In recent years there have been several successful examples of crowd-bases systems, used for very different purposes and built on a variety of technological artefacts, some ad-hoc, some generic. We presume it is possible to give a precise characterisation of what we call crowd-based sociocognitive systemsand postulate that it is possible to formulate a framework to model and implement actual crowd-based sociocognitive systemsin a principled way. In this paper we outline the research program, propose the main features of a metamodel for modelling crowd-based sociocognitive systemsand make a call to arms for future development and collaboration.[3]

1 Introduction

Despite the explosion of crowd-based systems and our increasing desire to engage in social activity online, there has been surprisingly little interest from the multi-agent community to use their methods to provide an analysis of such systems. This is surprising, as at the heart of these systems is the notion of coordination, cooperation, emergence, regulation, trust and reputation. This is especially true when we consider that such systems must be modelled from a *socio-cognitive* perspective where individual agents (human or computational) not only behave with some degree of rationality but where that rationality is based on the model that agents have of the other agents in the system. In this paper we will refer to such systems as *Crowd-Based Socio-Cognitive Systems* (or CBSCS) which at their most fundamental level have large numbers of rational agents, each with the ability to model the other agents in the system, and that interact in order to achieve shared or individual goals.

Thinking about the nature of crowd-based socio-cognitive systems in this was might lead readers from a multi-agent system (MAS) background to consider them as wonderful ideal testbeds, with huge amounts of data, to see whether our theories of regulated

[3] This paper reflects ideas from conversations with Harko Verhagen and Julian Padget. Both of them and Mark d'Inverno received support from the European Network for Social Intelligence, SINTELNET (FET Open Coordinated Action FP7-ICT-2009-C Project No. 286370) for short term visits to the IIIA. Work supported by projects PRAISE, ACE, COR, AT and by Generalitat of Catalunya grant 2009-SGR-1434. Some of the work was undertaken as part of the FP7 project in the Technology-Enhanced Learning Program called Practice and Performance Analysis Inspiring Social Education (PRAISE). The Principal Investigator for the project at Goldsmiths is Mark d'Inverno and the partners also include IIIA-CSIC, Barcelona

behaviour can be used to describe how people operate in the virtual worlds of crowd-based systems. It is our belief that there is a wonderful opportunity for our research community to make a concerted effort to try to understand these systems from an MAS perspective. Not only to test out own models and theories with lots of real examples of regulated multi-agent systems but to provide proper tools for the analysis of such systems that can provide us with the proper language for engaging with the huge body of work from sociologists, psychologists, anthropologists and cultural theorists investigating human behaviour in such systems. This paper is a first foray into trying to engage the MAS community in the specification, design, analysis and engineering of such systems as we believe that the community has much to offer.

Not only is there an opportunity to understand the social activity that happens online but also an opportunity to build a language to relate the virtual worlds of crowd-based systems to the real emotional and physical worlds we inhabit with our systems. We need tools and analysis techniques for exploring this relationship, for a technical language for describing and understanding the meaning and ramifications of various kinds of online social activity into the real world, and even to start map out the space to see if there are new kinds of opportunity for building new kinds of online systems supporting new kinds of social activity. Not surprisingly many systems are driven by technological possibility and financial gain rather than social good. If we, as researchers in this area, can provide a conceptual framework to map out what is happening within these systems from a multi-agent perspective might there be an opportunity to take part (as just one example) in the discussions about the social responsibility of such systems.

The current situation appears to us that we have no clear technical grounding to adequately describing such systems from an MAS socio-cognitive perspective. We do not have the models, language, theory or tools for the description, analysis and creation of such systems. In order to try to form a bridge between the work in MAS and regulated systems and the plethora of emerging systems this paper sets out to define a conceptual framework for such systems using the language of agents, norms and communication in order to do so. By doing so we hope to seed an emerging research area concerned with developing theories, tools, languages and methodologies for designing such systems. If we can demonstrate the applicability and usefulness of our modelling techniques then we may potentially provide a bridge to other subject areas, such as sociology, in order to have a more rounded understanding of the social and psychological responsibility that should be considered in the design of such systems.

1.1 Socio-Cognitive Systems

Socio-Cognitive Systems (SCS) are characterised across the following characteristics. The reader should consider this to be an indicative list of the qualities of the class of SCS rather than an exhaustive one.

- Dimension 1. The system contains *agents*. Agents are either computational or human and can exhibit purposeful behaviour.
- Dimension 2. The population with a system may be a *mix* of human and software agents.
- Dimension 3.. The agents have a *model* of the world in which they operate.

- Dimension 4. The agents within the system are *rational* in that they are capable of choosing different courses of action based on their own models (however simple or complex these may be).
- Dimension 5. The agents are *social* in that they interact with other agents.
- Dimension 6. The agents are have *social models* (either complex or simple) of some of the other agents in the system,
- Dimension 7. At least some of the agents are *socio-cognitive* in the sense that they based their decisions on some decision-making process which takes into account the models of the social world in which they are situated. This includes the capability to plan for future desire states in the environment whilst taking into account the motivations and models of other agents. Such agents can reason about who to collaborate with other agents to achieve individual and joint goals.
- Dimension 8. The agents have *social capabilities* including potentially awareness and models of others, an ability to understand the norms of a system and adopt attitudes relating to norm-compliance and the ability to have altruistic goals
- Dimension 9. Any such system is defined by the system of interacting agents which means that the state of the system can never been known in full as there is no general access to the internal state of agents. This is often referred to as *opacity*..
- Dimension 10. Agents may enter and leave an SCS at any time. It cannot be known either by the designer of the system or by other agents which agents may join or leave. Agents may often be able to join or leave without it being known to other agents.
- Dimension 11. Such systems as *regulated* either intrinsically because of the way the system is designed, the way that some agents have been specified to operate or naturally through the agreement of agents within the system. The point about regulated systems is that not all actions are available to all agents at all times which enables more effective *social coordination* to be facilitated.
- Dimension 12. Agents are *autonomous* and so march to the beat of their own drum and so are not necessarily socially-considerate, benevolent, or honest and so may fail to act as expected or desired or promised.
- Dimension 13. All interactions are mediated by technological artefacts and may therefore be wrapped as communicative acts or messages. Systems that have this property are referred to as *dialogical*.

This list includes the characteristics of systems that have been investigated by the research community looking at *regulated multi-agent systems* and attempts to reflect the recent discussion on Socio-Cognitive Technical Systems that is arising from the Sintelnet project (see Positon Papers in www.sintelnet.eu/wiki/index.php/Sourcebook and in particular[2, 1])

1.2 Towards a description of Crowd-based Socio-Cognitive Systems

According to the work of Surowiecki [9] crowd-based systems are concerned with connecting or collecting diverse collections of independently deciding individuals. The basic thesis is that diverse collections of independent autonomous agents with different models, perceptions, motivations and rationality can often analyse or predict scenarios

or data more effectively than individuals even when those individuals are specialists in their area of expertise.

He discusses three types of system advantages

1. Cognition. This is about how crowds can make judgements through thinking and information processing faster than individual experts
2. Coordination. This is whether social or physical coordination can emerge naturally in large communities of agents. It relates to how a shared view of the reactions of a community provide often accurate judgements about how the community will react to events.
3. Cooperation. Again this relates to the emergence of ways in which trust and reputation can emerge naturally without needing a top-down set of norms for social cooperation and co-ordination of activity.

There are then four criteria to distinguish wise crowds from unwise crowds that we summarise using our own MAS terminology as follows.

1. *Diversity of opinion.* Each agent has its own private information that cannot be known by others.
2. *Independence.* Agents' opinions are not completely determined by the opinions of those around them, agents also have a degree of autonomy in the way they form their opinions.
3. *Decentralization.* Agents have different local knowledge and different perceptions of their local environment.
4. *Aggregation.* Some mechanism exists for turning private judgments into a collective decision.

So the question becomes what further characteristics do we need to add to our descriptions of *Socio-Cognitive Systems*. In the four criteria above we already have items 1, 2 and 3 from our own definition. So we add two more

Crowd-based Socio-Cognitive Systems are Socio-Cognitive Systems which have the following three sometimes rather nebulous characteristics.

- Dimension 14. There is a significant population of agents.
- Dimension 15. The system allows for norms (for social cooperation and coordination), trust and reputation to arise natural.
- Dimension 16. The system provides mechanism for turning individual analysis, goals or work into collective analysis, goals and work.

Crowd-based Socio-Cognitive Systems are thus systems that exhibit some features of what is accepted as crowdsourcing or crowd-based behaviour systems but have the distinguishing characteristic that individuals need to reason about themselves and their social environment, because their behaviour is affected by that social environment and also because with their behaviour they may influence the social environment to some extent. A system which implements all of these characteristics also appears in this conference [12].

1.3 Motivation of our work

Such systems are a new phenomenon that involves thousands and sometimes millions of people. What is striking, and perhaps a little unnerving too, is that most such systems are being developed without any theoretical underpinning and in such a way that it is often not easy to see what is underneath the bonnet. Of course there are many kinds of definitions and descriptions but they are not necessarily conducive to a principled analysis or design of CBSCS. Our motivation is to want to understand them in principled ways and describe them in a systematic and formal way that can then be used for the design and implementation of such systems based on principles developed from work in regulated MAS. Our overall ambition is to be part of the design systems of these systems where we can more clearly articulate the social benefit for those participating within it Our wide-ranging goals to support this ambitions can be described as at least containing the following enumerated below.

1. How can the MAS community take part in the design of CBSCS?
2. What could we offer to the design of such systems in general?
3. How should we present our work in such a way that any system of designers would ever care to notice it?
4. Could we imagine collaborative research projects with designers where research could be developed through the process of design about the nature of designing such systems and understanding how MAS techniques could be applied?
5. In this light, does it make sense to define the universe of SCS (in terms of normative systems and institutions?)
6. If this is not possible (look at the thousands of agent definitions that derailed many scientific and investigations because of lack of a common conceptual and definitional framework from which proper scientific enquiry and engineering systems integration could take place) would it be possible to identify and define the key concepts of CBSCS systems that is useful, engaging and relevant?
7. How might we turn this round and highlight to the MAS community the potential of CBSCS systems for investigating social systems from an MAS perspective?
8. What is a good way to map out the key research issues for a regulated MAS approach to analysing CBSCS systems?
9. Could we identify the potential influence into the design of CBSCS systems for new kinds of collective activity for communities such as ours?

We believe that with a combined effort we can produce answers to these questions trough models., metamodels, tools, design methodologies for interaction and interfaces that would underpin the principled design, specification and analysis of such systems. The way to do this, we believe, is to undertake an empirical study of such systems, and attempt their characterisation using our conceptual framework (models, data structures and languages).

We want to consider the social needs of users and the kinds of actions users want, to coordinate within communities and use this as part of the design of new systems. If we can develop clear, useful and principled models of such systems that can be used by designers and communities then the emphasis can be much more focussed on the end use than the specific goals of the engineer who builds it.

Not only do we want principled design and tools and interfaces that make it easy to build these things, but also so that potential communities can understand the range of options that are available to them. So there is a political element here to develop models and design methodologies that can put the user and the community in control of the system they want to be a part of, rather than be part of a system that has been developed by sets of engineers from multi-national companies with less clear motives.

At heart is the question of not wanting to be left out of the design and investigation of crowd-based systems when the MAS community has put such effort into understanding them. Indeed many of us joined the research effort into understanding what social action is from a multi-agent system perspective because we wanted to understand and support cooperation and coordination. How could we coordinate the activities of agents with different personal goals coming together for a common need? Political and social activism, environmentalism, local community, learning, fun and games—we want to be able to build such systems to support a whole range of coordinated social activity as well as investigate the potential range of social activity that can be supported by such systems. If now is not the time that the work we, the MAS community, have developed over the last 20 years or so in understanding social systems from a computational perspective then, when will it ever be?

2 Background

The research programme we envision is to achieve an understanding of socio-cognitive systems, in general, so that we may eventually be able to design new systems with a principled approach. We propose to address the general problem, first by delimiting the universe to an explicit set of features that may allow us to decide whether a given system–existing or being designed—belongs to that universe, and second, developing an abstract understanding of what is common to these systems by separating two fundamental objects of concern: the actual agents (be they human or artificial) and the social space where these agents interact.

On modelling agents, we will assume that these agents may need to reason not only about themselves but also about that social space, because the social space influences and determines in some sense their actions, and also because agents do influence the social space. Thus agents would have to exhibit capabilities or cognitive dispositions to be aware of other agents, to interpret what is the state of the world, and to hold expectations of what possibilities of action are available (for itself or for other agents) and what the consequences of those actions may be. Likewise, the modelling of the social space determines what inputs and outputs will be accessible to the agents, and therefore one has to device the means to model what the social space "affords" agents to act upon and to be aware of, and the means by which the space may influence the activity of agents. In other words, what objects exist in that space, how agents communicate, how can activities may be coordinated, what types of organisations can an agent belong to, and so on. [4] Consequently, in abstract terms, we shall speak of meta-models of socio-cognitive agents and metamodels of social spaces. For each of these metamodels

[4] We use the notion of *affordance* in the spirit of Norman [8].

we would then attempt to produce precise, even formal, descriptions that would allow the specification (and formal analysis) of actual models of agents and of social spaces. Metamodels that in turn need to be accompanied by technological artefacts that enable the actual implementation of socio-cognitive systems where artificial or natural agents pullulate in an artificial social space.

We realise that attempting to find a single metamodel for agents and for social spaces is at best impractical. However we glimpse the possibility of sketching some generic metamodels for families of socio-cognitive systems (SCS). For example, on-line marketplaces, massive on-line role playing games, mixed-level simulation, or policy-making support systems. We postulate that one of these families are crowd-based SCS.

Consequently, we may rephrase the next steps in our research programme, in terms of CBSCS, in the following four steps:

1. Compile a corpus of CBSCS
2. Understand what is "structural" of crowd-based SCS.
3. Map the universe from these exemplars
4. Outline a conceptual framework, identify adequate tools and methodological guidelines

2.1 Our bias

In two previous papers [10, 7] we discussed the basic tenants of our research programme: (i) a three-fold view of socio-cognitive systems, the notion of shared context and the relationships among the three views, its abstract (*platform-independent*) model, and the implementable (*platform-specific*) model. (ii) The relationship between metamodel, environment, computational architecture and platform. (iii) The separation between agents and social space and between design environment and enactment environment of the socio-cognitive systems.

We are confident that the approach we propose has some hope to succeed because we already have done a similar task for the abstract notion of "electronic institutions" ([3]). In this case, the universe of systems it may model and implement is close to the very general notion of socio-cognitive systems that we have as the ultimate goal of our research programme.

After a decade of development of the conceptual framework, associated tools and a considerable number of application cases, we were able to produce a formal metamodel, and the associated technological artefacts integrated in a working development platform [3]. The electronic institution framework is actually only applicable to the social interaction space, since one of its key assumptions is that participating agents are black boxes who are able and willing to comply with the space conventions, thus there is no commitment to any agent model. Thus, the space itself is a regulated multiagent system where agents interact through speech acts that are organised as interrelated conversations (or "scenes") where the illocutionary exchange is prescribed with regimented procedural rules.

Figure 1 summarises the actual metamodel. Briefly speaking, it includes two parts: the "dialogical framework" that provide participants those elements that are involved in action (these are "dialogical" because interactions are understood as conversations); and

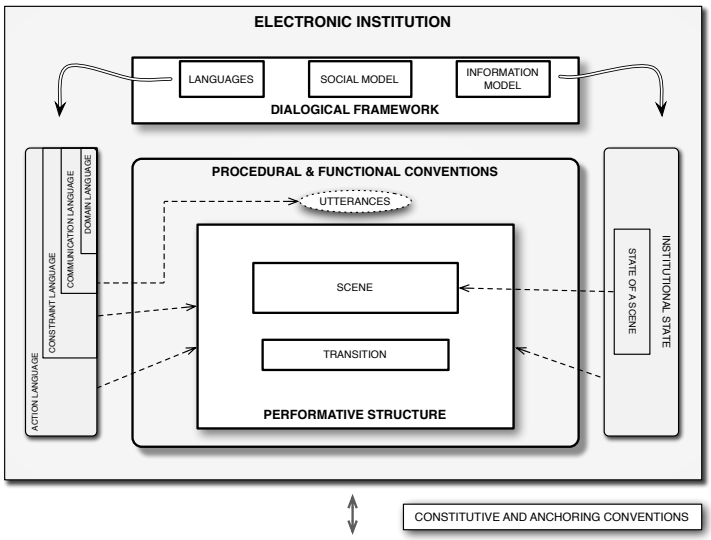

Fig. 1. The conceptual model of electronic institutions. After [6].

the regulations that govern those actions. It is beyond the scope of this paper to explain these components in detail but we shall make some quick remarks on the *affordances* of this metamodel, in order to give a flavour of what needs to be made explicit and properly formalised in order to provide *ad-hoc* metamodels for crowd-based SCS:

- Ontologies about the domain of interaction, of possible actions, its preconditions and effects and how these are related to the social model. These ontologies are captured and reflected through a collection of languages that are part of that dialogical framework.
- A social model that in this case is limited to a finite set of roles and relations among these.
- A means to specify and structure local contexts (scenes) where a subgroup of the participants may participate in a collective activity (conversation) and a structure that establishes how agents playing specific roles may move between scenes or participate simultaneously in different scenes (the performative structure)
- A normative model that allows to specify and regiment those conventions that regulate the conversational exchanges and the pragmatics of those exchanges.
- An information model that includes (i) a collection of variables and parameters that refer to the domain of interaction, participants in the interactions and interactions themselves, (ii) the "state of the scene" that keeps track of the evolution of the conversation and the "state of the institution" that includes those states and, in essence, keeps track of the values of those variables and parameters at every point in time.

3 Towards a framework for crowd-based SCS

3.1 Outlook

We aim to define a framework to model and eventually implement crowd-based SCS. Consequently, we distinguish between the conceptual framework that is used for modelling a crowd-based SCS and the artefacts that implement actual models. With the conceptual framework we intend to model social spaces that enable participating agents to perform a variety of social tasks, that are particular of crowd-based SCS. Namely, define and broadcast a collective challenge, accept a call, perform the individual tasks involved in the challenge, compile the responses to the challenge, assess whether the tasks are properly completed and the rewards entailed by accomplishing a task are being properly granted, etc. Along this conceptual framework we presume that there will be technological artefacts that serve to implement actual systems that are modelled with the conceptual framework.

In order to enable crowdsourcing tasks, the conceptual framework needs to "afford" the designer, and ultimately the participating agents, the means to specify and apprehend the social space where those tasks take place. Because we intend to use the conceptual framework to model crowd-based SCS, we will refer to it as a *metamodel*.

Although it is beyond the scope of this paper to give a formal definition, a metamodel is a collection of affordances each of these specified as a class of formalisms. Loosely speaking, the metamodel should *afford* agents those aspects of the social context that enable them to act proficiently. Hence, the metamodel should include means to formalise (i) the contents of the shared context, (ii) the features of the agents that participate in it, (iii) their interactions and (iv) the overall behaviour of the system.

These affordances should involve, for example, the means to establish and become aware of how much is revealed of the individual's identity, and how individuals become aware of other participants in a given activity. Moreover, agents would need to have means to communicate in non-ambiguous terms, and that implies that, at the very least, the metamodel shall afford some shared ontology, communication language and interaction model. Coordination is usually achieved by organising the crowdsourced system into activities, usually collective, that may be executed in parallel or following certain rules that link them by time, causality or whatever (for instance, in Wikipedia, editing articles, handling disputes, quality review,...). In most SCS one may reify a social structure, however primitive, where individuals are meant to play roles (e.g., *editor*, *administrator*, *bureaucrat* in *Wikipedia*) that entail some capabilities and are frequently subject to rules that apply to individuals only while they play that role. More sophisticated social structures involving groups, hierarchies and organisations are used to better coordinate complex projects.[5] Likewise there are usually means through which the designer may specify the "rules of the game" (see for example the "five pillars" of

[5] For instance in the crowd-sourced drafting of the Constitution of Iceland, Parliament created a Constitutional Council, whose members were citizens voted by the population, this Council was then organised in three working groups who produced recommendations to the council, that received comments from the public through the Council's webpage, and were submitted in turn to the Council and incorporated into a "process document" that was again open to public comments until a "draft proposition" was made by the Council. http://stjornlagarad.is/english/

Wikipedia and the associated guidelines and polices) and how to enforce and update them; and for participants to be informed of these rules, of their application and of any possibility of changing them.

3.2 A tentative list of affordances for Crowd-Based SCS

We presume it is possible to propose a single conceptual framework that allows the modelling of a large class of CBSCS. That is an empirical question that can only be answered with a systematic analysis of existing crowd-based sociocognitive systemsand a serious attempt at the design of new ones. Nevertheless, after our experience with electronic institutions [3], a discussion of normative MAS [6] and a superficial inspection of three other paradigmatic SCS classes (gaming, simulation and policy-making[10, 7]) we propose a tentative list of affordances that we claim should be included and made precise in an abstract conceptual framework for crowd-based sociocognitive systems.

Agent types and agent socio-cognitive models Probably three types are enough: (A1) humans or real world organisations that commission or execute a task; (A2) software agents that commission or execute a task; (B) server agents that perform management or support functions (for instance, enact a collective process, search for potential executioners, remind executioners of pending tasks, evaluate task performance, perform police-like and time-keeping functions). For each of these types we may want to make explicit the socio-cognitive dispositions that agents have or should have.

Domain ontology. This will include the elements the are used to define the content of collective contexts and interactions. For example, in `change.org` the ontology would include petition, signature, motivation, proposer, number of signatures. in Wikipedia, articles, review, update, guideline, editor, bureaucrat, dispute, etc.

Languages. These are needed to define the behaviour of the system and the way it is regulated. These may be organised as a hierarchy of languages that starts with a *domain* language (to refer to the domain ontology); *communication* language, *action* languages (description of an action); f *constraint* languages (preconditions and post-conditions of actions); *normative* languages (procedural, functional or operational directions; behavioural rules,...) and so on, depending on the complexity of the crowd-based system

Interaction contexts are needed to define separate collective activities and their interrelations. They are ideal locations or activities where several agents interact simultaneously, sharing the *same state*) (they correspond to EI's "scenes"). For instance a Turk challenge, a Wikipedia dispute over an article, the trading process of a prediction market. When a CBSCS involves several activities, it should be possible to specify how several local context may be connected and how individuals may move between them.

Actions . Atomic actions like "speak", "move to another interaction context"'; complex actions like "broadcast", "execute", e.g. in Wikipedia: create, edit, censor an article, introduce a guideline or a norm, participate in a dispute.

Information structures. The (shared) state of the system (the value of each and every variable that may change through the action of some agent or the passing of

time) and the shared state of local contexts (generally, subsets of the state of the system); profile of participants, performance indicators, data structures associated with composite actions, ...

Social constructs. Describe the way individuals are related among themselves and also serve as means to refer to individuals and groups of agents by the role they play, rather than by who they actually are. These may include: roles; relations among roles (n-ary relationships between individuals as well as higher-order relationships. i.e, groups, hierarchies of roles, power relationships and so on); organisations (groups plus coordination conventions).

Regulatory system. To allow top-down or bottom-up articulation of interactions: e.g., norms of different types (procedural, constitutional, rules of behaviour,..) with their associated features (relationships between norms, incentives, effects of compliance and non-compliance,...)

Inference. Assumptions about different ways of inferring intended or observed behaviour. Ways to model reasoning under *uncertainty* and alternatives to classical forms of inference like argumentation of coherence.

Social order mechanisms. To allow top-down or bottom-up governance. Among these: regimentation (rendering some actions impossible, strict application of sanctions,...); mechanisms for assessment, evaluation, prosecution and punishment of non-compliance; social devices (trust, reputation, prestige, status, gossip); policing devices (law enforcement),...

Performance indicators. To measure the behaviour of the system by the designer and qualified participants.

Evolution. Means by which the system may change over time (adaptation of agents, bottom-up, negotiated, external change of system regulations) and devices involved in producing that change happens: performance indicators, normative transition functions and such.

4 Concluding Remarks

4.1 The affordances challenge

The list of affordances we used in the previous section is biased by our previous attempts with games, simulation and policy-making. This list is inadequate for two main reasons: It is a mix of heterogenous notions : languages and ontology, for instance, are not like “inference” or “social order mechanisms”, which are easier to assimilate to a collection of formalisms each providing a different flavour to the same type of functionality. Second, it is not complete, or not explicit enough. For instance, What are the affordances that explain the main functionalities of the *Amazon Turk*? or Where does one capture the requirements of *crowdness* or the way one may filter the participation of given individuals?

There is another matter to ponder: Should the list of affordances be different in crowd-based SCS and other classes of SCS (say games or electronic markets)? What would the advanges be for one answer over the other? How may this issue be settled?

4.2 The expressiveness trade-off.

As suggested above, an affordance should entail those conceptual elements that when properly specified allow for a precise specification of the way the CBSCS enables certain functionalities. Ideally, that precise incarnation ought to be made operational through technological artefacts that implement those functionalities. As discussed in [6] the metamodel may choose among several available formalisations of a given affordance, depending on the functionality desired, and each formalisation will be implemented possibly in different ways choosing among available artefacts. However, there is a "whorfian" expressiveness paradox: once a formalism is chosen, then the implementation of the affordance is conditioned by the chosen formalism and the corresponding choice of artefacts, and *vice-versa*.[6]

When there is a collection of artefacts that are coherent, interoperate and are integrated on a working computational architecture, they are usually called a *platform*. Ideally, the metamodels we foresee should allow a formal representation of *platform-independent models* whose implementation will eventually depend on available artefacts. If one is lucky enough to have a platform that integrates those artefacts, including a specification language, the transcription of those platform-independent into platform-dependent implementation is a relatively simple task, modulo the"expressiveness paradox."

As with other representation of a class of problems (e.g. norms, planning, work-flows), there is a trade-off between the generality of the framework that is used to describe (and formalise) a sub-class of those problems and the ease with which a sub-class of those problems is represented in the framework.[7]

Currently, many working CBSCS are not platform-independent and some forms of crowd coordination have proliferated because practical platforms are available. Let's examine three paradigmatic cases, of platform dependent classes of systems: *Mechanical Turk*-enabled projects, prediction markets and *Ushahidi* crisis mapping .[4] Two obvious remarks apply to the three examples: First, the three examples are designed using platform-dependent models. That is, they are founded on particular technological artefacts (or platform) that restrict CBSCS design to involve only those features that are afforded and implemented by the particular artefacts (or platform). In other words, they can model only the systems that can be implemented with the corresponding artefacts. Second, these platforms are the nuclei of the corresponding crowd-based SCS but are not necessarily powerful enough to model and implement an actual crowd-based SCS and, as we shall show, not enough for the three examples at hand. Now, in particular,

- The Amazon *Mechanical Turk* (https://www.mturk.com) is a full-fledged *platform* that applies to different sorts of microwork, and its metamodel affords means for specifying, enacting and evaluating projects that have one single activity, performed

[6] The usual, but not altogether false, misreading of B.L. Whorf (in [11]) as the postulate that the structure of anyone's native language strongly influences or fully determines the worldview he will acquire as he learns the language.

[7] We have come across that trade-off in the case of electronic institutions where the rich language for describing transitions between scenes are unnecessarily cumbersome in work-flows that can be hard-wired(in e-commerce, for instance).

as a mutitude of micro-tasks by populations whose members may be filtered into the project from an open pool. It is Amazon, and not the Turk platform, who provides the additional services that allow the management of projects through other artefacts.
- Prediction markets may be implemented in different platforms—for example the Iowa Electronic Market (tippie.uiowa.edu/iem/) or iPredict (www.ipredict.co.nz/)—each of them is a regimented implementation of a particular futures market with its own notion of contracts, its trading protocols, entry and compensation requirements, the definition of an event to predict. and so on They are open to traders that fulfil some requirements but there is a unique model for each platform, and no more affordances may be included in their CBSCS.
- The *Ushahidi* platform (http://ushahidi.com/products/ushahidi-platform) gravitates around the *Usahidi* map, an artefact that consists of a graphical representation of geo-referenced data that belong to an explicit taxonomy of relevant events. This platform supports the collection, interactive mapping and visualisation of events but the implementation of an actual crisis follow-up SCS needs to be complemented with other ad-hoc artefacts for integrating, filtering and deciding on how to use incoming data. Thus, while the mapping activity has a single model, the metamodel for crisis management needs to afford other features depending on the organisational structure of the management organisation and the functions it assumes during the crisis.

We presume that it is possible and useful to strive for platform-independent models, and we believe that it is possible to meet this challenge. Evidently, a good analysis of currently available platforms is an immediate step to take towards the affordances challenge.

4.3 A call to arms

This is simply to say to our colleagues in the research field of normative multi-agent systems and in areas of the social sciences, that the theories, models and methodologies that we have developed in the last 20 years or so can be brought to the design and analysis of the increasing plethora of crowd-based socio-cognitive systems. There are several goals that we outlined above but in the immediate term there are several issues we need to address first that we consider here.

We propose to start with an empirical study of existing crowd-based sociocognitive systemsthat should enlighten the development of an abstract conceptual framework for modelling CBSCS, or support the convenience of developing several metamodels. We think that the achievement of a clear description of that conceptual framework should open the way to the assembly of technological artefacts and the development of crowd-based platforms that allow design and use of CBSCS whose properties can be ascertained formally and ideally proven to be correctly implemented.

The steps to follow, in our opinion, are:

1. Compile a set of CBSCS “examples” that map out the space of such systems clearly.
2. Produce a more rigorous description of the universe of CBSCS by developing the 16 dimensions we have posed above.

3. Identify precise descriptions of the entities (in each example) that permitted modelling and implementation of affordances (also to be described with precision) present in the models.
4. Develop this model and framework using formal methods in order to ensure clarity, rigour and portability of our ideas. (Using techniques for developing formal conceptual frameworks such as [5] for example.)
5. Contrast this analysis with the tentative list of characteristics we have developed above and refine these.
6. Based on that outcome, draft a single conceptual framework.
7. Classify existing and future CBSCS and formulate classification schemata to apply these ideas.
8. Identify useful technological artefacts to implement the components of the conceptual framework.
9. Develop proof-of-concept CBSCS, analyse and postulate preliminary methodological guidelines for those interested in designing such systems.

We hope that we have made it clear why our models and method are of interested to those participating in the workshop as well as a introducing a tentative road map of how an interdisciplinary community might emerge through contribution to the explicit components that we identify in the roadmap of research above.

References

1. Alexander Artikis Andrew Jones and Jeremy Pitt. The design of intelligent socio-technical systems. *Artificial Intelligence Review*, 39(1):5–20, May 2013.
2. Cristiano Castelfranchi. Minds as social institutions. *Phenomenology and the Cognitive Sciences*, pages 1–23, forthcoming.
3. Mark d'Inverno, Michael Luck, Pablo Noriega, Juan A. Rodriguez-Aguilar, and Carles Sierra. Communicating open systems. *Artificial Intelligence*, 186(0):38 – 94, 2012.
4. Jessica Heinzelman, Rachel Brown, and Patrick Meier. Mobile technology, crowdsourcing and peace mapping: New theory and applications for conflict management. In Marta Poblet, editor, *Mobile Technologies for Conflict Management*, volume 2 of *Law, Governance and Technology Series*, pages 39–53. Springer Netherlands, 2011.
5. M. Luck and M. d'Inverno. Structuring a Z specification to provide a formal framework for autonomous agent systems. In J. P. Bowen and M. G. Hinchey, editors, *ZUM'95: 9th International Conference of Z Users, Lecture Notes in Computer Science*, pages 47–62, Heidelberg, 1995. Springer-Verlag.
6. Pablo Noriega, Amit K. Chopra, Nicoletta Fornara, Henrique Lopes Cardoso, and Munindar P. Singh. Regulated MAS: Social Perspective. In Giulia Andrighetto, Guido Governatori, Pablo Noriega, and Leendert W. N. van der Torre, editors, *Normative Multi-Agent Systems*, volume 4 of *Dagstuhl Follow-Ups*, pages 93–133. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2013.
7. Pablo Noriega and Julian Padget. Approaching social coordination from a normative perspective. In Harko Verhagen, editor, *Analytical Sociology, Social Coordination and Informatics*, 06/06/2013 2013.
8. Donald A. Norman. Affordance, conventions, and design. *interactions*, 6(3):38–43, May 1999.

9. James Surowiecki. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Little, Brown, 2004. ISBN 0-316-86173-1.
10. Harko Verhagen, Pablo Noriega, and Mark d'Inverno. Towards a design framework for controlled hybrid social games. In Harko Verhagen, Pablo Noriega, Tina Balke, and Marina de Vos, editors, *Social Coordination: Principles, Artifacts and Theories (Social.PATH)*, pages 83–87, Exeter, UK, 03/04/2013 2013. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.
11. Benjamin Lee Whorf. The relation of habitual thought and behavior to language. In J.B. Carroll, editor, *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*, pages 134–159. MIT Press, 1956. ISBN 0-262-73006-5.
12. Matthew Yee-King and Mark d'Inverno. Pedagogical agents for social music learning in crowd-based socio-cognitive systems. In *Crowd Intelligence: Foundations, Methods and Practices*. European Network for Social Intelligence, Barcelona, January 2104.

IT Enabled Crowds: Leveraging the Geomobile Revolution for Disaster Management

Marta Poblet[1,2], Esteban García-Cuesta,[3] Pompeu Casanovas[1]

RMIT University, Melbourne VIC 3001 Australia
[1]UAB Institute of Law and Technology, Campus UAB, 08193 Bellaterra, Spain
[3]iSOCO, Madrid, Spain

marta.pobletbalcell@rmit.edu.au, egarcia@isoco.com,
pompeu.casanovas@uab.cat

Abstract. This paper offers an exploratory approach to crowdsourcing methods, tools, and roles based on different levels of involvement of users, skills required, and types of data being processed (from raw data to highly structured dada). The paper also aims at refining different crowdsourcing categories and opening up a theoretical discussion on the advantages and limits of using crowdsourcing methods and technologies in disaster management activities.

Keywords: crowdsourcing, crowdsensing, micro-tasking, data, online platforms, disaster management

1 Introduction

Mobile technologies, location-based services, and geospatial data are currently fuelling the geomobile revolution that brings up to the front the relational dimension of space. A growing plethora of sensors and applications in our mobile devices are constantly producing data, both for us and about us: geospatial coordinates in digital maps, routes, check-ins, etc. Such geospatial data are the digital anchors from which we interact with our immediate context. These data also contribute to augment our reality with added layers of information. Likewise, our real-time geopositioning in a given space triggers the digital footprints that we leave as we interact with our immediate context.

Combined with different crowdsourcing approaches and methods, the geomobile revolution also creates unprecedented opportunities for research, industry, and social change. This paper explores how state-of-the-art technologies are opening up new avenues for citizens' involvement in disaster management initiatives with different crowdsourcing roles.

2 Crowdsourcing: the power of the crowds

The term crowdsourcing was first coined by Jeff Howe in 2006 when referring to "the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call" [11]. Since Howe's first definition, different crowdsourcing categories, dimensions, and typologies have been discussed in the literature [5,18,7,6,19,8,17]. Other studies consider crowdsourcing as part of the broader paradigm of collective intelligence [12] and review the similarities, overlapping and gaps between human computation, crowdsourcing, social computing and data mining [16].

The three key elements intersecting in Web-based crowdsourcing are the crowd, the outsourcing model, and advanced Internet technologies [17]. According to their definition, "crowdsourcing is a sourcing model in which organizations use predominantly advanced Internet technologies to harness the efforts of a virtual crowd to perform specific organizational tasks" [17]. Another recent definition by Chamales also highlights the technological component of crowdsourcing [3]:

Crowdsourcing technology brings together a distributed workforce of individuals in order to collect resources, process information, or create new content. The implementation of a crowdsourcing system can vary widely, from complex online websites that coordinate a million simultaneous workers to low-tech, ad hoc approaches that use a shared spreadsheet." [3].

At present, Web 2.0 technologies have expanded the range of available crowdsourcing methods to the point that the concept has become an umbrella term that covers multiple ways to collect and share information, respond to labor offers or contests, volunteer for a number of tasks, etc. Reviewing some of the currently available tools will provide the basis for some useful distinctions.

3 Open source crowdsourcing platforms

In the last few years, crowdsourcing platforms have sprouted to leverage the resources of the crowds in crisis and disaster management efforts [14]. Most of these tools have embraced open source licenses from their inception. The first generation of open-source platforms, Ushahidi, OpenStreetMap, and Sahana are among the most popular, with large communities of developers and users. Ushahidi was initially launched as a Google Maps mash-up to map reports of violence after the Kenyan post-election fallout at the beginning of 2008.[1] Ushahidi and Crowdmap (its hosted version) have been used in over 30,000 deployments in 156 countries [9].

[1] http://www.ushahidi.com/about-us

OpenStreetMap is an editable map with more than 1,350,000 registered users (as of August 2013).[2] The platform, started in 2004 by Steve Coast, allows free access to the full map dataset via the Open Data Commons Open Database License (ODbL).

The origins of Sahana ("relief" in Sinhala) are also grounded in the response to a critical event (the Indian Ocean Tsunami of 2004) and the need to coordinate organizational efforts. The newer version of the platform (Eden) specifically addresses disaster management tasks and includes dedicated modules for organization registry, volunteer management, and online mapping.[3]

CrisisTracker, initially developed by Jakob Rogstadius in 2011, combines automated processing with crowdsourcing to quickly detect new events in Twitter. The CrisisTracker platform uses an automated real-time clustering algorithm based on Locality Sensitive Hashing (LSH) to group together tweets that are textually very similar.[4] Volunteers are then invited to refine the topical clusters or create new ones.

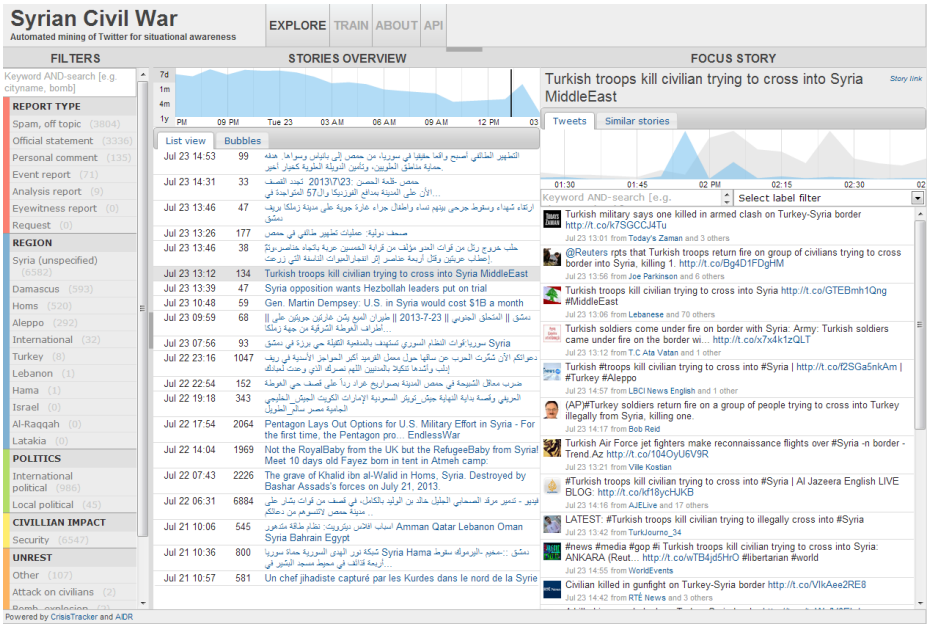

Fig. 1. Clustering tweets related to the Syrian Civil War with CrisisTracker

Deployments and projects using any of the above platforms typically require help from volunteers with different skill levels in information and data management (i.e. media monitoring, categorization, reporting, etc.), GIS (geolocation and mapping) or disaster management (logistics, volunteer management, etc.). In recent years, a number of volunteer communities from different backgrounds and domains have gathered

2 https://openstreetmap.org

3 https://eden.sahanafoundation.org/

4 https://github.com/jakobrogstadius/crisistracker

around these tools to contribute to crisis and disaster management efforts.[5] Frequently, volunteer communities provide initial training to any incoming contributor via different channels: skype chats, IRC channels, Google hangouts, tutorials, etc.

A second generation of open-source tools mainly consists of lightweight, easy-to-navigate mobile applications. In this mobile environment, the tasks (or micro-tasks) that volunteers are usually requested to complete are bite-size chunks (e.g. translate a sentence, tag a tweet, assess different levels of damage as seen in a picture, etc.). TaskMeUp is an application initially developed in 2010 by InSTEDD.org where users can request the help of volunteers on tasks such as text message translation or categorizing information.[6] Crowdcrafting, defines itself as a "framework for developing and deploying crowd-sourcing and microtasking apps".[7] Recently, two of these microtasking apps—TweetClicker and ImageClicker—have been used as part of the response to typhoon Laura in the Philippines in a partnership between UNOCHA and digital volunteer organizations (i.e. the Standby Task Force and Humanitarian Open Street Map). The tasks requested to volunteers—in an open call open to the general public—were fairly simple. TweetClicker asked to tag a tweet at a time (from a set of tweets filtered with machine learning algorithms) either as not relevant to the disaster, as a request for help, as reporting infrastructure damage, or a population displacement. Similarly, ImageClicker proposed three categories of damage (none, mild, or severe) to tag images extracted from social media. Each app included a mini-tutorial to guide volunteers, who could also participate in a skype chat if they wanted to share questions or comments. The two apps have been developed by Micromappers, a project led by Patrick Meier at the Qatar Computing Research Institute (QCRI).[8]

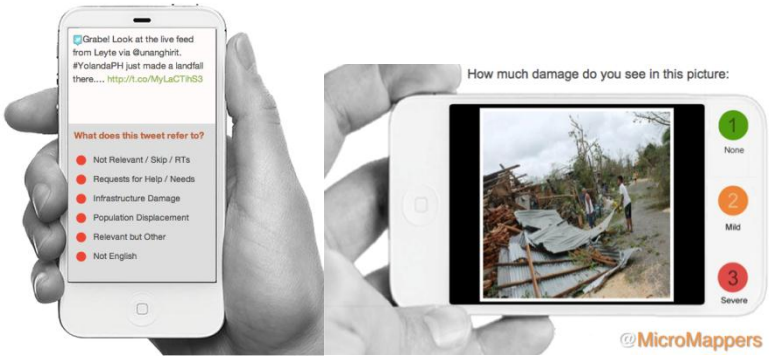

Fig. 2. TweetClicker and ImageClicker (by MicroMappers)

[5] See the Digital Humanitarian Network, a network of volunteer organizations working on disaster management efforts from different backgrounds, http://digitalhumanitarians.com/

[6] https://bitbucket.org/instedd/taskmeup

[7] http://crowdcrafting.org/about

[8] http://www.qcri.com/

4 Crowdsourcing roles

The size and composition of the crowd can also help to determinate whether the crowdsourced effort is unbounded (anyone can participate) or bounded to "a small number of trusted individuals" [13]. We can further distinguish the role of the crowd based on the type of data being processed and the level of participation involved. This leads to four types of crowdsourcing roles based on: (i) type of data processed (raw, semi-structured, and structured data), (ii) participants' level of involvement (passive or active) and, (iii) skills required to fulfill the assigned task (basic or specialized skills). Figure 3 below shows these four roles based on how the crowd is involved in the process of generating and adding value to the knowledge chain process.

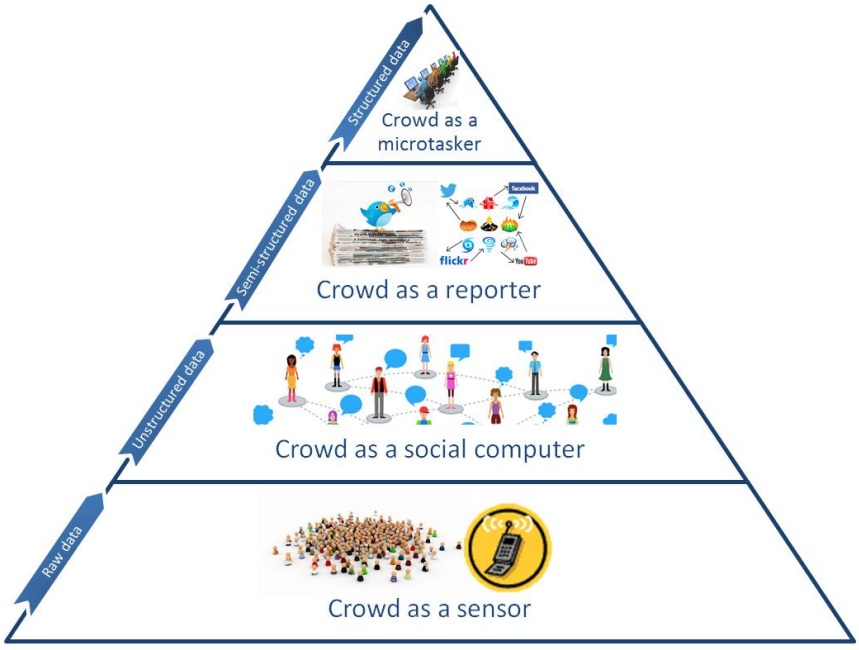

Fig. 3. Crowdsourcing roles based on users' involvement and level of data processing

The lower tiers of the pyramid represent users who generate raw or unstructured data by the mere use of mobile phones, tablets, etc. (crowd as a sensor) or their either occasional or regular use of social media (crowd as a social computer). In contrast, the two top tiers include users with an explicit, conscious use of a priori knowledge to achieve a specific goal (crowd as a reporter and crowd as a microtasker). Moving from lower to higher levels in the pyramid also implies a shift in the quality of the obtained data. From a knowledge generation and data processing point of view we are ranging from raw data, unstructured data, or semi-structured data, to structured data (which also become interpreted data resulting from the execution of the process).

Whereof, lower roles in the pyramid produce raw data and higher roles high valued data which are related with the action of solving a specific problem (e.g. labeling an image). Such a categorization also implies different levels of effort by the crowds involved:

i) Crowd as sensors: people generate raw data just because some processes are automatically performed by sensor-enabled mobile devices (e.g. processes run in the backend by GIS receivers, accelerometers, gyroscopes, magnetometers, etc.) which can be later on used for a purpose (i.e. mobile phone coordinates for positional triangulation, traffic flow estimates, etc.). This type of data collection has been defined elsewhere as "opportunistic crowdsourcing" [30]. Opportunistic crowdsourcing requires very low data processing capabilities (if any) on the side of participants and is the most passive role in the contributing information chain.

ii) Crowd as social computers: people generate unstructured data mostly by using social media platforms for their own communication purposes (e.g. sharing contents or socializing in Facebook, Twitter, Instagram, etc.). Social media users do not process information in any specific form, but these data can later be reused to extract semantically structured information. As in the previous role, there is no explicit participatory effort in any crowdsourced initiative or project).

iii) Crowd as reporters: people offer first-hand, real-time information on events as they are unfolding (e.g. they tweet about a hurricane making landfall and the reporting damages in a specific location). This user-generated content already contains valuable metadata added by users themselves (e.g. hashtags) than can be used as semi-structured, preprocessed data.

iv) Crowd as microtaskers: people generate structured, high quality, interpreted data by performing some specific tasks over raw data (e.g. labeling images, adding coordinates, tagging reports with categories, etc.). This role requires an active participation of users in the crowdsourcing effort and it may exploit special skills or require different levels of previous training.

5 Conclusion

As new tools and technologies enable citizens to participate in crowdsourced efforts with different roles and skills, new opportunities emerge for projects and initiatives involving the management of large amounts of data. Disaster management, environmental sensing, scientific research, business, and marketing are among the areas than can benefit from crowdsourced input or microtasking activities. The efficient allocation of tasks to a largely distributed online workforce can produce immediate outcomes that would be otherwise difficult to obtain with traditional outsourcing methods. Yet, bringing such a large crowd-force into organizational workflows raises a

number of issues that need to be taken into account: management of the crowd, accuracy, reliability, quality control of the outcomes, etc. Compliance with data protection and privacy rules (including different types of consent) will also help to delimitate how crowdsourced data can be aggregated, shared, used, and reused. Finally, an appropriate ethical framework can be brought into the picture to interact and complement rules, principles, and standards whenever needed [1,2,15]. Further research will explore further the connections between different crowdsourcing roles and the corresponding regulatory frameworks.

6 Acknowledgements

This research has been supported by two research grants from the Spanish Ministry of Economy and Competitiveness (MINECO) to the projects "CrowdCrissControl" (IPT-2012-0968-390000) and "Crowsourcing: instrumentos semánticos para el desarrollo de la participación y la mediación online" (DER 2012- 39492 -C02 -01).

7 References

1. Casanovas, P., Poblet, M. 2008. Concepts and Fields of Relational Justice. In: Sartor, G., Casanovas, P., Casellas, N., Rubino, R. (Eds) Computable Models of the Law: Languages, Dialogue, Games, Ontologies. LNAI 4884, Springer Verlag, Berlin, Heidelberg: 323-339.
2. Casanovas, P., Poblet, M. 2009. The Future of Law: Relational Law and next Generation of Web Services. In: M. Fernández-Barrera et al. (Eds), The Future of Law and Technology: Looking into the Future. Selected Essays. European Press Academic Publishing, Florence: 137-156.
3. Chamales, G. 2013. Towards trustworthy social media and crowdsourcing. Wilson Center Commons Lab. Available at http://www.wilsoncenter.org/sites/default/files/TowardsTrustworthySocialMedia_FINAL.pdf.
4. Chatzimilioudis, G., Konstantinidis, A., Laoudias, C., Zeinalipour-Yazti, D. 2012. Crowdsourcing with smartphones. Internet Computing, IEEE, 16(5), 36-44.
5. Doan, A., Ramakrishnan, R., and Halevy, A. Y. 2011. Crowdsourcing systems on the World-Wide Web, *Communications of the ACM*, Vol. 54 (4) 86-96.
6. Estellés-Arolas, E., González-Ladrón-de-Guevara, F. 2012. Towards an integrated crowdsourcing definition. *Journal of Information Science*, Vol. 38 (2): 189–200.
7. Geiger, D., Seedorf, S., Schulze, T., Nickerson, R. C., Schader, M. 2011. Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. AMCIS-Proceedings of the Seventeenth Americas Conference on Information Systems (Detroit, Michigan August 4th-7th). Available at http://schader.bwl.uni-mannheim.de/fileadmin/files/schader/files/publikationen/Geiger_et_al._-_2011_-_Managing_the_Crowd_Towards_a_Taxonomy_of_Crowdsourcing_Processes.pdf.
8. Haklay, M. 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In *Crowdsourcing Geographic Science*, D. Sui, S. Elwood, M. Goodchild (eds.): 105-122, Springer Netherlands.

9. Hersman, E. 2012. Which is better, 10,000 reports or 10,000 maps? (Part 1/2), http://blog.ushahidi.com/2012/07/10/which-is-better-10000-reports-or-10000-maps-part-12/
10. Hetmank, L. 2013. Components and Functions of Crowdsourcing Systems: A Systematic Literature Review. *Wirtschaftsinformatik Proceedings 2013*. Paper 4.
11. Howe, J. 2006. The rise of crowdsourcing. *Wired*, June 14 2006. Available at http://www.wired.com/wired/archive/14.06/crowds.html
12. Malone, T. W., Laubacher, R., Dellarocas, C. N. 2009. Harnessing crowds: Mapping the genome of collective intelligence. MIT Sloan Research Paper 4732-09.
13. Meier, P. 2011. Why Bounded Crowdsourcing is Important for Crisis Mapping and Beyond. Available at http://irevolution.net/2011/12/07/why-bounded-crowdsourcing/
14. Poblet, M., Garcia-Cuesta, E., Casanovas, P. 2104. Crowdsourcing Tools for Disaster Management: A Review of Platforms and Methods. In *Lecture Notes in Artificial Intelligence*, Springer Verlag, Berlin, Heidelberg (forthcoming)
15. Poblet, M.; Leshinsky, R., Zeleznikow, J. 2012. Digital neighbours : Even Good Samaritan crisis mappers need strategies for legal liability, Planning News, 38, 11, Dec 2012: 20-21.
16. Quinn, J. A., Bederson, B. B. 2011. Human Computation: A Survey and Taxonomy of a Growing Field", CHI Conference, May 7-12, 2011 Vancouver, BC, Canada: 1403–1412. Available at http://alexquinn.org/papers/Human%20Computation,%20A%20Survey%20and%20Taxonomy%20of%20a%20Growing%20Field%20(CHI%202011).pdf
17. Saxton, G.D., Onook, O., Kishore, R. 2013. Rules of Crowdsourcing: Models, Issues, and Systems of Control Information Systems Management, Vol. 30(1): 2-20.
18. Schenk, E., Guittard, C. 2011. Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics* 1/2011 (n°7). Available at www.cairn.info/revue-journal-of-innovation-economics-2011-1-page-93.htm
19. Zhao, Y., Zhu, Q. 2012. Evaluation on crowdsourcing research: Current status and future direction. *Information Systems Frontiers.* Published online: 11 April 2012.

Crowdsourcing as an Activity System: Online Platforms as Mediating Artifacts

A Conceptual Framework for the Comparative Analysis of Crowdsourcing in Emergencies

Gregory Asmolov

Media and Communications Department, London School of Economics and Political Science, Houghton St, London WC2A 2AE, United Kingdom
g.asmolov@lse.ac.uk

Abstract. This paper explores a number of challenges in the analysis of crowdsourcing platforms, relying on major theoretical approaches. In order to address these challenges, it suggests applying cultural-historical activity theory (CHAT) to the analysis of crowdsourcing projects. Accordingly, it suggests that crowdsourcing projects can be analyzed as tools that contribute to the construction of activity systems. Applying CHAT allows addressing a number of central questions, including the relationship between subjects and objects as well the dynamics of the power relationship around crowdsourcing deployments. It also allows the conducting of a comparative investigation of crowdsourcing projects, while "activity" is considered as the major level of analysis. The paper also introduces a number of methods that can be used to investigate crowdsourcing applications as a manifestation of an activity system.

1 Introduction

Since Jeff Howe coined the term "crowdsourcing" in 2006 [33], a broad and interdisciplinary body of academic literature dedicated to this concept has been constantly developing. The research reflects numerous disagreements on a variety of issues related to crowdsourcing. The major field of battle is the definition of crowdsourcing. For instance, Estellés-Arolas and González Ladrón-de-Guevara [20] counted more than 40 different definitions of crowdsourcing and suggested their own, which made a comprehensive effort to integrate the others. Recently, Brabham dedicates special attention to discussions of "what crowdsourcing is and is not – strictly speaking." [9].

There are a number of layers of disagreement in discussions around crowdsourcing. The first layer concerns the purpose of crowdsourcing applications. Some researchers approach them primarily either as a new business model for the production of material and immaterial goods or as a form of digital labor that allows increasing profit for firms [34,8,38,22,26]. Another body of literature expands the potential applications of crowdsourcing to problem-solving [8]. Depending on the disciplinary affiliation of the writer, the role of crowdsourcing is discussed in a context of crisis

response [43], the production of volunteer geographic information (VGI) [29,7], governance [6] or citizen science [30], among others. A number of papers suggest mapping the types of function and types of crowd in an application to a specific field e.g. crisis response [44].

Another layer of debate is the structure of the relationship between actors participating in crowdsourcing. For instance, Brabham [9] discusses the spectrum of control between the organization and the crowd. He defines crowdsourcing in very specific terms as a project where the purpose is defined by the organization, while the potential crowds that can be engaged have, in the process of achieving this purpose, a limited degree of freedom in their participation. According to Brabham, on the one hand, "when the locus of control is too much on the side of the organization – the crowd becomes a mere pawn." On the other hand, "the opposite end of the spectrum when the locus of the control resides more on the side of the community" leads to self-governance, while a situation where "the organization is merely incidental to the work of the crowd" is also not considered by Brabham [9] as crowdsourcing.

Some researchers, however, use as examples crowdsourcing projects where the crowd is not only responding to a request, but also defines the purpose of the deployment (e.g. Wikipedia). Some research explores the obstacles to the collaboration of institutional and informal actors around the same project, in particular in the field of scientific research [48].

The nature of the resources that are mobilized is also disputed. Some researchers suggest relying on Surowecki's concept that what is mobilized is the "wisdom of crowds" [60]. Others approach it as "crowd capital" [56]. The discussion around the nature of resources also differentiates between those that are used for simple mechanical tasks and those that can address complicated tasks [55]. A concept of thin and thick engagement [28] can be helpful in differentiating between various forms of participation by the crowd in crowdsourcing projects. Relying on analysis of the nature of resources and the nature of tasks, a number of researchers [55] suggest models for the optimization of the crowdsourcing process and for matchmaking between the crowd and those who seek to engage it for a particular purpose.

The layers mentioned above present primarily instrumental research that is focused on how crowdsourcing is used, what impact it has on different fields and how its value can be optimized. Another stream of research on crowdsourcing is critical analysis. On the one hand, some researchers who have an optimistic attitude to information technologies approach crowdsourcing as a concept that can empower people. Concepts like "participatory culture" [36], peer production [4] "long tail" [1], "cognitive surplus" [59] allow us to discuss crowdsourcing as a concept that supports generosity, creativity and the agency of individuals. Meier [49] suggests that crowdsourcing can be used for the mobilization of "global goodwill". On the other hand, some research suggests a dystopic vision of ICTs in general and of crowdsourcing in particular. Neo-marxist scholars approach crowdsourcing as another form of "immaterial labour" [39] and as exploitation of the digital labor of crowds in order to gain profit for firms. Researchers such as Fuchs and Sevignani [26], who discuss the ICTs in classical Marxist vocabulary, suggest that crowdsourcing should be freed from the control of capital and transformed into "digital work" that serves the interests of people and not

the interests of capitalist structures. As Brabham [8] points out, "It is easy for critics to bemoan the oppressive exploitation of labor taking place in the crowdsourcing process, but narratives from superstars in the crowd indicate more agency than Marxist critiques would allow."

Analysis of crowdsourcing can also rely on a number of major theoretical frameworks that are often applied to the analysis of ICTs. The notion of connective action developed by Bennett and Segerberg [5] can significantly contribute to understanding the dynamic of the process behind crowdsourcing. According to Bennett and Segerberg, unlike collective action, which relies on coordination by organizational structures and hierarchical institutions that suggest a specific frame of action, "connective action networks are typically far more individualized and technologically organized sets of processes that result in action without the requirement of collective identity framing or the levels of organizational resources required to respond effectively to opportunities" [5]. While crowdsourcing can be approached as a manifestation of connective action, the concept does not allow for a distinction to be made between crowdsourcing and any other type of ICTs that also support the loosely organized action of many individuals.

Crowdsourcing can also be approached as a manifestation of networking power [13] and analyzed in terms of programming and switching power. In this case the major level of analysis is networks and crowdsourcing platforms are tools for the formation of networks around a particular purpose that can be analyzed in terms of "programming" and "switching" [12]. The purpose of crowdsourcing platforms reflects programming power. The coalition of groups that emerges as a part of collaboration around the platform's purpose can be addressed through switching power.

According to the network power concept, ICTs do not necessarily favor horizontal actors and lead to a change in the power relationship. According to Castells, the power of networks can be used by the traditional power-holders – corporations and governments. That said, Castells also introduces the concept of mass self-communication, which suggests how networks can challenge the traditional hegemonic actors. According to Castells [11], mass self-communication is "[t]he building of autonomous communication networks to challenge the power of the globalized media industry and of government and business controlled media."

The counter-power that relies on mass self-communication and uses the "opportunity offered by new horizontal communication networks of the digital age" is defined by Castells as "the capacity by social actors to challenge and eventually change the power relations institutionalized in society" [11]. Consequently one can suggest that crowdsourcing can be approached not only as a form of programming/switching power but also, potentially, as a technology that enables new forms of mass self-communication.

However, there are a number of challenges in using Castell's theory to analyze crowdsourcing. First, it does not allow the addressing of any unique features of crowdsourcing projects. From this point of view of networks there is no substantial difference between social networks and crowdsourcing platforms. Both can be approached as manifestations of networking power. The focus on networks also does not leave space for other elements of crowdsourcing – the tools (crowdsourcing plat-

forms) and the purpose of crowdsourcing projects. Furthermore, it does not address the nature of the resources that are mobilized, while focusing primarily on the process of mobilization.

Neither the networking power concept nor collective/connective action and social mobilization theory can address the complexity of crowdsourcing or differentiate between it and other ICT-based applications. We can also see that the concepts that investigate crowdsourcing while relying on a specific notion of its purpose (be this production, problem-solving, the generation of generosity or governance) limit the scope of research in a way that can lead to missing a substantial part of the project.

The conceptualization of crowdsourcing requires a framework that will allow for the addressing of the actors and their relationships, the structure of resources, the process of mobilization of these resources and the purpose of mobilization. The concept needs to be neutral in terms of an optimistic or dystopian view of the ICTs. At the same time it should allow not only an instrumental, but also a critical analysis that explores the structure of power relationships in crowdsourcing projects.[1]

Consequently, what is necessary first is a definition of crowdsourcing that distinguishes it from other forms of ICTs used for production and/or social mobilization, but at the same time does not limit its understanding to a particular form of relationship among the actors, a particular purpose or a particular definition of the nature of the resources mobilized. It should also avoid embedding critical interpretations of the nature of crowdsourcing as a process.

2 Definition of Crowdsourcing

The conceptualization of crowdsourcing requires a definition that will distinguish it from other ICT applications and address the challenges described above. This paper would like to suggest that the unique feature of crowdsourcing is that in any crowdsourcing project there is a link between communication and action. The structural properties of the crowdsourcing tools and deployments always link communication to mobilization. In other words, the major characteristics that distinguish crowdsourcing projects from a number of other online tools, including social networks and blogs, is that crowdsourcing projects are action-oriented tools which by definition are used to mobilize and engage Internet users and a variety of potential audiences.

In the case of crowdsourcing, the symbolic power of representation and the material power of action are interrelated because of the design of the system. The type of social construction and discourse that are mediated and produced through crowdsourcing platforms have a link to the specific types of action defined by a particular platform. Consequently, a crowdsourcing platform is a framework that relies

[1] One of the other theoretical frameworks that can link people to tools and approach crowdsourcing as a form of enrollment is actor-network theory (ANT). However, this also fails to distinguish crowdsourcing platforms from other tools. Moreover, the approach of ANT to power relationships could create significant challenges for critical analysis of crowdsourcing deployment (in particular the power relationships).

on a link between communication and the mobilization of a crowd in order to carry out a specific type of action defined through the platform. The way in which the situation is framed always appears in the context of a potential action.

Accordingly, I would like to limit the comprehensive definitions and rely on a definition of crowdsourcing as the ICT-mediated mobilization of networked individuals' (the crowd's) resources in order to achieve a particular goal. This definition does not limit the nature of the purpose, but suggests that crowdsourcing is always purpose-oriented. It does not suggest a specific mode or relationship between crowd and organization, but argues that the mobilization of the crowd's resources by any type of actor is always the core of a crowdsourcing project. Finally, it emphasizes that crowdsourcing is always mediated though ICTs.

Instead of defining the potential forms of application of crowdsourcing platforms, this paper suggests mapping the potential resources that can be mobilized through the mediation of the Internet in order to achieve a particular purpose. Every crowdsourcing platform seeks to mobilize a particular set of crowd resources.[2] It can include:

- Sensor resources (mobilization of the crowd in order to collect information around a specific topic)
- Intellectual resources (knowledge and experience)
- Analytic resources (data-mining and the curating of information that does not require prior knowledge)
- Financial resources (money, also known as crowdfunding)
- Commodity resources (any type of goods or objects that have value)
- Physical resources (any type of activity that requires physical action, participation, demonstration or volunteering)

The core element of crowdsourcing is not the structure of the actors' relationship and not the purpose, but the action that is enabled by the mobilization of the resources of the crowd and mediated through ICTs. Therefore the major level of analysis should be the process - the action, and the system of resources, actors and purposes that emerge around it.

When applying this notion to Castells' argument about the capacity of horizontal networks to challenge traditional power structures, we have to distinguish crowdsourcing platforms from other forms of "mass self-communication." Consequently, in order to incorporate crowdsourcing within a discussion of power, I would suggest that crowdsourcing is a specific form of mass self-communication that should be conceptualized as "*mass self-mobilization*." That said, "mass self-mobilization" is only a private case of crowdsourcing, while the crowd can be also mobilized by external entities and organizations.

[2]Resources of any type can also be measured in terms of time – how much time needs to be spent in order to have the resources required for specific tasks or for completing a specific task. This means that the value of similar resources has a relative nature and can be different for different people.

The purpose of this paper is to suggest a conceptual framework that addresses activity as the major level of analysis of crowdsourcing platforms. This framework allows investigation of a linking of the technology to the action through the notion of *mediated activity*. However, at the same time it approaches crowdsourcing not as a technological platform, but as a new social phenomenon enabled by ICTs. In order to achieve the purpose of this methodological project and address the challenges described above, I would suggest using cultural-historical activity theory (CHAT). The following sections will describe the principles of CHAT and elaborate on how it can be applied to the analysis of crowdsourcing projects.

3 Cultural-Historical Activity Theory and the Analysis of Crowdsourcing Projects

3.1 Activity Theory and Mediation

The foundation of activity theory relies on Karl Marx's Theses on Feuerbach [46]. Vygotsky and his colleagues Luria and Leontiev "used Marx's political theory regarding collective exchanges and material production to examine the organism and the environment as a single unit of analysis" [62]. According to this concept, there is no direct linkage between an individual and his/her environment. The relationship between a subject (individual) and an object is always mediated. The mediation is conducted through variety of cultural means, including tools and signs.

The role of the tools in a relationship between the subject (the person) and his environment (object), as developed by Marx, became a basis for the concept of artifact-mediated and object-oriented action [61]. While it is out of the scope of this paper to trace the origins of CHAT, it is worthwhile to highlight this central element: the notion of mediation. In this perspective, "all human experience is shaped by the tools and sign systems we use" [52]. As Engeström has also put it, "subject's actions are mediated through tools/instruments and directed at a particular object" [18,19]. According to Kaptelinin and Nardi [37] "the structure of a tool itself, as well as learning how to use a tool, changes the structure of human interaction with the world. By appropriating a tool, integrating it into activities, human beings also appropriate the experience accumulated in the culture."

While Vygotsky was the first to introduce the notion of mediated activity, his interests were focused primarily on the individual level and the development of human consciousness through the mediated interrelation of subjects and their environments. Leontiev [41] treated activity as a holistic unit of analysis that not only could be applied to individuals, but also "broadened the scope of Vygotsky's mediated action by introducing human activity as the unit of analysis that is distributed among multiple individuals and objects in the environment" [62].

Relying on the latter notion Engeström [16] developed an analytical framework for the analysis of activity systems, while defining his primary level of analysis as a "joint activity or practice" and the activity system as a "systems of collaborative human practice" [17]. Engeström's model identifies a number of new components that

had not been conceptualized previously, including rules, community and division of labour. The elements at the top of the triangle remain the same: subjects, mediating artifacts and objects.

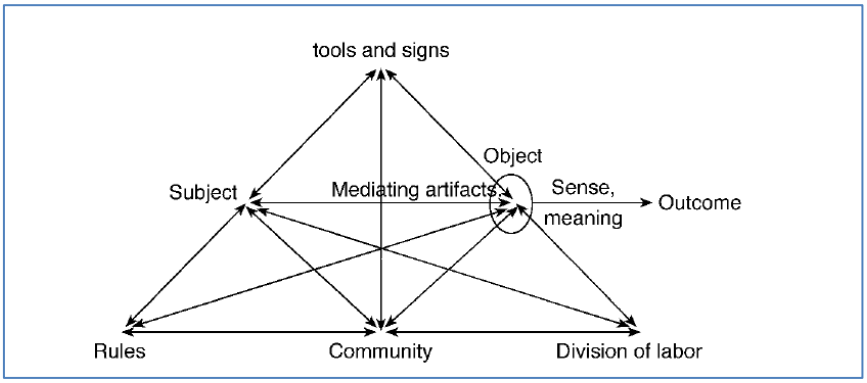

Fig. 1. The structure of a human activity system, by [16]

The notion of community allows for the argument that a division of labor takes place in a particular community that collaborates around a particular object. At the same time, an activity system is also based on a set of rules and norms that are shared by its members.

3.2 Activity and Natural Disasters

The traditional field of application of activity theory is education and concepts of child development. CHAT is often used to analyze organizations. It has started often to be used for the analysis of human-computer interaction [37]. Here, however, I would like to illustrate the advantages of the application of CHAT to an analysis of crowdsourcing based on addressing emergency situations.

Leontiev [40,41,42] writes that when we talk about the social environment of a person we mean the environment that the person is adapted to living in. The closest environment is the social group that a person belongs to and his circle of communication. Leontiev, however, emphasizes that adaption to the surrounding environment is not the core of personal development. On the contrary, at the core of development is the capacity associated with opportunities to go outside the comfort zones of the surrounding environment. Leontiev argues that the development of new activity systems is "caused by dialectical contradictions between organisms and their environments" [37].

In some cases a person can go out of his/ her own comfort zone. However, in other cases, his/her comfort zones are destroyed when a person has not changed their environment. This happens in crisis situations, and in particular natural disasters. Accordingly, a natural disaster is not only a tragedy, but also an opportunity for development. A disaster suggests a new form of relationship between a person/ collective and

nature, and this relationship – emergency response – is mediated through a variety of tools. Consequently, a natural disaster leads to the definition of new objects of activity and the transformation of everyday life activity systems.

There are two layers of analysis relying on CHAT that can be applied to emergency situations. The first layer is that of analyzing professional emergency response organizations, which present institutionalized forms of activity systems created in order to respond to emergencies. The second layer addresses the general population, including affected communities and potential volunteers who are not affiliated with formal emergency response institutions.

For instance, activity theory was used for an analysis of NASA's response to the Challenger disaster [32]. Owen [53] uses activity theory for an investigation of the emergency response to bushfires in Tasmania. A group of researchers used activity theory in order to investigate the emergency response to attacks in Mumbai [57].

Mishra and others [50] provide a case study of using activity theory as a conceptual and methodological framework for the analysis of organization-based emergency response. Their paper investigates the contradictions and tensions in an emergency response system as a potential trigger for innovation. It relies on a number of methods, including training observation and semi-structured interviews with tactical commanders in the UK Police, Fire and Rescue Services and Ambulance Service.

Mapping the activity system of emergency response allowed the examining of "the role of tools within the activity system and the way in which they mediate behavior" [50]. The research provides an example of how the question of mapping an emergency response activity system can be formulated by relying on a triangle of the activity system (including rules, community, division of labor). The analysis suggests contradictions between emergency response officials and the technologies they use, as well as contradictions between subject and rules, subject and community, and subject and division of labor. The paper concludes that Activity Theory is "a valuable methodological and analytical tool" for the investigation of emergency response. It also suggests that we can rely on the analysis that "tensions and contradictions are considered as a source of innovation" [50].

As we can see, most of the applications of CHAT to emergency response analysis are focused on institutional structures and not on the general population. However, according to Leontiev [42], we would expect a natural disaster to be a moment of transformation in particular for those who do not expect it. CHAT allows the conceptualizing of the relationship between nature and people, in a context of collective activity.

In addition to the Engeström model, which can be used for mapping an emergency response system (including the community of responders, the division of labor between responders and the tools that mediate the response), the framework also allows us to focus on tensions and contradictions within an activity system that emerges in a case of disaster. At the same time, CHAT allows us to approach emergency response as a situation of development for a society whose members are forced to find themselves outside their comfort zone.

3.3 CHAT as a methodological framework

As Nardi has pointed out [52], "[a]ctivity theory is a powerful and clarifying descriptive tool rather than a strongly predictive theory." According to Yamagata-Lynch (2010), many studies use CHAT as a descriptive tool for mapping activity systems as a part of qualitative research without relating to its conceptual implications. In other words you do not have to be a CHAT scholar in order to apply CHAT. Therefore it is possible to separate CHAT as a theory from activity systems analysis as a methodology [62].

As a methodology for mapping systems, CHAT allows the identification of what Engeström suggested were bounded systems of activity. As Yamagata-Lynch points out, "activity theory researchers and practitioners need to examine interactions shared among multiple activities and the boundaries of those activities to identify the potential development and changes in both human activity and societal systems" [62]. CHAT provides the methodological framework that allows us to draw the boundaries of a system for the purpose of an analysis.

4 Crowdsourcing as Mediation of Activity

This paper argues that ICT, and in particular crowdsourcing platforms, can give rise to different types of new activity systems. In different socio-political environments we can expect the emergence of different types of activity system. This difference is mediated through tools (crowdsourcing platforms).

Additionally the paper argues that, as a methodology, CHAT can provide a framework for analyzing crowdsourcing platforms and responding to a number of central questions about the structure of power relationships and the association between crowdsourcing deployments and their socio-political environment. It suggests that "activity" can be identified as the major level of analysis as a part of the investigation of crowdsourcing platforms.

Accordingly, crowdsourcing platforms can be conceptualized as mediating artifacts of activity systems that suggest a particular structure of potential action. In other words, crowdsourcing platforms can be approached as a mode of governance and a technique of power [23,24,25]. Relying on CHAT, the purpose of this investigation is the deconstruction of crowdsourcing platforms as platforms that suggest a particular range of actions and define a particular type of activity system.

According to Engeström, various activity systems inherit various types of tension between the nodes, and as a consequence we can expect to identify various types of tension in various environments that lead to the emergence of various types of activity system as they are mediated through crowdsourcing platforms. The triangle also allows us to analyze the internal contradictions and conflicts within an activity system where the "'nodes' pull and push against one another" [47]. These tensions can be considered as a process of constant mediation and renegotiation of the boundaries of activity systems, while the dynamic of tensions can be followed through crowdsourcing platforms.

In other words, contradictions are the driving force of change and development. However, once users are able to participate in the development of crowdsourcing systems from within, it may allow the users to resolve the contradictions without a need to create a new activity system. The way a contradiction is resolved can suggest who is dominating in a particular activity system, whether it is institutions (structure) or individuals (agency). Contradictions can also lead either to the polarization of different activity systems or to the integration of citizen and state in joint activity systems.

CHAT methodology and terminology allows us to ask and address the following questions:

- What are the boundaries of an activity system and their purpose? What is the degree of flexibility/ generativity within the system?
- What is the structure of community/ division of labor and what are the rules in a particular activity system?
- Who plays a dominant role in the definition/ mediation of boundaries of the activity system and the purpose of this system? Is it a structure-driven or an agency-driven system?
- What are the major tensions within the activity system, how does it develop and what is its proximal zone of development?
- Are there any competing activity systems around the same objects (e.g. natural disasters)?
- How can the same technologies give a rise to different activity systems in different cultural-historical/ socio-political environments?

Mediation of the division of labor in a particular activity system is one of the most important roles of crowdsourcing platforms. This refers to the way the labor is divided in crowdsourcing platforms (e.g. what can be done by skilled and unskilled volunteers, how the division of functions is defined and who defines the framework for division).

The structure of community itself (whether it is an open or a bounded community, who is excluded and included, what the criteria are for becoming part of the community) is particularly important for the mapping of crowdsourcing platforms as activity systems. These elements are embedded in the structure of the mediating tools (platforms).

5 The Power and Construction of Activity Systems

Cultural-historical activity theory does not discuss the role of power or the power relationship. At the same time, it emphasizes the linkage between the cultural-historical context and the nature of the activity system, which provides a window of opportunity for an investigation of the association between a particular type of environment and the particular structure of an activity system.

One can suggest that, since the roots of activity theory grew from the concept developed by Karl Marx, the Marxian approach to power relationships should be ap-

plied as a part of activity theory. This argument would be misleading since one have to differentiate the way Marx understood the nature of the social world through the dialectical relationship of subject and object from the particular situation described by Marx as capitalism.

Indeed, the production that takes place in capitalist society can be described in terms of activity theory. One can argue that capitalist powers use ICTs in order to construct activity systems and in order to control them, define the object of activity and gain from what is produced. For instance, Mosco [51] defines outsourcing as a "multifaceted phenomenon, one vector in an increasingly complex international division of labor involving far more than simply the transfer of service jobs from high to low wage nations." In other words, in case of outsourcing, ICTs enable the creation of global activity systems that connect and divide labor between developed and developing countries. One of the examples of outsourcing as a construction of global activity systems is call centers, which were conceptualized by Brophy [10] as a form of communicative capitalism.

However, a global activity system does not necessarily have to be constructed as a form of capitalist abuse. Activity theory does not necessarily make an argument about the exploitation of labor and alienation of a subject from the means/fruits of production. Indeed, social media and crowdsourcing platforms can be used to construct activity systems that serve the interests of large firms and support exploitation, but that does not necessarily mean that this happens in every case. Moreover, some neo-Marxist scholars [27] tend to see a form of exploitation in any online platform and in any activity system, while ignoring the fact that the same tools can serve a variety of interests and favor variety of actors, and that in some cases there is no opposition between the interests of users and the interests of the owners of a particular platform. Besides, not every effort to gain profit from users is a form of exploitation.

Information technologies, and in particular crowdsourcing platforms, can be constructed in different ways and conceptualized as tools that mediate activity and allow the emergence of new activity/ development of existing activity systems. A crowdsourcing platform is an example of a mediating artifact. Consequently, various online platforms enable the creation of various types of mediated activity system.

The core thesis concerning power relationships that can be argued on the basis of activity theory is that the structure of activity systems can favor the interests of particular actors, and that activity systems can be constructed in different ways to serve the interests of different types of actor. Consequently, one can argue that the structure of specific activity systems can embed a particular structure of power relationships.

Relying on the triangle of activity systems, we can ask a number of questions - what the object of the activity system is, how the labor is divided, who is excluded from and included in the community, what kind of rules exist within the systems. The responses to these questions are reflected in the structure of tools that serve as mediating artifacts for the activity systems. Consequently, an analysis of the structure of mediation tools can allow the deconstruction of a particular mode of power relationship.

Since the tools are developed and created by someone, we can argue that activity systems are also the object of construction and therefore can serve the interests of

different actors. One can differentiate between *agency-driven activity systems* created from the bottom up within horizontal networks and *system-driven activity systems* created from the top down within hierarchical structures. There are also options that can be seen as situated between these two extremes.

A neo-Marxist perception of power can explore only one side of the relationship between actors. There is, however, another notion of power that allows us to address the complexity of the power relationship while focusing on activity as a primary unit of analysis.

In his late works concerning governance as a disciplinary mode of power, Michel Foucault argues that the main subject of a power relationship is possibilities of action by other people: "To govern, in this sense, is to structure the possible field of action of others" [25]

A concept of *power as governance* was introduced by Foucault in his last works and in particular in *The Subject and Power* [25] where he suggests that government is a designation of the way "in which the conduct of individuals or of groups might be directed**"**. This can suggest various groups and topics for government e.g. the government of children, of souls, of communities, of families, of the sick." Foucault approaches the move from a variety of possibilities of action to one singular outcome as governance [21].

Activity is the major object of disciplinary regulation, when activity systems can be imposed and enforced from the top by institutions. At the same time, however, the notion of governance allows us to identify a field of opportunities where agency can flourish through new activity systems. Constructing activity systems can be approached as a "technique of power" [23].

One of the major advantages of Foucault's approach is that the power relationships are not fixed or stable, but ongoing through permanent change and struggle. As a part of the decentralized nature of power, Foucault denies the notion of power that comes from a particular center. According to Foucault, power is range of effects "that run through the social body as a whole". Power is inherited in and reproduced through every action [24].

Activity as a level of analysis allows the conceptualizing of the nature of this struggle, which takes place within activity systems as well as around the construction and introduction of new activity systems through new forms of activity mediation, and the definition of the boundaries of activity systems.

This notion allows crowdsourcing platforms to be approached as a field of struggle. As a technology that enables us to construct new activity systems, we may expect that institutional actors will try to use it as a new technique of power and disciplinary framework for activity, while the horizontal or bottom-up actors will try to use this opportunity to construct independent activity systems in order to allow what was conceptualized above as mass self-mobilization. This is why limiting crowdsourcing to a relationship between institutions and the crowd [9] can miss the analysis of the role of crowdsourcing in a reconsideration of power relationships, where the crowd participates not only in activity itself, but also in the definition of the activity framework.

From this perspective ICT, and in particular crowdsourcing, and the architecture of online platforms can be conceptualized as forms of "governance of crowds" that

through their structure suggest "the possible field of action of others." Every platform may have embedded a different "possible field of actions." The purpose of analysis in this case is to deconstruct the possible field of actions and the "possibilities of action of other people" that are embedded in a particular architecture and suggested to the crowd.

6 The Generativity of Activity and "Vertical crowdsourcing"

In order to explore the potential of crowdsourcing as a form of activity mediation to challenge the existent power structure, it is useful to apply the notion of generativity. Zittrain [63] defines generativity as "a system's capacity to produce unanticipated change through unfiltered contributions from broad and varied audiences." Applying this notion to crowdsourcing allows us to ask to what extent crowdsourcing systems are capable of allowing activity that will produce unanticipated change. Consequently, we can formulate a number of potential paths for the impact of the agency on crowdsourcing as activity systems:

1. The system of activity and its boundaries are constructed by individual/ horizontal agents.
2. The system of activity has been changed from within in order to allow new forms of activity.
3. The boundaries of the system are flexible enough to allow various forms of activity, including unanticipated outcomes. The degree of flexibility of the activity system can also be conceptualized as the degree of generativity – the potential capacity of the system to produce unanticipated outcomes.

An example of an issue that can be examined in order to evaluate the degree of generativity of activity systems, and the power relationship within this, is the structure of categories in crowdsourcing platforms (e.g. Ushahidi). The structure of categories suggests a particular form of activity if this is a gathering of particular types of data or/and a facilitation of particular types of offline action. In this case the question is who defines the categories, to what extent they are flexible and diverse, and who is able to change these definitions.

In fact, the categories define the boundaries of the activity system. This can be conceptualized as the taxonomy of an activity system. But once the users are able to participate in the definition of categories, or once the creator of the platform is not an institutional actor, or the structure of categories is vague enough that the lack of clarity allows a diversity of activity forms, we can argue that crowdsourcing allows us to move from a taxonomy to a folksonomy of activity, where the structure of activity systems is defined by those who participate in these systems (by the community, in terms of Engeström's model).

By contrast, we can introduce a model of "*vertical crowdsourcing*" where the structure of activity is defined by the institutional actor, without any space for the

influence of agency on the system's structure. In this case the purpose of the system, the boundaries, the structure of categories, the rules, the right to participate in community and the division of labor are dictated by the agent that created the platform. In many cases the major purpose of this type of activity system is not to produce the expected outcome, but primarily to control the activity of the crowd and neutralize the potential for independent forms of activity. This is also the situation where we can expect alienation between the community, the tools and the purpose of activity – as introduced by Marx.

Since CHAT links the structure of mediation to a particular cultural-historical environment, activity theory allows us to investigate the association between the role of ICTs in the mediation of activity and the socio-political/ cultural environment. Accordingly, it allows us to argue that the same type of technology can lead to the emergence of different forms of mediation in different socio-political environments. At the same time, however, while talking about the cultural-historical context, CHAT does not address specifically the political context of mediation. Therefore there is a space for addressing not only the cultural and historical dimension, but also the political contexts, in terms of the development of activity systems. This will allow us to focus, in the comparative analysis of crowdsourcing applications as activity systems, not only on the structure of these systems, but also on the identity of those who construct these systems and the dynamics of the power relationship around the development of a system.

7 Crowdsourcing-based Emergency Response as an Activity System

According to Leontiev, crisis can be approached as a change in the "comfort zone" of the surrounding environment that forces us to adapt to a new situation [41]. A disaster is a situation where, in order to respond, the development of new forms of activity is required. Thus emergency situations are a particularly suitable case for an analysis of the development of activity systems and an examination of the role of crowdsourcing platforms in the mediation of activity in particular and how ICTs can give a rise to new activity systems in general.

Emergency response is a system of activity where people (subjects) use tools towards objects (nature) in order to struggle against a disaster. A crisis response crowdsourcing platform cannot be analyzed by itself, but only in the context of an activity system that is mediated through the platform. While the crowdsourcing platform belongs to a tool that mediates response to disaster and negotiates the range of actions that can be applied to an object by the subject, one should question how this tool can be associated with rules (and norms), community and a division of labor that regulates the structure of collective action as part of the response.

In other words, we need to ask who the responders are (e.g. full-time workers or volunteers, professional or unskilled responders, local community or national/ global population), what functions they fulfill (e.g. mapping, coordination, humanitarian response, firefighting) and how these functions are divided between the members of

the community. Moreover the division of labour can take place between professional and unskilled responders (and then we can expect integration of organization-based and citizen-based resources into one activity system) or we might see that emergency response organizations and citizens fail to collaborate, and create separate activity systems and respond independently to the emergency. This separation should be reflected in the structure of crowdsourcing projects. Additionally, it is important to distinguish between two layers of activity: responding to the problem as a form of activity (e.g. providing food) and coordination of activity as a form of activity (e.g. allocation of resources between different needs).

Barton [3] suggests that, in a case of disaster, the everyday social system is replaced by an emergency social system. Relying on the notion of an activity system, one could suggest that we should focus on a shift between "everyday life activity system" and "emergency activity system". The major question that should be asked is whether the emergency activity system introduces new types of norm, new forms of community or a different division of labor, and if the form of this change can be associated with the role of the system/ state. For instance, activity in everyday life can be more regulated, with a clear division of labor, while an activity system in response to emergency has a different structure of rules, communities and in particular division of labor.

To conclude, the analysis of crowdsourcing and how it mediates action can help to understand the entire activity system of disaster response, and conversely the nature of an activity system is embedded within the structure of a crowdsourcing platform.

8 Methods for Mapping Activity Systems

8.1 Online Mapping

The online mapping of activity systems is focused on an analysis of online platforms as mediating tools of activity systems. There are two layers of analysis of online platforms: content and structure. For instance, content analysis of the messages on a crowdsourcing platform can allow us to identify the major types of activity mediated through this platform [1]. The structural analysis focuses on the design and various properties of the platform, e.g. categories, protocols of mobilization of community, the criteria for joining and membership (open or closed), the structure of moderation and the criteria for activation/ mobilization of the community defined by the platform. The structure (e.g. division into teams) can also teach us about the division of labor.

An additional method is joining/observing the online teams of users of crowdsourcing systems and analyzing their activity by relying on virtual ethnography methods.

8.2 Offline Mapping

The purpose of offline mapping is to investigate the role of ICTs, and in particular crowdsourcing tools, as mediating artifacts for activity systems. The purpose in this case is to look at the offline dimension of activity and to see to what the contribution

of crowdsourcing platforms was and to what extent it was significant. This type of research can be conducted by relying on ethnographic observation (e.g. participatory observation through joining responders in emergency or coordination centers for emergency response) or interviews with developers of platforms, volunteers/ crowdsourcing platform users and members of relevant organizations (e.g. emergency agencies in the case of analysis of emergency response).

9 Conclusion

This paper has suggested that cultural-historical activity theory and focusing on activity as the major level of analysis can significantly contribute to an analysis of crowdsourcing projects. The application of a framework is able to address a number of conceptual challenges that were identified by using other theories as part of the investigation of crowdsourcing. It has also been suggested that ICTs in general and crowdsourcing platforms in particular can be approached as tools that mediate activity and contribute to the construction of activity systems.

CHAT can assist us in conceptualizing the relationship between subject and object, as well as in analyzing power relationships around crowdsourcing platforms. It also enables us to investigate the association between crowdsourcing and the socio-political environment, which makes it possible to conduct a comparative analysis of a crowdsourcing project that addresses the same issues in different cultural and political systems.

10 References:

1. Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*. New York: Hyperion.
2. Asmolov, G. Russian wildfires, 2010: Challenging power relationships in emergency situations through crowdsourcing platforms. . Paper presented at IAMCR Conference, 25-29, June, 2013 http://www.iamcr2013dublin.org/content/russian-wildfires-2010-challenging-power-relationship-through-crowdsourcing-platforms-emerge
3. Barton, A. H. (1970). Communities in disasters: A sociological analysis of collective stress situations. Garden City, NY: Doubleday, Anchor Books.
4. Benkler, Y., & Nissenbaum, H. (2006). Common-based peer production and virtue. Journal of Political Philosophy, 14(4), 394–419.
5. Bennett W. L., & Segerberg, A. (2012). The logic of connective action. Information, Communication & Society, 15(5), 739-768.
6. Bott, M., Gigler, B.-S., & Young, G. (2011). The role of crowdsourcing for better governance in fragile state contexts. Washington DC: World Bank Publications. http://www.scribd.com/doc/75642401/The-Role-of-Crowdsourcing-for-Better-Governance-in-Fragile-State-Contexts
7. Boulos, M. N. K., Resch, B., Crowley, D. N., Breslin, J. G., Sohn, G., Burtner, R., … Slayer Chuang, K-Y. (2011). Volunteered Geographic Information Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and

crisis management: trends, OGC standards and application examples. International Journal of Health Geographics, 10:67, http://www.ij-healthgeographics.com/content/10/1/67
8. Brabham D. C. (2008). Crowdsourcing as a model for problem solving. An introduction and cases. Convergence, 14(1), 75-90
9. Brabham D. C. (2013) Crowdsourcing. Boston, MA: MIT Press.
10. Brophy E. (2010).The subterranean stream: Communicative capitalism and call centre labour. Ephemera, 10(3/4), 470-483. http://www.ephemerajournal.org/sites/default/files/10-3brophy.pdf
11. Castells, M. (2007). Communication, power and counter-power in the network society. International Journal of Communication, 1, 238-266.
12. Castells, M. (2011). A network theory of power. International Journal of Communication, 5, 773–787.
13. Castells, M. (2009). Communication power. Oxford: Oxford University Press.
14. Cole, M. (1985). The zone of proximal development: Where culture and cognition create each other. In J. V. Wertsch (ed.), Culture, communication and cognition (pp. 146-161). Cambridge, MA: Cambridge University Press.
15. Coté M., & Pybus, J. (2011). Learning to immaterial labour 2.0: Facebook and social networks. In M. A. Peters . & E. Bulut, E. (eds.), Cognitive capitalism, education and digital labor. New York: Peter Lang.
16. Engeström, Y. (1987). Learning by expanding: An activity-theoretical approach to developmental research. Helsinki: Orienta-Konsultit.
17. Engeström, Y. (1988) How to do research on activity? Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 10, 30-31.
18. Engeström Y. (1999) Communication, discourse and activity. The Communication Review, 3(1-2), 165-185.
19. Engeström Y. (2001) Expansive learning at work: Toward an activity theoretical reconceptualization. Journal of Education and Work, 14(1), 133-156.
20. Estellés-Arolas, E. &, González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. Journal of Information Science, 38(2), 189–200.
21. Felluga, D. (2011) Modules on Foucault: On power. Introductory guide to critical theory. http://www.purdue.edu/guidetotheory/newhistoricism/modules/foucaultpower.html. [retrieved 9 December 2013].
22. Fish A., & Srinivasan, R. (2012). Digital labor is the new killer app. New Media & Society, 14(1), 137-152.
23. Foucault, M. (1975 [1995]). Discipline & punish: The birth of the prison (A. Sheridan, trans.). New York: Vintage Books.
24. Foucault, M. (1978). The History of Sexuality, Volume 1: An Introduction, New York: Random House.
25. Foucault, M. (1982). The subject and power. In H. Dreyfus & P. Rabinow (eds.), Michel Foucault: Beyond structuralism and hermeneutics (pp. 208-226). Chicago: University of Chicago Press.
26. Fuchs C., & Sevignani, S. (2013). What is digital labour? What is digital work? What's their difference? And why do these questions matter for understanding social media? TripleC, 11(2). http://www.triple-c.at/index.php/tripleC/article/view/461
27. Fuchs, C. (2010). Labour in informational capitalism. The Information Society, 26(3). 176-196.
28. Graeff, E. (2013). Ethan Zuckerman's DML keynote talk "Beyond the crisis in civics". [blog post]. http://civic.mit.edu/blog/erhardt/ethan-zuckermans-dml-keynote-beyond-"the-crisis-in-civics". [retrieved 9 December 2013].

29. Haklay, M. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap Ordnance Survey Datasets. Environment and Planning B: Environment and Design 37(4):682–703.
30. Haklay, M. (2011). Classification of Citizen Science Activities. http://povesham.
31. wordpress.com/2011/07/20/classification-of-citizen-science-activities/[retrieved 9 December 2013].
32. Holt, G. R., & Morris, A. W. (1993). Activity theory and the analysis of organizations. Human Organization, 52(1), 97-109.
33. Howe, J. (2006). Crowdsourcing: A definition. Crowdsourcing.com, [online, June], http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html [accessed: December 9, 2013].
34. Howe, J. (2008). Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business. New York: Crown Business.
35. Il'enkov, E. V. (1977). Dialectical logic: Essays in its history and theory. Moscow: Progress.
36. Jenkins H. (2006). Convergence culture: Where old and new media collide. New York: NYU Press.
37. Kaptelinin, V., & Nardi, B. A. (2006). Acting with technology: Activity theory and interaction design. Cambridge, MA: MIT Press.
38. Kleemann, F. G., Voss, G., & Rieder, K. (2008). Un(der)paid innovators: The commercial utilization of consumer work through crowdsourcing. Science, Technology & Innovation Studies, 4(1), 5-26.
39. Lazzarato, M. (1996). Immaterial labour (P. Colilli & E. Emory, Trans.). In P. Virno & M. Hardt (eds.), Radical thought in Italy (pp. 132-146). Minneapolis: University of Minnesota Press.
40. Leontiev, A. N. (1981). Problems of the development of the mind. Moscow: Progress Publishers.
41. Leontiev, A. N. (1978). Activity, consciousness, and personality. Englewood Cliffs, NJ: Prenice Hall.
42. Leontiev, A. N. (1981). The problem of activity in psychology. In J. Wertsch (ed.), The concept of activity in Soviet psychology (pp. 37 - 71). Armonk, NY: Sharpe.
43. Liu S. B., & Palen, L. (2010). The new cartographers: Crisis map mashups and the emergence of neogeographic practice. Cartography and Geographic Information Science, 37(1), 69-90.
44. Liu, S. B., Poore B., & Earle, P. (In press). Geospatial crowdsourcing: A typology of crowdsourcing for the emergency management domain.
45. Luria, A. (1976). Cognitive development: Its cultural and social foundations. Cambridge, MA: Harvard University Press.
46. Marx, K. (1969). Theses on Feuerbach. In Marx/Engels Selected Works, Volume One (pp. 13 – 15). Moscow: Progress Publishers. . http://www.marxists.org/archive/marx/works/1845/theses/theses.htm
47. Massanari, A. L. (2007). In context: Information architects, politics, and interdisciplinarity. PhD dissertation, Department of Communication, University of Washington.
48. Mansell, R. (2013). Employing digital crowdsourced information resources: Managing the emerging information commons. International Journal of the Commons, 7(2), 255-277.
49. Meier, P. (2013). Opening keynote address at Crisis Mappers 2013, November 20th, 2013, http://irevolution.net/2013/11/20/opening-keynote-crisismappers-2013/ [retrieved 9 December 2013].

50. Mishra J. L., Allen, D., & Pearman, A. (2011). Activity theory as a methodological and analytical framework for information practices in emergency management. Proceedings of the 8th International ISCRAM Conference, Lisbon, Portugal, May 2011.
51. Mosco V. (2005). Here today, outsourced tomorrow: Knowledge workers in the global economy. Javnost - The Public, 12(2), 39-56.
52. Nardi, B. A. (1996). Activity theory and human computer interaction. In B. A. Nardi (ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 1-8). Cambridge, MA: MIT Press.
53. Owen C. A. (2007). Analysing the activity of work in emergency incident management. Bushfire Co-operative Research Centre University of Tasmania, @ctivités, 4(1). http://www.activites.org/v4n1/owen-EN.pdf
54. Palen, L., &Liu, S. B. (2007). Citizen communications in crisis: anticipating a future of ICT supported public participation. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, pp. 727-736.
55. Petrick I., Erickson, L. B., & Trauth, E. M. (2012). Hanging with the right crowd: Matching crowdsourcing need to crowd characteristics. Proceedings of the Eighteenth Americas Conference on Information Systems, Seattle, Washington, August 9-12.
56. Prpić J. , & Shukla, P. (2012). The theory of crowd Capital. Proceedings of the Hawaii International Conference on Systems Sciences #46. January 2013, Maui, Hawaii, USA. IEEE Computer Society Press, Forthcoming.
57. Shankar, D., Agrawal, M., & Raghav Rao, H. (2010). Emergency response to Mumbai terror attacks: An activity theory analysis. In Cyber security, cybercrime and cyber forensics: Applications and perspectives.
58. http://www.igi-global.com/chapter/emergency-response-mumbai-terror-attacks/50713 [retrieved 9 December 2013].
59. Shirky, C. (2010). Cognitive surplus: Creativity and generosity in a connected age. New York: Penguin.
60. Surowiecki, J. (2005). *The wisdom of crowds*. Anchor Books.
61. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.
62. Yamagata-Lynch, L. C. (2010). Activity systems analysis methods: Understanding complex learning environments. Springer.
63. Zittrain, J. (2009). The future of the Internet--And how to stop it. Penguin.

The Future of Democratic Participation:
my.con: An Online Constitution Making Platform

Sean Deely and Tarik Nesh Nash

Postwar Reconstruction and Development Unit, Dept. of Politics, University of York, York, England
University of Mundiapolis, Aéropole Mohammed V, Nouacer, Morocco
sean.deely@gmail.com, tarikn@gmail.com

Abstract. During the Arab Spring revolutions protestors used mobile communications technology and social media platforms to share information, mobilize supporters, and organize activities to bring about the political transformation of their countries. In each case the drafting of a new constitution was the next step adopted to continue that transformation. We ask whether the digital revolution that powered the overthrow of old regimes during the Arab Spring can also be used to facilitate a similar level of participation in the constitution making process and we present "my.con", an online platform allowing citizens to collaborate in constitutional drafting.

Keywords: participation, constitution making, social media platforms

1 Introduction

Recent attempts to promote participatory constitution making in post-authoritarian Arab states have struggled to overcome the limitations imposed by citizens' lack of familiarity with substantive issues, or the lack of experience of incorporating citizens' views into deliberations and drafting. Increasingly however these limitations are being overcome by advances in digital communications technology. Today young people get their news from social media and Internet platforms and they are more likely to get political and social information and advice from blogs or microblogs. Globally, Arab youth are the most likely to exchange political views on line. This paper addresses post-conflict constitution making as a key stage in the transition to democracy in the context of the Arab Spring revolutions. It analyses the findings from studies of constitution making practice in post-conflict states and identifies a number of prerequisites for meaningful participation. After reviewing recent developments in mobile communications technology and social media use in the Arab region, the paper presents "my.con" an online platform allowing citizens to participate in different stages of the process and collaborate in constitution drafting.

2 Defining Participation

If meaningful participation is the key to legitimacy in developing and sustaining democratic political systems, what constitutes meaningful participation? For the vast majority of citizens participation was traditionally limited to voting – either to elect a constituent assembly at the beginning of the process, or to endorse or reject the constitution when drafting was completed [20]. Since the end of the Cold War, efforts have been made to broaden participation in the constitution-making process in recognition of the growing importance of popular engagement as a basic right and a source of legitimacy. While there have been some outstanding examples of successful processes, a review of the literature suggests that progress has been uneven. This seems to be partly as a result of the daunting logistical challenges, partly a consequence of poor planning, and sometimes due to a lack of conceptual clarity. A number of prerequisites for meaningful participation are identified.

2.1 Forms of Participation

Constitution making is participatory if it incorporates opportunities for the broader public to engage in the process through some combination of oversight, direct input, and ratification [13]. Traditionally the most common forms of participation have involved voting: either at the beginning of the process when citizens might have the opportunity to elect representatives of a deliberative body to draft the constitution, or in a referendum to endorse or reject the draft constitution produced by an assembly or commission [15,14,31,6,9,20].

2.2 Electing Representatives to Prepare the New Constitution

The members of the entity responsible for preparing the new constitution can either be appointed to a commission or elected to a representative assembly [15]. In the case of a commission membership is usually based on technical expertise but may also reflect political affiliation or social diversity. The assembly is a democratic and representative body, and is usually elected. Depending on the specific process this type of participation has both advantages and drawbacks. In some cases a proportion of the delegates may be nominated to represent special interests [15,20]. The representative body may be a constituent assembly elected specifically for the purpose of preparing the new constitution or it may take the form of a regular legislature with an additional mandate to produce the new charter. In the context of a post-conflict transition the election of a constituent assembly is seen as providing an important opportunity for reconciliation through national dialogue [31].

2.3 Approval by Referendum

The draft charter is made public and citizens vote in a referendum to approve the proposed constitution, usually in a ‘yes’ or ‘no’ vote on the charter as a whole. Since

the early twentieth century public ratification has been the most common form of public participation in the process of producing a new constitution [20]. The knowledge that the charter will have to gain the formal approval of the citizenry helps to ensure that the drafters give due consideration to the expressed interests and aspirations of the people [15,7]. Processes in which public participation in a referendum was required to approve the constitution also tend to produce constitutions that adopt the referendum as a mechanism for public participation in governance decisions in the future [7,31,34]. Some experts believe that while the referendum is a useful legitimizing devise, it may not actually be necessary if the constituent assembly is fully representative, and can actually complicate matters and produce fresh divisions in society [15].

2.4 Civic Engagement and Outreach Campaigns

In recent years constitution-making processes have experimented with a growing range of strategies and methods that seek to educate citizens on the basic elements of constitutionalism and to survey their views or provide formal opportunities for consultations with groups representing various political, religious, professional, economic, cultural, and social interests and rights [21,34,32]. Brandt et al. [9] make the distinction between direct and indirect forms of participation. Direct participation includes traditional approaches that range from face-to-face meetings, community gatherings, and national conventions focusing on key interest groups or themes, to debates on specific issues and options, and public opinion polling using digital technology. Mechanisms to solicit and process written submissions are designed and incorporated in the formal process. Submissions can come from individuals or formally constituted bodies. Indirect forms of participation can vary from demonstrations in support of particular interests to other forms of lobbying, to any form of community or group mobilization or collective action to create pressure for a particular demand. It also includes written submissions and petitions submitted by individuals and groups where no formal process or mechanism exists [15,21].

2.5 Combining Forms of Participation: The South African Process

Most processes today use a combination of voting with civic engagement and outreach to promote popular involvement. Probably the most successful participatory constitution-building campaign ever conducted was the South African process that took place between 1989 and 1996 [21,9,1]. In April 1994, voters elected representatives to a constitutional assembly. From 1994 through 1996 these representatives engaged in an intensive outreach campaign to educate the public and provide opportunities for them to express their views and make submissions to the drafting body. A multi-media campaign provided information and awareness raising material via newspapers, radio and television, billboards, and public buses. The assembly also published its own newspaper reaching a circulation of 160,000. Humour was also employed to spark interest and fuel debate using cartoons. A website was designed to provide up-to-date information, and public meetings were organized to provide in-

formation, share opinions, and solicit input. Altogether these efforts are estimated to have reached some 73 percent of the population. The success of the campaign in generating public participation can be gauged from the two million opinions, petitions, and other contributions submitted to the assembly by individuals, civil society organizations, advocacy groups, professional associations, and other bodies between 1994 and 1996.

3 Post-Conflict Constitution Making

The many ways in which a constitution-making process contributes to the development of a democratic system are well documented [7,16,32,14,21]. A constitution establishes a system for the distribution of power and resources in society, regulating political institutions, constraining executive power, and protecting fundamental rights and privileges. By reaching out to the various communities and constituencies and bringing them together around the goal of developing a new constitution, the process can contribute to peacebuilding and reconciliation, educating the population and engaging them in a national dialogue on the form and function of the future state and their place in it. In the immediate aftermath of a revolution a participatory constitution-making process can serve two particularly urgent functions: (a) it provides a platform for engaging the major groups in society in the development of a new political system; and (b) it establishes a foundation for a culture of democratic political behavior without which the democratic transition is unlikely to survive.

3.1 The Challenge of Transition

In a revolution, those demanding radical change confront the reactionary forces of the old order and seek to replace a political system that does not meet their needs. Many internal conflicts occur when this political system fails to resolve differences between major groups in society. It may be that the old arrangements are no longer acceptable to some groups, or that the regime has taken on new powers and underestimated the depth of popular discontent and frustration. The overthrow of an autocratic regime tends to result in varying degrees of disruption and disorder as newly empowered actors seek to establish the legitimacy of the new political system. System legitimacy is derived from the belief that the existing political institutions are the most appropriate ones for the society [19]. It is generated from one of three sources - traditional, charismatic, or rational authority, or a mixture of the three Weberian types [35]. The main challenge of transition is to agree on a new system that satisfies the needs of these groups, sufficiently at least to allow the transition to proceed. If the proposed political system or the roadmap to produce the new system is not perceived as adequately accommodating to the claims of the major groups in society, a new crisis of legitimacy may develop during the transition [19]. The challenge for the interim authorities is to ensure the legitimacy of the new political system, its institutions and working arrangements, in the eyes of the major groups in society?

3.2 Legitimizing the New Order

The legitimacy of political institutions in Arab states tends to be based more on charismatic, strongman leadership or traditional arrangements rather than rational or legal authority (Lewis 2005). The overthrow of the old regimes in the Arab Spring revolutions removed the strongman rulers and sought to replace their political institutions with more rational legal arrangements accountable to the people. The most important challenge they face is how to establish a system that reconciles the demands of major groups with competing claims, historical grievances, questions of authority over particular communities and geographical areas, and issues of representativeness or legitimacy of transitional institutions or arrangements [19].

Proponents of participatory constitution-making point to a growing body of evidence that popular participation can build consensus among the main political groups about the type of political system and institutions [21,34,14,31]. If the public consultation process produces a broad consensus, it can become a cornerstone for the legitimacy, acceptance and stability of the new regime [23]. A representative process can significantly reduce the demand for renegotiation or the resistance of groups who claim that their interests have been neglected. In the post-revolution vacuum it can also help guard against manipulation by dominant or politically adept actors seeking to impose a particular agenda or to consolidate power. Experience suggests that even if people have not fully understood the issues, or the process was largely ceremonial they still feel a degree of ownership of and commitment to the resulting constitution [14].

The right to public participation in democratic governance exists in international law, notably in the International Covenant on Civil and Political Rights, which establishes minimum obligations for participation in public affairs [21]. In addition, the United Nations Committee on Human Rights has recognized a specific right to participate in constitution making [21]. In the context of a post-authoritarian transition the right to participate goes well beyond the legal dimension: the emotional importance of consulting the people who have fought for their right to decide what type of state and government they want cannot be overstated. Participation in the constitution-making process acknowledges the role they have played and the sacrifices they have made, and gives them the opportunity to build a legitimate new constitutional order.

3.3 Developing a Democratic Culture to Sustain the Transition

A second fundamental question for post-revolutionary Arab countries is how can they develop the democratic political behavior to sustain the transition? Experience shows that the violent overthrow of a regime tends to produce a "commandist" political culture that favors those who have more radical, militant, extreme, unquestioning, totalistic agendas. These groups tend to be driven by an ideal that fuels a belief in their monopoly of legitimacy, which is used to justify a lack of willingness to compromise or moderate demands. The singularity of purpose that inspired their military campaign against the old regime metastasizes into intolerance of other political forces whose opposition views are characterized as support for the former regime [11].

Democracy depends upon having not only the proper political institutions but also a democratic political culture – the values, attitudes, and perceptions that determine the way citizens think, believe and behave socially and politically [3]. High levels of interpersonal trust, political interest, involvement in community and civic organizations, and tolerance of others are all essential components of a democratic political culture [11]. These characteristics tended to be low among Arab citizens in the five societies that were surveyed twice in the Arab Barometer and actually decreased between the first and second surveys [33]. How, after decades of dictatorship, can they develop the culture of moderation, accommodation, cooperation, and bargaining among political elites that will allow them to sustain the transition and not revert to the strongman, "clever personality" [2] authoritarian rule of the past? Transforming the political culture of a society is one of the most difficult aspects of any post-conflict transition. Deeply engrained political and social practices built up over decades do not change with the holding of an election and the adoption of a new constitutional framework – it requires long-term strategies that engage citizens through cycles of civic education, dialogue and participation [32].

Research undertaken by the International Institute for Democracy and Electoral Assistance (IDEA) in twelve countries emerging from civil conflict or authoritarian rule suggests that constitution making processes that adopted participatory strategies and methods contributed to democratic education in societies that had not had political freedom or the chance to shape the governance of their state in the past. This led directly to the political empowerment of wide sections of the population and contributed to constitutions favoring free and fair elections, greater political equality, more social justice provisions, human rights protections, and stronger accountability mechanisms [31,22]. This is supported by Ghai and Gali [14]: "A constitutional review process with a careful scheme for public participation can, to a considerable extent, familiarize the people with the concept and procedures of political authority, and win support for the idea of a limited government that is bound by rules and accountable to the people". Research by the Comparative Constitutions Project also supports the contention that participatory processes produce constitutions that create better conditions for democratic consolidation by requiring governments to hold a national referendum when major changes in governance are proposed, and by guaranteeing a public role in approving constitutional amendments [20].

Overall, findings and analysis from a range of studies indicate that a participatory constitution-making process contributes to the transformation of the political culture by raising people's expectations through education about their rights and responsibilities and how they can engage with fellow citizens to bring about changes in government without resort to violence [5,21,14,15,20,26,32].

3.4 Limitations

For some scholars the risks of high levels of popular participation outweigh the benefits. Landau [25] cites Bolivia, Venezuela, and Egypt as examples of transitions where participation has created opportunities for powerful individuals and groups to manipulate popular demands, polarize participation and impose authoritarian agendas

at a time when institutional order is weak. Other researchers point to experiences where consultation has fuelled unrealistic expectations that cannot be met from states' limited resources, enshrining unattainable aspirational rights into the constitution, and subsequently failing to deliver on these rights in practice, thereby undermining the credibility of the participation process and jeopardizing the legitimacy of the constitutional order [26]. Nor are outcomes such as legitimacy and commitment as assured as some proponents might suggest: the IDEA study also found that while participatory processes resulted in constitutions enshrining rights for previously marginalized or excluded groups, or include provisions addressing issues of social and economic justice, and accountability "[t]hese provisions did tend to render the adoption and enforcement of the constitutions more controversial, as they were often perceived by the elites as a threat to their power or privilege" [32,22]. Analyzing the high level of participation in the Ugandan process, Moehler [27] concluded that those who had participated actively in the process were no more likely to support the constitution than were other citizens.

Overall, however, there is a large and growing body of research to support the contention that properly constituted popular participation in the making of a democratic constitution serves to legitimize the new political order and to initiate the development of the civic culture that is needed to sustain the transition to democracy.

4 Towards a Formula for Meaningful Participation

What constitutes meaningful participation by citizens is contentious at the best of times [4]. There is no set pattern for public participation in constitution making processes. In most cases it is seen as an element of the design of the overall process, in some it is an afterthought: "actual constitutional design processes employ scattered and usually rather anemic forms of popular participation and oversight to substitute for actual consent" [7]. Highlighting the absence of established standards for assessing whether a constitution-making process has been "free and fair" Brandt et al. [9] note that many processes are undertaken with little reflection about what constitutes a genuine and effective public consultation campaign. Large sums of money are spent only to have the views ignored or never analyzed.

In this section we look at some of the basic requirements for people to be able to understand and engage in constitution making. We then suggest a combination of elements that taken together could constitute a basic formula for meaningful participation in the process.

4.1 No Participation without Education

For their vote in the constituent assembly election to represent some form of meaningful participation, citizens need to be provided with a basic education on what a constitution is and what they can expect from it, what functions it performs in a democratic state, and how it can help build state institutions that better meet their needs as citizens and solve the governance problems that restrict their political, social and

economic development. This education is even more important in countries emerging from authoritarian rule where people are unlikely to be familiar with the concept of constitutional government or understand how a constitution can be used to protect their rights and fundamental freedoms and hold government accountable. Assembly elections and referenda are the most common traditional forms of consultation[1] [20]. However unless citizens have some understanding of how a constitution serves to determine their rights and status, the functions and limitations of government, the type of state they live in, and how their vote is likely to influence these provisions, citizens' participation in any constitution drafting election or referendum is likely to be a tokenistic exercise.

4.2 Forming and Aggregating Opinions

Deciding whether the provisions are acceptable and should be endorsed implies having had the opportunity to consider and discuss the alternative options in order to have formed at an opinion in the first place. Opportunities need to be ensured for citizens to access impartial information about the range of options to choose between in designing the constitutional framework – whether in relation to the type of government, levels of decentralization, executive power and constraints, fundamental rights, etc. – and their relative merits and drawbacks. Debated in public, the worth of different options can be seen by the strength of the arguments supporting them rather than which proposals are supported by the most powerful representative or the largest number of people [5]. It also contributes to reconciling different points of view, and is an essential stage in a process of reaching the compromises that permit workable solutions to political and social dilemmas [8]. Through informed discussion and debate citizens test and challenge the opinions and arguments of others, and may be compelled to accept a particular conclusion.

4.3 Incorporating Citizens Views

Providing channels for citizens to communicate their views to the constituent assembly is essential for a participatory process. When citizens vote to elect a representative it is reasonable to expect that there will be some formal process or mechanism to communicate their views on particular issues or provisions to that representative or to the body to which she has been elected. It is also important to note that in many developing countries both professional experts and technocrats come almost exclusively from the more privileged sections of society, the elite class with preferential access to education and the upper echelons of the public service. At a minimum this means they are less inclined to know about or understand the particular problems that arise from systematic exclusion and marginalization. It almost certainly introduces a conservative bias to their approach to policy and social development. It also has an impact on

[1] The Comparative Constitutions Project (CCP): Analyzing data on the content of 806 constitutions promulgated between 1789 and 2005 by Ginsburg et al detected a significant trend since the early twentieth century, toward public ratification making it the most common form of public participation in constitution making processes [20].

the public's perception of democracy: developing a democratic civic culture involves nurturing a belief on the part of the individual citizen that their participation can have an impact; that even as a single individual their voice will be heard and at least considered in the deliberation process, particularly when it coincides with those of others [11]. Ensuring formal channels for communication with the assembly are built into the official process is essential to the effectiveness and credibility of the process.

4.4 Special Arrangements for Traditionally Marginalized Groups

What about special arrangements to solicit feedback from traditionally excluded or disadvantaged communities or groups in society? Where these groups or communities have been historically excluded from political participation exclusion becomes self-fulfilling as a result of disaffection and apathy towards the political system. Special representation measures are needed to eliminate this systematic discrimination [36]. Special group representation measures may also be warranted in a multi-cultural society where an indigenous community has the right to some form of self-government [24]. There is a strong case for special arrangements for these groups in the process [9,14,15,22]. Ghai [15] recommends that states allow representatives of minorities and indigenous peoples, and minority-representative institutions, a special role – such as initiation, prior consultation and special voting rights – regarding provisions that have a major bearing on minority rights. This would imply a process of prior consultation, and review of articles or provisions intended to address their marginalization, or at a minimum the opportunity to highlight which articles did not meet their expectations, either prior to or as part of a referendum.

4.5 Democratizing the Drafting Process

Going even further, a process that allowed specially convened citizens' working groups to collaborate in drafting articles about particular issues that were of profound importance to them would be an even more meaningful form of participation. Although some scholars baulk at the prospect of methods that involve direct input from civil society citing "the magnitude of the challenges involved in absorbing public suggestions" or "the challenges of writing-by-committee, much less writing-by-nation" [20], it is difficult to accept that the advances made in digital media and communications technology do not offer more collaborative forms of drafting. In her book "Wiki-Government" former U.S. Deputy Chief Technology Officer Beth Noveck notes that while some activities require technical expertise, professionals, bureaucrats or lawyers do not have a monopoly on expertise: "a person may be expert on wetlands because she possesses professional credentialing. Another person may be an expert because she lives near one […] for every project there is a different kind of expertise, which could be sought" [28]. Allowing members of particular groups, communities or professional bodies to self-select to participate in a drafting group of their choosing is one way. Indeed as Beth Noveck notes "The ability to self-select to participate in the arena of one's choosing is what makes collaborative democracy egalitarian" [28].

5 A Formula for Meaningful Participation

These requirements suggest some basic prerequisites, which – taken together – could be seen as a formula for meaningful participation in the constitution making process. They have been treated in varying degrees of detail in various studies of recent constitution making processes. However, a review of these documents suggests that the key to this formula is in the interdependence of the different elements [7,8,9,20]. A process that provides all of these options for engagement at some level constitutes what we believe is a platform for meaningful participation in the constitution making process.

1. *Information about the Process*: Citizens cannot be expected to engage in the constitution-making process if they are not given basic information about how it will be conducted: what it is, who is responsible, how much time has been allocated, whether it is divided into different phases or stages to facilitate agreement on fundamental principles or major considerations prior to addressing the details of individual provisions, whether there will be a civic education campaign, whether or not citizens or interest groups will be consulted or given the opportunity to participate, how the draft will be approved and adopted, etc.
2. *Resources for Education*: A well-designed scheme for public participation should provide people with digestible civic education material through a variety of appropriate channels and products. Topics should be tailored to people's interest and ensure a basic education on a range of political concepts and procedures of political authority, forms of government, accountability, and how they can participate in the affairs of the state and protect their constitutional rights [9].
3. *Forums for Opinion Formation and Aggregation:* Meaningful participation also implies individual citizens forming opinions about what they feel are the best options for them as individuals and for their family, and community. Opinions are formed when people receive balanced or impartial information about an issue, and have the opportunity to express their interests and concerns and to question one another, respond to criticisms raised, and critique the arguments and proposals of others. Citizens need to be given the opportunity to debate contentious issues, to understand what options are available, to form and aggregate opinions.
4. *Channels for Communication*: Citizens also need to be given the opportunity to express their opinions, demands, expectations and priorities and to know that their opinion will somehow be communicated to the body responsible for drawing up the new charter. For this to happen the official procedure needs to include processes for soliciting citizens' input at three levels: a) passive monitoring of public opinions about general issues (system of government, unitary or federal state structures, etc.) as expressed in debates, public discussions, using online and traditional media; b) consultation with interest group representatives such as civil society organizations, professional associations, trade unions, cultural associations and rights groups, and so forth, about specific provisions; and c) a process of actively soliciting individual citizens' and interest group submissions through a formal dedicated mechanism for submissions to the assembly.

5. *Inclusive Deliberation Mechanisms:* During the drafting phase, these monitoring and consultation processes need to be complemented by mechanisms for formally engaging representatives and experts in deliberations about specific provisions. It is at this level that democratic participation acquires its highest expression. Self-selecting representatives of civil society organizations representing minorities, people with disabilities or other traditionally marginalized groups or communities, and people with specific expertise can be invited to contribute to the drafting of particular provisions or articles or to provide feedback on drafts prepared by the responsible thematic committee or subcommittee.

6 Empowering Meaningful Participation in Constitution Making

This section of the paper looks at the rapid growth of Internet and social media use in the Arab Region in recent years and the ways in which Internet and mobile technologies are contributing to unprecedented social networking and activity. Here we pose the fundamental question: can the mobile technology and social media platforms that powered the Arab Spring revolutions also facilitate meaningful participation in the making of new constitutions for these countries? Can they be harnessed to meet the five prerequisites for meaningful participation outlined above? A review of how young people are using Internet and social media in the Arab Region suggests that all of these needs can be met much more effectively through Internet and social media platforms.

A 2013 survey conducted across the Arab Region by the Dubai School of Governance found that the region had more than 125 million Internet users with an average annual growth rate close to 30%. The average penetration rate in the region was almost 30 percent while some countries (UAE, Bahrain, Qatar and Kuwait) reached 50 percent or above. Even in countries like Morocco, Sudan and Yemen that had significantly lower penetration rates growth rates were among the highest in the region [12].

- Eighty-six percent of the 15 to 35 years age group access the Internet on a daily basis, compared with 63 percent of the 49 to 65 years group.
- More than one third (36%) of respondents reported spending between 3 and 4 hours on the Internet everyday.
- Forty-nine percent of respondents spend most of their time on the Internet after 6pm on workdays.
- Eighty-eight percent indicated that they access the internet from home, while 54% access the internet from work, with only 12% of respondents accessing the internet from school or university.

6.1 Access to information

Seventy-one percent of respondents in the Dubai School of Governance survey agree that online communication has replaced traditional communication, and 85% believe

that social media has enabled their social activity [12]. Forty-three percent connect with friends several times a day and 23% said they use instant messaging several times a day. The 2013 DSG report also supported the contention that a dedicated Internet platform could significantly enhance citizens' awareness of developments in the constitution making process [12]. The survey found that Internet is the primary source of news about current events for 36% of respondents , with more users (29%) getting their news from social media than from traditional media sources (28%). Facebook is the most popular social network with 54% of respondents indicating they use Facebook more than once a day, followed by Google+ (30%) and then Twitter (14%) [12]. 69% of respondents research their interests at least once a day and only 1% have never done so [12].

6.2 Customised education resources

Findings from the 2013 DSG survey strongly support the hypothesis that Internet can provide an important channel for education on the basic principles of democracy and constitution making, with one person in four taking online educational courses while one in three take language courses on the Internet [12]. Thirty-two percent of respondents use language learning platforms at least once a day and more than one-in-four (27%) take formal online courses several times a week, while 22% reported taking free online courses at least several times a week, 14% of these on a daily basis. The use of educational videos was particularly popular: 31% of respondents reported watching instructional videos at least once a day.

6.3 Opportunities to receive impartial analysis to help form and aggregate opinions

The use of weblogs or blogs during the Arab Spring revolutions showed the ease with which young people in the Arab region could share information using social media in an effort to influence their peer's opinions and shape their social and political reality. A 2012 survey of 3,000 Internet users conducted by the Ideation Centre in nine countries in the Arab Region also suggests that Internet platforms that incorporate blogging have the potential to play an important role in informing and shaping opinions on complex concepts related to values and beliefs [30]. This was reinforced by the finding that almost all religious figures in these countries now provide online guidance through their own blogs allowing lay people to access different schools of thought and effectively removing the hierarchical aspect of religious discourse. Seventy percent of respondents reported using Internet to explore different aspects of religion and find answers to questions [30]. Responses to the 2013 DSG survey support this contention with 41% of respondents reporting that they read educational blogs at least once a day [12].

6.4 Channels for expressing these opinions and communicating with those responsible for preparing the draft charter

A 21-nation survey conducted by the Pew Research Center's Global Attitudes Project highlighted the fact that social media users in Egypt, Tunisia, Jordan, and Lebanon were much more likely to express their opinions about politics, community issues, and religion. Between 60 and 68% of social media users have posted comments expressing their opinions about politics online. This is up to twice the median (34%) across 20 of the nations surveyed. Similarly while a cross-national median of 46% reported sharing their views about community issues on social media, the figure rises to 82% of users in Tunisia, closely followed by 81% in Lebanon, 80% in Jordan, and 74% Egypt. In the same survey users from 63% of users from a median of 14% reported sharing views about religion on social networking sites – the figures for three of the four Arab Region countries were almost four times higher. Although the majority of Arab Internet users surveyed in the 2012 Ideation Centre study have not used any online platform to participate in government processes many say they would if they were given the opportunity. More than half respondents (56 percent) in post-revolution Egypt said that they would use such a service, compared with 28 percent in Algeria. Interestingly, 43 percent of respondents said they would use such a platform is to give a suggestion, while 39 percent would use it to voice opinions about political or social matters, and 29 percent would use it to compliment the leadership [30].

6.5 Collaborative Drafting Mechanisms to allow citizens or representatives to collaborate in drafting specific elements of the new constitution

On 9 March 2011, in response to a wave of demonstrations across by the country demanding democratic change, the King of Morocco appointed a commission to draft a new constitution. Moroccan citizens would be given the opportunity to vote in a referendum on the constitution as a whole at the end of the review process [22]. There were very limited opportunities for meaningful engagement.[2] In an effort to promote meaningful participation in the process activists [3] launched the website www.reforme.ma on which they uploaded the entire existing constitution and asked citizens to show their support for – or dissatisfaction with – each article or part of an article in the constitution thereby indicating clearly to the members of the appointed commission exactly where they wanted changes to be made. People were also able to rewrite articles or add new ones. To attract people to the site, Reforme partnered with Facebook groups created by activists and CSOs. Full Facebook and Twitter integration allowed anyone who was a "friend" or "follower" of anyone already commenting

[2] The commission was headed by an advisor to the king, as was a consultative body whose role was to liaise between the drafters of the constitution and political parties, labor unions, professional and business associations, NGOs and any individuals interested in making a submission to the new constitution. In a classic example of "participation without power" there was no follow up or debate once the submissions had been made and the members of the consultative body were only shown a written draft of the new constitution the day before the king presented it to the nation in a televised speech [29].

[3] Led by Tarik Nesh Nash.

on the site to see what her friend on Facebook or Twitter opted to "Like" – essentially creating a viral effect.

On the strength of this review it would seem that social media and mobile communications technology have significant potential to overcome the limitations experienced by traditional forms of participation. If this is the case, how would advocates of online participation go about building an Internet platform that delivers the five elements identified in the formula presented above?

7 "My.Con" – Options for Digital Participation

This section presents a proposal for an interactive Internet platform that deploys social media tools and applications that have revolutionized the way people learn, communicate, and network to maximize the opportunities for citizens to keep up to date with the process, learn about the issues, seek advice, debate options and share opinions, propose submissions, and collaborate in drafting specific provisions. Each of its main features is designed to address one of the prerequisites in the formula for meaningful participation so that together they constitute a comprehensive online platform for popular participation on an unprecedented level.

7.1 Internet Notice Board

The greatest utility would surely be served by an online notice board that concentrates in one site information about the official constitution-making process and also features all the major news stories about related developments and events as they happen. Ideally this part of the My.Con platform would have a formal link to the constituent assembly secretariat to receive official information on the programme and updates on the work of the assembly and its various committees and sub-committees. In addition to news sourced through RSS feeds the Online Notice Board would engage traditional media to both share relevant news stories and re-broadcast content that has been developed through the platform. This would both ensure a regular flow of relevant news while extending the reach of the platform to citizens who do not have access to Internet. A linked page would provide NGOs with an online platform to publicize their civic awareness and outreach work on the constitution-making process. Information about events such as public debates, conferences, town-hall meetings, training workshops, youth forums, etc. would be solicited from NGOs and published in a calendar of events. Planned events would be shown on an interactive map so that people could see what is going on in their area. This would also allow the constituent assembly and NGOs to see what areas are underserved with civic education activities and take action to organize events for people in these areas.

7.2 Online Video Tutorials

The objective of the online education section of the platform is to empower citizens with education about the main constitutional topics and issues so that they can participate meaningfully in the process. A basic education on the key issues would allow people to engage in informed debate, advocate for specific rights and interests, and contribute to decision-making about the future of the state. This section would feature a series of 2-to-3 minute video tutorials introducing basic concepts related to democratic government, constitutions, and the constitution making. The mini-tutorials would be presented by a qualified communicator and illustrated with visual representations using images, graphics, art, etc. to help the viewer remain engaged and better understand the concepts being presented. Users would be able to watch tutorials online (streaming) or download to a device. The download feature is included so that people who do not have access to Internet can also benefit from the tutorial. It also allows the tutorial to be shown to a group – e.g. students in a class or lecture setting, family members at home, NGO staff or members, etc. This function is typically undertaken through workshops or printed material. Working in partnership with universities, secondary schools, NGOs, and traditional media using a carefully designed dissemination strategy would help ensure the tutorials would benefit a broad section of the popular.

7.3 Expert Discussion Forum

At the heart of the My.Con platform's online discussion forum is the idea that information or analysis from credible sources helps people form their own opinion about an issue, particularly if the information is accompanied by a discussion or commentary from a range of perspectives. The forum centres on a blog with regular posts that shed light on a constitutional topic or recent development, highlight the main points in a process, or offer useful suggestions about how to improve knowledge or take action. Short posts like "Five Things the Constituent Assembly Needs to Do in Its First Meeting" or "Three Ways You Can Participate in the Constitution Making Process" empower users quickly with information they need to get involved in the process. Readers can also leave comments in an interactive format that acts as a kind of discussion forum. Research suggests that the inclusion of a discussion forum where users can agree or disagree with the position taken in the blog post and write comments contributed significantly to the popularity of many blogs. It also helps inform users' opinions about the major questions as they arise or the choices that need to be made in relation to specific options in the constitution making process. Expert contributors would also identify the critical questions about the blog theme that would constitute the subjects for the discussion forum, and facilitate an online question and answer session in real time on the issues highlighted in the blog. These sessions would be widely publicised on the blog page and though dedicated platform Facebook and Twitter accounts to attract user participation.

7.4 Upload Your Submission Channel

The objective of the "Upload Channel" part of the platform is to mobilize members of the public to record and upload their personal submission - or that of a relative, friend or neighbor - to the constituent assembly. Anyone with access to a smartphone will be able to submit their opinion by recording and uploading a short (30 second) video explaining the priorities to be included in the constitution or addressed by the Drafting Assembly. Uploaded on the constitution making platform, YouTube, and social media these videos become a form of submission to the constituent assembly. Users will be able to select an option to have their submission automatically posted to their Facebook page and shared with their friends, and to be entered in a weekly competition where videos are posted on the platform Facebook page and users vote for their preferred upload. To make it easy for people to record their submissions, a specially designed app providing a one-touch record-and-upload service could be developed and made available for download free-of-charge. This can be a particularly powerful way of giving voice to people who might not otherwise be able to interact with the constituent assembly or submit any proposal to them. At different stages in the process a video-collage could be compiled bringing together the video submissions that best express the views, concerns and aspirations of citizens for the new constitution. Special screenings could also be organized for the assembly both individual videos and the final collage could be widely disseminated through social media, YouTube, television, and traditional media.

7.5 Collaborative Drafting Forum

This component of the My.Con platform would engage online teams composed of civil society representatives, interest group advocates, and academic experts in the collaborative drafting of articles using a specially designed wiki. The primary users would be self-selecting experts and representatives of civil society organizations who speak on behalf of a CSO, professional association, NGO, academic body, or group of students. The teams would be structured with set roles assigned according to different types of experience and qualifications matched to key tasks. For example, roles could include facilitator, drafters, advocates, subject specialists, and researchers who review and analyze proposed articles, research, upload, annotate and cross reference relevant articles from other constitutions draft new formulations and submit for review. Inspired by the Peer-to-Patent initiative [28] the forum will make extensive use of a visualization process by which participants can see on their screen the roles they have accepted and tasks they have been assigned. Draft articles could be shared on Facebook and voted on by the public prior to submission to the constituent assembly. This facility could also be used to organize public commentary on individual articles following the publication of a draft constitution prior to referendum. Citizens would be able to identify articles that they did not agree with and propose alternative formulation or provisions, or to vote for the constitution on an article-by-article basis.

8 Implementation

Would it work? Constitutional law is a dull subject at the best of times. Would people make use of a social platform about constitution making? A 2012 Stanford University study of the impact of social media on social unrest in the Arab Spring countries could provide the key. Researchers found that "In the hands of civil society members, the Internet has enormous effects upon protest probability" [10]. Results from the study provide strong evidence that the probability of having participated in the protests increased up to ten times among people who were both members of civil society groups and Internet users. Most interesting is the finding that it is the combination of civil society group membership reinforced with virtual community membership that seems to be the strongest motivating factor in bringing about participation in protest activity. The study supports the theory that Internet communities can serve a similar function to civil societies in that they provide a platform to connect individuals who share a common interest. Online collaboration is enhanced when participants have already developed shared interests through membership in actual civil society networks. The online platform allows them to discuss common socioeconomic grievances and political concerns in the same way as a physical meeting. In the case of the online collaboration however the obstacles presented by security, transport, time, and limited physical information resources are removed and the potential for participation is limitless. This finding, that social media can best facilitate participation when it builds upon existent social ties, such as those created in civil society groups, may be the key to a successful platform. Given the potential of social media platforms offer to meet the basic prerequisites for meaningful participation this would seem to suggest that a dedicated platform that made the information, education, opinion formation and communication opportunities available would, in the first instance, have the maximum impact if it were made available through civil society organizations with an existing membership base.

9 Conclusion

Traditional forms of engagement in constitution-making processes have had limited success in delivering a basic combination of prerequisites for meaningful participation. But Internet and digital communications technology are already overcoming these limitations. Today there are social media platforms that respond directly to the need for information about the process and offer online education about substantive issues; blogs and micro-blogs that provide guidance on contentious issues and options; and any number of new apps that could facilitate consultations with citizens and ensure that the outcomes of these consultations are synthesized and incorporated into deliberations and drafting. Internet and mobile communications technology are fuelling a level of networking and social participation that used to be the stuff of science-fiction movies. The paper hypothesizes that participation in constitution making can be revolutionized with an online platform that offers information and targeted education features and deploys tailor-designed apps to allow users to express their opinions

and communicate them to the official drafting body. A proposal for the design of such a mash-up platform was presented as a model for testing inappropriate constitution making contexts.

10 References

1. Afsah, I. (2012) "Constitution-Making in Islamic Countries – A Theoretical Framework" in Grote, R. & Roder, T.J. (Eds.) Constitutionalism in Islamic Countries: Between Upheaval and Continuity. Oxford University Press: New York
2. Al-Azm, Sadik (1968) "Self-criticism after the defeat" Saqi Books: London
3. Almond, G. A. & Verba, S. (1969) The Civic Culture: Political Attitudes and Democracy in Five Nations. Princeton University Publications: NJ
4. Arnstein, Sherry R. (1969) A Ladder of Citizen Participation. JAIP, Vol. 35, No. 4, July 1969, pp. 216-224.
5. Banks, A. M. (2008) Expanding Participation in Constitution Making: Challenges and Opportunities, 49 Wm. & Mary L. Rev. 1043 (2008) http://scholarship.law.wm.edu/wmlr/vol49/iss4/2
6. Blount, J. (2011) "Participation in Constitutional Design" in Ginsburg, T. and Dixon, R. (Eds.) Comparative Constitutional Law. Edward Elgar: MA
7. Blount, J., Elkins, Z., & Ginsburg, T. (2012) "Does the Process of Constitution-Making Matter?" in Ginsburg, T. (Ed.) *Comparative Constitutional Design* Cambridge University Press: New York
8. Böckenförde, M., Hedling, N. & Wahiu, W. (2011) A Practical Guide to Constitution Making. International Institute for Democracy and Electoral Assistance (International IDEA). IDEA: Stockholm
9. Brandt, M., Cottrell, J., Ghai, Y., and Regan, A. (2011). Constitution-making and Reform: Options for the Process Date Interpeace: Geneva
10. Dewey, T., Kaden, J., Marks, M., Matsushima, S., & Zhu, B. (2012) The Impact of Social Media on Social Unrest in the Arab Spring. Research report prepared for Defense Intelligence Agency, March 20, 2012. Stanford University: Stanford, CA
11. Diamond, L. (1999) Developing Democracy: Towards Consolidation. Baltimore: John Hopkins University Press
12. DSG (2013) The Arab World Online: Trends in Internet Usage in the Arab Region, http://www.dsg.ae/en/publication/index.aspx?&PriMenuID=11&mnu=Pri#sthash.IAg43cXo.dpuf
13. Elkins, Z., Ginsburg, T., & Blount, J. (2008) "The Citizen As Founder: Public Participation In Constitutional Approval" in Temple Law Review Vol. 81:2008, pp101-121
14. Ghai, Y. & Gali, G. (2006). Constitution Building Processes and Democratization. International IDEA:Stockholm
15. Ghai, Y. (2004) The Constitution Reform Process: Comparative Perspectives. Presented at the conference "Toward Inclusive and Participatory Constitution Making" 3-5 August, 2004, Kathmandu (Nagarkot)
16. Ghai, Y. (2005) The Role of Constituent Assemblies in Constitution Making. IIDEA: Stockholm
17. Ginsburg, T. (2011) "Constitutional Endurance" in Ginsburg, T. & Dixon, R. (Eds.) *Comparative Constitutional Law*. Edward Elgar: Cheltenham
18. Ginsburg, T. & Dixon, R. (2011) Comparative Constitutional Law. Edward Elgar: Cheltenham

19. Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. American political science review, 53(1), 69-105.
20. Ginsburg, T., Elkins, Z., & Blount, J. (2009) "Does the Process of Constitution-Making Matter?" in The Annual Review of Law and Social Science 2009.5:201-223
21. Hart, V. (2003) Democratic Constitution Making (July 2003), USIP Special Report 107, United States Institute of Peace *available at* http://www.usip.org/pubs/specialreports/sr107.pdf
22. IDEA (2012) The 2011 Moroccan Constitution: A Critical Analysis. IDEA: Stockholm
23. Klein, C. & Sajo, A. (2012) Constitution-Making: Process and Substance. In Rosenfeld, M & Sajo, A. (eds) The Oxford Handbook of Comparative Constitutional Law. Oxford University Press: Oxford
24. Kymlicka, W. (1996) Multicultural Citizenship. Oxford University Press: Oxford
25. Landau, D. (2013). Abusive Constitutionalism. *UC Davis Law Review*, Fall 2013, Forthcoming; FSU College of Law, Public Law Research Paper No. 646. Available at SSRN: http://ssrn.com/abstract=2244629
26. Liolos, J. (2013) Erecting New Constitutional Cultures: The Problems And Promise Of Constitutionalism Post-Arab Spring. Boston College International & Comparative Law Review [Vol. 36:1 2013]
27. Moehler, Devra C. (2008) Distrusting Democrats: Outcomes of Participatory Constitution Making University of Michigan Press: Michigan
28. Noveck, B. (2010) Wiki Government: How Technology Can Make Government Better, Democracy Stronger, and Citizens More Powerful Brookings Institution Press: Washington
29. Ottaway, M. (2011) The Moroccan Constitution: Real Change or More of the Same? Carnegie Endowment for International Peace. Available at: http://carnegieendowment.org/2011/06/20/new-moroccan-constitution-real-change-or-more-of-same/
30. Sabbagh, K., Mourad, M., Kabbara, W., Shehadi, R., & Samman, H. (2012) Understanding the Arab Digital Generation. A report for the Ideation Centre. Booz & Co. Accessed online at http://www.booz.com/media/file/BoozCo_Understanding-the-Arab-Digital-Generation
31. Samuels K. 2006. Constitution Building Processes and Democratization: A Discussion of Twelve Case Studies. Geneva: IDEA
32. Samuels, K. (2006) 'Post-Conflict Peace-Building and Constitution-Making' in Chicago Journal of International Law Vol. 6 No. 2
33. Tessler, M., Jamal, A., & Robbins, M. (2012) New Findings On Arabs And Democracy. In Journal of Democracy Volume 23, Number 4 pp. 89-103, October 2012. Baltimore. JHU
34. Voigt, S. (2003) "The Consequences of Popular Participation in Constitutional Choice - Towards A Comparative Analysis" in Deliberation And Decision: Economics, Constitutional Theory, And Deliberative Democracy. Anne van Aaken, Christian List, Christoph Lutge, eds., Aldershot: Ashgate, 2003
35. Weber, M. (1968) Economy and Society Bedminster Press: New York
36. Young, I. M. (2007) Inclusion and Democracy. Oxford University Press: Oxford

A Multi-Agent Experiment on the Acquisition of a Language System of Logical Constructions

Josefina Sierra-Santibáñez[1]

Universidad Politécnica de Cataluña, Spain
(Maria.Josefina.Sierra@upc.edu**)

Abstract. This paper analyses an experiment which studies the acquisition of the linguistic competence required to communicate logical combinations of categories from the wisdom of the crowds perspective. The acquisition of such competence encompasses both the construction of a set of logical categories by each individual agent and of a shared language by the population. The processes of conceptualisation and language acquisition in each individual agent are based on general purpose cognitive capacities such as discrimination, invention, adoption and induction. The construction of a shared language by the population is achieved using a particular type of linguistic interaction, known as *the evaluation game*, which gives rise to a shared language system of logical constructions as a result of a process of self-organisation of the individual agents' interactions, when these agents adapt their languages to the expressions they observe are used more often by other agents.

1 Introduction

The wisdom of the crowds main thesis is that a diverse collection of independently deciding individuals is likely to make certain types of decisions and predictions better than individuals or even experts. This principle seems to work for many naturally occurring systems such as ant colonies, bird flocks or moving traffic flows, and it has been successfully applied to market prediction [1, 2] and multi-agent computer systems as well [3]. However not all crowds (groups) are wise, and it is therefore important to identify some criteria which separate wise crowds from irrational ones. Four such criteria are described in [4]: (1) diversity of opinion, enough variance in approach, thought processes and private information is necessary; (2) independence, agents' decisions should not be determined by other agents; (3) decentralisation, agents should be able to specialise and draw on local knowledge; and (4) aggregation, some mechanisms should be provided for turning individual decisions into collective ones.

Two additional important aspects of the wisdom of the crowds approach are also pointed out in [5]: the necessity of designing methods for describing how a group thinks as a whole; and the importance of disagreement and contest as mechanisms that enable the generation and selection of optimal decisions.

** The research leading to these results has received funding from the Spanish Ministry of Economy and Competitiveness under project TIN2011-27479-C04-03.

Experiments studying the effectiveness of the wisdom of the crowds approach often incorporate some functions which allow assessing the performance of a group in a given task, thus making it possible to establish a comparison between the collective performance and that of its individuals. Some of these functions are sometimes referred to as *collective intelligence quotient* (or *cooperation quotient*) and compared with the *individual intelligence quotient (IQ)*.

There are, however, other definitions of collective wisdom which not only focus on consensus-driven decision making, but on other aspects of it such as: shared knowledge arrived at by individuals and groups; shared intelligence that emerges from the collaboration, collective efforts and competition of many individuals; or collective learning over time. For example, [6] defines the collective intelligence phenomenon as 'the capacity of human communities to evolve towards higher order complexity and harmony, through such innovation mechanisms as differentiation and integration, competition and collaboration'. A step forward in this direction is *crowdsourced crisis mapping* [7, 8], which tries to bridge the gap between the creation and sharing of knowledge by global communities and the necessary action to solve social problems based on that information. Interesting projects addressing related issues such as the construction of a democratic political culture [9] or generalised access to education using crowd-based socio-cognitive systems [10, 3] are additional examples.

The rest of the paper is organised as follows. Firstly, we present the results of a multi-agent experiment in which a group of autonomous software agents try to construct at the same time a set of logical categories and a shared language. Then, we analyse such results from the wisdom of the crowds perspective, i.e. taking into account the definitions of wisdom of the crowds and criteria for distinguishing wise crowds from irrational ones introduced in this section.

The multi-agent experiment is not described in detail in this paper, although its main characteristics have been outlined in the abstract. A complete description of the evaluation game and the mechanisms the agents use for discrimination, induction and adaptation can be found in [11]. A summary of the main steps of the evaluation game, the induction rules and the adaptation strategies used by the agents are also given in appendixes A y B.

2 Results of the Experiment

As mentioned above, the multi-agent experiment analysed in this paper studies the acquisition of the linguistic competence required to communicate logical combinations of basic categories, such as 'up and to the left' (i.e. [and, up, left]), 'not up or to the right' (i.e. [if, up, right]) or 'either up or to the right, but not both' (i.e. [xor, up, right]). The acquisition of such competence encompasses both the construction of a set of logical categories by each individual agent and of a shared language by the population. In particular, the set of logical categories the agents can construct in this experiment is the set of Boolean functions of one or two arguments *not, and, nand, or, nor, xor, iff, if, nif, oif* and *noif*. Boolean functions *not, and, or, if* and *iff* correspond to the connectives of propositional

logic $\neg, \wedge, \vee, \rightarrow$ and $\leftrightarrow$ respectively. The semantics of Boolean functions *nand, nor, xor, nif, oif* and *noif*, assuming they are applied to propositions A and B, can be defined by the following formulas $\neg(A \wedge B), \neg(A \vee B), (A \vee B) \wedge \neg(A \wedge B), \neg(A \rightarrow B), B \rightarrow A, \neg(B \rightarrow A)$ respectively.

The experiment involves a population of autonomous software agents which are made to interact with each other playing language games. The particular type of language game used in the experiment analysed in this paper is called the evaluation game. It is played by two agents, a speaker and a hearer. It requires the agents to communicate about subsets of objects of the set of all the objects in a given context. In order to do so, the speaker must construct a logical combination of categories that is true for the subset of objects it tries to communicate about and false for the rest of the objects in the context, i.e. a conceptualisation of the subset. Then, it should transform this conceptualisation into an utterance using its lexicon and grammar, and communicate that utterance to the hearer. The hearer then tries to parse the utterance, reconstruct its meaning and use it to identify the subset of objects the speaker had in mind. Depending on the outcome of the game speaker and hearer use different strategies to expand and adapt their internal languages in order to be more successful in future language games. All agents in the population play both the role of speaker and that of hearer in different language games.

In the experiment, the agents are initially endowed with a set of cognitive abilities for discrimination, invention, adoption and induction that are hypothesised to be necessary for seeing the emergence of possible language strategies to be successful in the evaluation game. Then, they are made to play a series of language games, where they configure possible strategies and try them out. The goal of the experiment is to find out whether the population as a whole succeeds in the evaluation game, i.e. communicates effectively, and to observe the conceptualisations and language strategies that emerge in the population as a result of the processes of collective invention and negotiation.

In the particular multi-agent experiment analysed in this paper we have performed several simulation runs. In each simulation the agents first play 700 evaluation games about subsets of objects which can be discriminated using a single category or the negation of a category. In this part of the simulation the population reaches a *communicative success* of 94% after playing 100 games (see figure 1). *Communicative success* is the average of successful evaluation games in the last ten games played by the agents. Next, the agents play 6000 evaluation games about subsets of objects which require logical combinations of one or two categories for their discrimination. In this part of the simulation the population reaches a communicative success of 100% after playing 3600 evaluation games. As it can be observed in figure 1, this level of communicative success is maintained until the end of the simulation. The results shown in the figure are the average of ten independent simulation runs with different random seeds.

At the end of a typical simulation run the set of logical categories and grammatical constructions built by each agent are not necessarily equal to the set of logical categories and grammatical constructions built by other agents. However

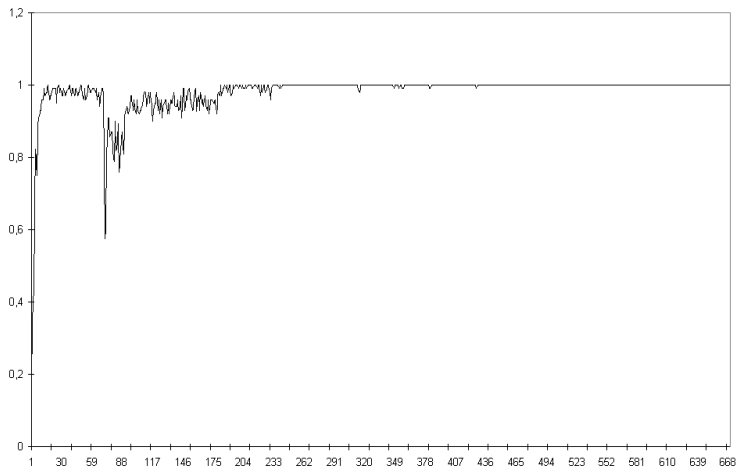

Fig. 1. Evolution of communicative success for a population of three agents.

they are compatible in the sense that they guarantee the unambiguous communication of logical combinations of one or two categories.

Let us focus now on the set of logical categories and grammatical constructions built by three agents at the end of a particular simulation run (see table 1). All the agents have constructed a grammar rule for expressing negations, and all of them use the same expression (i.e. *cp*) for referring to logical category *not*.

All the agents have constructed logical categories for all **commutative boolean functions of two arguments** (i.e. *iff, xor, and, nand, or* and *nor*) as well; and all of them prefer the same expressions for naming such categories (*j, wbt, y, nb, dol* and *ssq* respectively).

In order to express logical formulas constructed with binary Boolean functions, the agents use two types of grammar rules. Which rule is used for expressing a given formula depends on the Boolean function appearing in that formula and the syntactic category of the expression associated with such Boolean function. Syntactic category $c1$ is used in grammatical constructions which place the expression associated with the first argument of a Boolean function in the second position of the sentence and the expression associated with the second argument of the Boolean function in the third position of the sentence. Syntactic category $c2$ is used in grammatical constructions which place the expression associated with the first argument of a Boolean function in the third position of the sentence and the expression associated with the second argument of the Boolean function in the second position of the sentence. The expression associated with a Boolean function is always placed in the first position of the sentence in this experiment.

We now consider **non-commutative binary Boolean functions** (i.e. *if, nif, oif* and *noif*). All the agents have constructed logical category *nif*, which

Grammar a1
s([not,X],Q) → cp, s(X,P), {Q is P · 1}
s([X,Y,Z],T) → c1(X,P), s(Y,Q), s(Z,R), {T is P · Q · R · 1}
c1(nif,R) → ml, {R is 1}
c1(nor,R) → nb, {R is 1}
c1(or,R) → y, {R is 1}
s([X,Y,Z],T) → c2(X,P), s(Z,Q), s(Y,R), {T is P · Q · R · 1}
c2(and,R) → j, {R is 1}
c2(xor,R) → dol, {R is 1}
c2(iff,R) → ssq, {R is 1}
c2(nand,R) → wbt, {R is 1}
c2(if,R) → why, {R is 1}
Grammar a2
s([not,X],Q) → cp, s(X,P), {Q is P · 1}
s([X,Y,Z],T) → c1(X,P), s(Y,Q), s(Z,R), {T is P · Q · R · 1}
c1(nif,R) → ml, {R is 1}
c1(nor,R) → nb, {R is 1}
c1(or,R) → y, {R is 1}
s([X,Y,Z],T) → c2(X,P), s(Z,Q), s(Y,R), {T is P · Q · R · 1}
c2(and,R) → j, {R is 1}
c2(xor,R) → dol, {R is 1}
c2(iff,R) → ssq, {R is 1}
c2(nand,R) → wbt, {R is 1}
c2(if,R) → why, {R is 1}
Grammar a3
s([not,X],Q) → cp, s(X,P), {Q is P · 1}
s([X,Y,Z],T) → c1(X,P), s(Y,Q), s(Z,R), {T is P · Q · R · 1}
c1(nif,R) → ml, {R is 1}
c1(nor,R) → nb, {R is 1}
c1(or,R) → y, {R is 1}
c1(oif,R) → why, {R is 1}
s([X,Y,Z],T) → c2(X,P), s(Z,Q), s(Y,R), {T is P · Q · R · 1}
c2(and,R) → j, {R is 1}
c2(xor,R) → dol, {R is 1}
c2(iff,R) → ssq, {R is 1}
c2(nand,R) → wbt, {R is 1}

Table 1. Logical categories and grammatical constructions built by each agent at the end of a particular simulation run. In principle, the agents can construct logical categories *not, and, nand, or, nor, if, nif, oif, noif, iff* and *xor*, although they do not necessarily construct all of them. Boolean functions *not, and, or, if* and *iff* have the standard interpretation ($\neg, \wedge, \vee, \rightarrow$ and $\leftrightarrow$ respectively). The rest can be defined as follows: (A nand B) is equivalent to $\neg(A \wedge B)$, (A nor B) to $\neg(A \vee B)$, (A nif B) to $\neg(A \rightarrow B)$, (A oif B) to $(B \rightarrow A)$, (A noif B) to $\neg(B \rightarrow A)$, and (A xor B) to $\neg(A \vee B) \wedge \neg(A \wedge B)$.

corresponds to *the negation of an implication*, all of them use the same expression (i.e. *ml*) for referring to it, and all of them associate the expression *ml* with syntactic category $c1$.

None of the agents has constructed logical category *noif*. But this does not prevent them from characterising any subset of objects, because formulas [noif, A, B] and [nif, B, A] are logically equivalent and all the agents have constructed logical category *nif*.

Let us focus now on differences. Agents a1 and a2 have constructed logical category *if* (i.e. logical implication), whereas a3 has not. On the other hand, agent a3 has constructed logical category *oif*, while agents a1 and a2 have not. However the lack of only one of these two logical categories does not prevent any agent from characterising any subset of objects, because formulas [if, A, B] and [oif, B, A] are logically equivalent. Furthermore, the three agents can always understand each other. Because the word agents a1 and a2 use for referring to logical category *if* (namely *why*) is the same word agent a3 uses for referring to logical category *oif*; and the syntactic category agents a1 and a2 associate with such word (i.e. $c2$) is different from the syntactic category agent a3 uses for it (i.e. $c1$), which means that agent a3 does not invert the order of the expressions associated with the arguments of *oif* in the sentence whereas agents a1 and a2 invert the order of the expressions associated with the arguments of *if*.

3 Discussion

The experiment described in this paper constitutes an example of *collective learning and coordination*. As we have explained above the agents construct a language system of logical constructions that allows them to communicate logical combinations of categories. This language system includes a common vocabulary for logical categories, and a set of grammatical constructions which allow them to order the expressions associated with the components of logical formulas in sufficiently similar ways as to ensure unambiguous communication.

In this section we try to analyse the results of the experiment from the wisdom of the crowds perspective, focusing on the definitions of wisdom of the crowds and criteria for distinguishing wise crowds from irrational ones introduced in section one.

First of all, does the population make better decisions than individual agents in the experiment? It might be difficult to answer such a question without knowing in detail the mechanisms each agent uses to construct logical categories, invent new words and induce grammatical constructions (which are described in detail in [11]), but we think it does. The population is able to recognise that certain binary Boolean functions are redundant. For example, *nif* and *noif* can be used for discriminating the same subsets of objects, and the same happens with *if* and *oif*. Consequently, the language of the population contains only two words for the four logical categories (*if, nif, oif, noif*). The mechanisms the individual agents use for constructing a set of logical categories and grammatical constructions do not allow them to discover such redundancies for themselves. It

is the interaction with other agents, who use different formulas for conceptualising the same subset of objects, what generates the opportunity of first using the same word for two different categories, and then selecting a single meaning for that word as a result of the selection process that takes place among competing associations between expressions and meanings both in the individual languages constructed by each agent and in the language constructed by the population.

The population is also able to discover that the word order of the expressions associated with the arguments of commutative Boolean functions is irrelevant for language understanding. This cannot be observed in the grammars shown in table 1. But in other simulation runs we have performed the word associated with a commutative Boolean function such as *and* can be associated with syntactic category $c1$ in an agent's grammar and with syntactic category $c2$ in the grammar of a different agent of the same population. In the current experiment, the agents themselves are not aware of this fact. Because each agent uses a particular word order for expressing formulas constructed with each commutative Boolean function. But the external language spoken by the population shows that they perfectly understand each other in spite of using different word orders for the expressions associated with the arguments of commutative Boolean functions.

With respect to the four criteria proposed by [4], the agents in the population have some degree of *diversity of opinion*, in the sense that they can invent different words for referring to logical categories and order the constituents of sentences in different ways. But they basically use the same approach for constructing logical categories and for expressing logical formulas, i.e. they all use word order as the only syntactic mechanism for disambiguation and non-recursive formulas of one or two arguments for discrimination. They are *independent* of each other, because each agent chooses the words and grammatical constructions it uses for communication taking into account only the scores of the rules in its own grammar. The scores of such rules depend on the interaction history of the agent, which is always different from the interaction history of the others, thus providing each individual agent with a *different perspective* of the language used by the population. The *aggregation mechanism* used in the experiments is, as we explained above, the shared language system that emerges as a result of the self-organisation process of the interactions that take place among the agents in the population. The mechanisms the agents use to adapt their languages to the expressions they observe are used more often by other agents favour such self-organisation process, because each agent tries to use the same expressions as the others.

The necessity of designing methods for describing how a group thinks as a whole, pointed out by [5], is addressed in multi-agent experiments studying language emergence and evolution using a number of functions which evaluate the performance of the group as a whole. In the present experiment we have used *communicative success*, but other functions which compute the similarity of the agents' grammars, the discriminating capacity of the set of categories constructed by the population, or the complexity of the vocabulary and gram-

matical constructions of the language spoken by the population, can be used as well [12].

Finally, we think that the agent interaction mechanisms used to construct compatible conceptualisations and a shared language system in the experiment analysed in this paper might be appropriately adapted and applied to crowd-based socio-cognitive systems [3] addressing issues such as crowdsourced crisis mapping [7], the construction of a democratic political culture [9] or the generalisation of access to education [10]. Because in each of these domains sets of new concepts and linguistic constructions need to be constructed in order to accurately reflect the reality their users are dealing with and to enable communication, and the best way of constructing such new language systems is using mechanisms that enable meaning coordination.

References

1. Jahedpari, F., Padget, J., De Vos, M., Hirsch, B.: Artificial prediction markets as a tool for syndromic surveillance. In: Crowd Intelligence: Foundations, Methods and Practices. (2014)
2. Buckley, P.: A research agenda for prediction markets. In: Crowd Intelligence: Foundations, Methods and Practices. (2014)
3. Noriega, P., D'Inverno, M.: Crowd-based socio-cognitive systems. In: Crowd Intelligence: Foundations, Methods and Practices. (2014)
4. Surowiecki, J.: The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business. Economies, Societies and Nations (2004)
5. Oinas-Kukkonen, H.: Network analysis and crowds of people as sources of new organisational knowledge. In: Knowledge Management: Theoretical Foundation. Informing Science Press (2008) 173–189
6. Por, G.: Blog of collective intelligence (2013)
7. Poblet, M., Casanovas, P.: Crowdsourced crisis mapping: how it works and why it matters. The Conversation (http://theconversation.edu.au) (2012)
8. Poblet, M., García-Cuesta, E., Casanovas, P.: It enabled crowds: Leveraging the geomobile revolution. In: Crowd Intelligence: Foundations, Methods and Practices. (2014)
9. Nesh-Nash, T., Deely, S.: The future of democratic participation: Online collaboration to build a democratic political culture. In: Crowd Intelligence: Foundations, Methods and Practices. (2014)
10. D'Inverno, M., Yee-King, M.: Pedagogical agents for social music learning in crowd-based socio-cognitive systems. In: Crowd Intelligence: Foundations, Methods and Practices. (2014)
11. Sierra, J., Santibáñez, J.: Experiments on the acquisition of cognitive and linguistic competence to communicate propositional logic sentences. In: Papers from the AAAI Fall Symposium (FS-09-01). (2009) 153–158
12. Steels, L.: Modeling the cultural evolution of language. Physics of Life Review **8** (2011) 339–356
13. Vogt, P.: The emergence of compositional structures in perceptually grounded language games. Artificial Intelligence **167(1-2)** (2005) 206–242
14. Kirby, S.: Learning, bottlenecks and the evolution of recursive syntax. In: Linguistic Evolution through Language Acquisition: Formal and Computational Models, Cambridge University Press (2002) 96–109

15. Sierra, J., Santibáñez, J.: The acquisition of linguistic competence for communicating propositional logic sentences. In: Engineering Societies in the Agents World VIII, Lecture Notes in Computer Science, volume 4955 (2008) 175–191

A The Evaluation Game

The emergence of a shared language system of logical constructions in the population results from a process of self-organisation of the linguistic interactions that take place among the agents in the population. The particular type of linguistic interaction used in the experiment discussed in this paper is called *the evaluation game*. It is played by two agents, a speaker and a hearer, and its main steps can be summarised as follows.

1. Conceptualisation Firstly both agents, speaker and hearer, are given a description of a set of objects which constitute *the context* of the evaluation game. Then the speaker picks a subset of objects from the context which will be *the topic* of the evaluation game. The rest of the objects in the context are called *the background*.

The speaker tries to construct a *conceptualisation* of the topic, that is, a logical formula which is true for all the objects in the topic and false for all the objects in the background. It does so by finding a unary or binary tuple of categories such that its evaluation on the topic is different from its evaluation on any object in the background. Once it has found a discriminating category tuple, the speaker tries to find a logical category which is associated with the subset of Boolean values or pairs of Boolean values resulting from evaluating the topic on that category tuple, and constructs a conceptualisation of the topic applying this logical category to the discriminating category tuple.

In general an agent can build several conceptualisations for the same topic. For example, if the context contains objects 1, 2 and 3 such that object 1 is up and to the left, object 2 is down and to the left, and object 3 is down and to the right, and the topic consists of objects 1 and 3, then both formulas [iff, up, left] and [xor, up, right] can be used as conceptualisations of the topic.

2. Generation The speaker tries to generate a sentence for each of its conceptualisations of the topic using its lexicon and grammar. It tries to maximise the probability of being understood by other agents by selecting the sentence with the highest score, and communicates that sentence to the hearer. The algorithm for computing the score of a sentence from the scores of the grammar rules used in its generation is explained in [13].

The agents in the population start with an empty lexicon and an empty grammar. Therefore they cannot generate sentences for most formulas (conceptualisations) at the early stages of a simulation run. In order to let language to get off the ground, they are allowed to invent new sentences for those meanings (conceptualisations) they cannot express using their lexicon and grammar. As

the agents play language games they learn associations between expressions and meanings, and induce linguistic knowledge from such associations in the form of grammar rules and lexical entries.

3. Interpretation If the hearer can parse the sentence communicated by the speaker using its lexicon and grammar, it extracts a formula (a meaning) and uses that formula to identify the topic. At the early stages of a simulation run the hearers usually cannot parse the sentences communicated by the speakers, since they have no prior linguistic knowledge. In this case the speaker points to the topic, and the hearer adopts an association between its conceptualisation of the topic and the sentence used by the speaker. Note that the conceptualisations of speaker and hearer might be different, because different formulas can be used to conceptualise the same topic.

4. Adaptation The evaluation game is successful if the hearer can parse the sentence communicated by the speaker, and its interpretation of that sentence identifies the topic (i.e. the subset of objects the speaker had in mind) correctly. Depending on the outcome of the evaluation game, speaker and hearer take different actions. We have explained some of them already (*invention* and *adoption*), but they also *adapt their grammars* to communicate more successfully in future games.

Coordination of the agents' grammars is necessary, because different agents can invent different words to refer to the same categories, and because the invention process uses a random order to concatenate the expressions associated with the components of a given formula. In order to understand each other, the agents must use a common vocabulary and must order the constituents of sentences in sufficiently similar ways as to avoid ambiguous interpretations. The following *adaptation mechanisms* are used to coordinate the agents' grammars.

We consider the case in which the speaker can generate a sentence and the hearer can parse it. If the speaker can generate several sentences for its conceptualisations of the topic, the sentence with the highest score is chosen for communication and the rest of the sentences are kept as *competing sentences*. Similarly if the hearer can obtain several formulas (meanings) for the sentence communicated by the speaker, the formula with the highest score is selected as its interpretation of the sentence and the rest of the formulas are kept as *competing meanings*.

If the topic identified by the hearer is the subset of objects the speaker had in mind, the evaluation game succeeds. The speaker increases the scores of the grammar rules it used for generating the sentence communicated to the hearer and decreases the scores of the grammar rules it used for generating competing sentences. The hearer increases the scores of the grammar rules it used for obtaining its interpretation of the sentence and decreases the scores of the rules it used for obtaining competing meanings. This way the grammar rules which have been used successfully get reinforced, and the grammar rules which have been used for generating competing sentences or competing meanings are inhibited.

If the topic identified by the hearer is different from the subset of objects the speaker had in mind, the evaluation game fails and both agents decrease the scores of the grammar rules they used for generating and interpreting the sentence used by the speaker respectively. This way the grammar rules used without success are inhibited.

The scores of grammar rules are *updated* replacing the rule's original score S with the result of evaluating expression 1 if the score is *increased*, and with the result of evaluating expression 2 if the score is *decreased*.

$$minimum(1,\ S + 0.1) \tag{1}$$

$$maximum(0,\ S - 0.1) \tag{2}$$

B Induction

Besides inventing expressions and adopting associations between sentences and meanings, the agents use some *induction mechanisms* to extract generalisations from the grammar rules they have learnt so far. The induction mechanisms used in this paper are based on the rules of *simplification and chunk* in [14], although we have extended them so that they can be applied to grammar rules which have scores attached to them [13]. The induction rules are applied whenever the agents invent or adopt a new association to avoid redundancy and increase generality in their grammars.

Instead of giving a formal definition of the induction rules used in the experiment, which can be found in [15], we give an example of their application. We use *Definite Clause Grammar* to represent the internal grammars constructed by the individual agents. *Non-terminals* have two arguments attached to them. The first argument conveys semantic information and the second is a *score* in the interval [0, 1] which estimates the usefulness of the grammar rule in previous communication. Suppose an agent's grammar contains the following rules.

$$s(\text{light}, S) \rightarrow \text{clair}, \{S\ is\ 0.70\} \tag{3}$$

$$s(\text{right}, S) \rightarrow \text{droit}, \{S\ is\ 0.25\} \tag{4}$$

$$s([and, light, right], S) \rightarrow etclairdroit, \{S\ is\ 0.01\} \tag{5}$$

$$s([or, light, right], S) \rightarrow ouclairdroit, \{S\ is\ 0.01\} \tag{6}$$

The induction rule of **simplification**, applied to 5 and 4, allows generalising grammar rule 5 replacing it with 7. In this case *simplification* assumes that the second argument of logical category *'and'* can be any meaning that can be expressed by a 'sentence', because according to rule 4 the syntactic category of expression 'droit' is s (sentence).

$$s([and,light,B], S) \rightarrow etclair, s(B,R), \{S\ is\ R{\cdot}0.01\} \tag{7}$$

Simplification, applied to rules 7 and 3, can be used to generalise rule 7 replacing it with 8. Rule 6 can be generalised as well replacing it with rule 9.

$$s([and,A,B], S) \rightarrow et, s(A,Q), s(B,R), \{S\ is\ Q{\cdot}R{\cdot}0.01\} \tag{8}$$

$$s([or,A,B],S) \rightarrow ou, s(A,Q), s(B,R), \{S\ is\ Q{\cdot}R{\cdot}0.01\} \quad (9)$$

Induction rule **chunk I** replaces a pair of grammar rules such as 8 and 9 with a single rule 10 which is more general, because it makes abstraction of their common structure introducing a syntactic category $c2$ for binary connectives. Rules 11 and 12 state that the expressions *et* and *ou* belong to syntactic category $c2$.

$$s([C,A,B],S) \rightarrow c2(C,P), s(A,Q), s(B,R), \{S\ is\ P{\cdot}Q{\cdot}R{\cdot}0.01\} \quad (10)$$

$$c2(\text{and},S) \rightarrow \text{et}, \{S\ is\ 0.01\} \quad (11)$$

$$c2(\text{or},S) \rightarrow \text{ou}, \{S\ is\ 0.01\} \quad (12)$$

Suppose the agent of previous examples adopts or invents the following rule.

$$s([\text{if}, le, up],S) \rightarrow \text{siclairdroit}, \{S\ is\ 0.1\} \quad (13)$$

Simplification of rule 13 with rules 3 and 4 would replace rule 13 with 14.

$$s([\text{if}, Q, R],S) \rightarrow \text{si}, s(Q,SQ), s(R,SR), \{S\ is\ SQ{\cdot}SR{\cdot}0.1\} \quad (14)$$

Then induction rule **chunk II**, applied to 14 and 10, would replace rule 14 with rule 15.

$$c2(\text{if},S) \rightarrow \text{si}, \{S\ is\ 0.1\} \quad (15)$$

Pedagogical agents for social music learning in Crowd-based Socio-Cognitive Systems

Matthew Yee-King and Mark d'Inverno

Goldsmiths, University of London, UK
dinverno@gold.ac.uk

Abstract. This paper considers some of the issues involved in building a crowd-based system for learning music socially in communities. The effective implementation of building such systems provides several fascinating challenges if they are to be sufficiently flexible and personal for effective social learning to take place when they are large number of users. Based on our experiences of building the infrastructure for a crowd-based music learning system in Goldsmiths called MusicCircle we address several some of the challenges using an agent based approach, employing formal specifications to articulate the agent design which can later be used for software development. The challenges addressed are: 1) How can a learner be provided with a personalised learning experience? 2) How can a learner make best use of the heterogenous community of humans and agents who co-habit the virtual learning environment? We present formal specifications for an open learner model, a learning environment, learning plans and a personal learning agent. The open learner model represents the learner as having current and desired skills and knowledge and past and present learning plans. The learning environment is an online platform affording learning tasks which can be carried out by individuals or communities of users and agents. Tasks are connected together into learning plans, with pre and post conditions. We demonstrate how the personal learning agent can find learning plans and propose social connections for its user within a system which affords a dynamic set of learning plans and a range of human/ agent social relationships, such as learner-teacher, learner-learner and producer-commentator.

1 Introduction

2012 has been called the 'year of the MOOC', the massive, open, online course [13]. Indeed one of the authors of this paper was part of a team which delivered a course to an enrolled student body of around 100,000 in 2013. The obvious problem with MOOCs is that there is a very high student to tutor ratio. This means it is not feasible to provide students with direct tutor support when they have problems with their learning and complex assessments which cannot be automated become impractical. The current solutions seem to be the use of forums and other social media wherein peer support can take place, and the use of peer assessment techniques such as calibrated peer assessment [7]. Running our MOOC, we noticed that the forum seemed to be an inefficient tool through which students could find information, where the same questions would be asked and answered repeatedly, and where the constant churn pushed old answers

away[1]. It was not clear who would bother to answer a given question, or who would be the ideal person to answer it. Regarding the assessment, there was a tendency to assess others' work superficially - to simply fulfill the most basic requirements of the peer assessment task. This was probably an instance of strategic learning, where the learner does the minimum to meet the apparent requirements. Another problem is a high drop out rate on courses. For example, we had around 10% of our 100.000 students still active at the end of our MOOC; Norvig and Thrun's famous Stanford AI CS211 course in 2011 went from 160,000 enrolments to 20,000 completions [14]. These figures improve if we instead consider the number of students actively accessing learning materials at the start of the course; in our case, 100.000 becomes 36,000. So motivation to complete the course is another area that needs work.

But how might one motivate a learner, given the particular characteristics of a MOOC, i.e. the high learner to teacher ratio, the presence of a large, heterogeneous peer group, the distance learning component and so on? Might motivation be amplified by leveraging the learner's peers - the social network? What might a 'networked learner' gain from being part of an active learning community? How can the learner be made aware of the structure and members of the community, and how that might help them achieve their learning goals?

In summary, guidance for learners, feedback to learners (on their work) and general learner motivation are areas for improvement for MOOCs. These are the key points we aim to address in our wider research work. In this paper we present our work on a representative component of this: the invention of a type of pedagogical agent called a *personal learning agent* which can provide a more intuitive and efficient route through the learning materials and information, which can hep the learner to network to find help or to provide help and feedback to others.

1.1 Pedagogical agents

Skiar et al. present a review of research where agents are used [15] and we propose the reader look at this for more details than we are able to present here. According to Soliman and Guetl, Intelligent Pedagogical Agents (IPAs) are agents which help learners by providing narrations and guidance in order to resolve difficulties and improve motivation' [16]. Magus et al. describe a math tutoring game which includes a conversational agent [11] which has some similar agent characteristics.

Animated pedagogical agents (APAs) operating in realtime virtual learning environments allow learners to request information that helps build an understanding of a student's thought processes and methods of knowledge acquisition [6]. Lester et al. trialled a 3D animated character with 100 middle school children. They discuss the *persona effect*, which encompasses the agent's encouragement (of learners), utility, credibility, and clarity, and which is much enhanced by the use of an animated character [8]. Johnson et al. provide a list of technical issues for designers of animated pedagogical agents to consider [5] and the interested reader is recommended to look at this paper for more details.

[1] this is somewhat alleviated by up-and down-voting of questions and answers but this is far from perfect

Xiao et al. empirically assessed the effect of pedagogical agent competency where learners were learning how to use a text editor supported by pedagogical agents with varying competency at the task [18]. Baylor et al. present an initial study where agents take on different roles (as in Electronic Institutions [4] when supporting learners: Motivator, Expert, or Mentor. More knowledgeable agents were more credible and seemed to transfer more knowledge but motivating agents were more engaging [1]. In [17], an agent based approach is used to simulate interactions between learners within a group.

2 The Music Circle System

The word presented in this paper is part of the efforts to build an online learning system for massively online learning of music within normative communities of practice. It is funded under the FP7 Technology-Enhanced Learning Program called Practice and Performance Analysis Inspiring Social Education (PRAISE). The first author is the Technical Project manager the project and the second author is the principal investigator. The major research questions of the project are as follows

1. How to evidence increase participation in musical learning activity?
2. Is giving and receiving feedback related to engagement with practice?
3. What is the right level of social coordination?
4. How to evidence musical learning?
5. How can automatic techniques be used to evidence feedback?
6. How to build a personal learning agent for personalised learning experiences?

In order to achieve this one of the key technologies we are developing is that of the *personal learning agents* that can enable pathways to and through information for learners, better feedback to learners (on their work) and general learner motivation, and as a resource for finding fellow students within which we can learnt together. In this paper, we will address the following questions which fall within this wider remit:

1. How might one formally specify a human learner that enables an autonomous personal learning agent to represent them in a crowd-based system for music learning?
2. What kind of operations might be useful, given the wider research goals articulated above?

The specific online system we are developing within the PRAISE project is called Music Circle and next we describe Music Circle and why it meets all the criteria for a Crowd-based Socio-technical System.

3 Music Circle as an example of a Crowd-based Socio-technical System

Before moving onto specify the personal learning architecture we first wish to state why this system fits the definition of a crowd-based socio-technical system that is described in a sister paper in this workshop by Pablo Noriega and Mark d'Inverno entitled "Crowd-Based Socio-Cognitive Systems" [12]. In this section we take their description of a *Crowd-Based Socio-cognitive System* and item by item explain why our Music Circle system matches the description given by Noriega and d'Inverno.

- Dimension 1. The Music Circle system contains *agents* which are both computational or human and are autonomous in that they can exhibit purposeful behaviour.
- Dimension 2. The population is a *mix* of human and software agents.
- Dimension 3. The human and computational agents have a *model* of the system in which they operate.
- Dimension 4. The agents within the Music Circle system are *rational* in that they are capable of choosing different courses of action based on their own models (however simple or complex these may be).
- Dimension 5. The human and computational agents in Music Circle are *social* in that they interact with other agents.
- Dimension 6. The human and computational agents in Music Circle have *social models* of the other agents in the system.
- Dimension 7. The human and computation agents are *socio-cognitive* in the sense that they base their decisions on some decision-making process which takes into account the models of the other agents in the Music Circle system.
- Dimension 8. The agents in Music Circle have *social capabilities* including awareness and models of others, and the ability to understand the norms of the learning community in which they are situated.
- Dimension 9. The Music Circle system is defined by the system of interacting agents which means that the state of the system at any stage can never been known by us as the designers and engineers of the systems (opacity).
- Dimension 10. Agents may enter and leave our system at any time and cannot be known in advance.
- Dimension 11. Music Circle communities are *regulated* in order to enable the social coordination of music learning in communities.
- Dimension 12. The human agents are *autonomous* and may not necessarily want to provide helpful feedback to the other music learners in the community and so we need to regulate communities in order to be sensitive to how such agents could destroy the trust within a learning community.
- Dimension 13. A Music Circle system is dialogical as all interactions are mediated by technological artefacts and may therefore be wrapped as communicative acts or messages typically in relation to audio media.
- Dimension 14. We currently have several hundred of students enrolled with the hope and expectation of getting this to thousands of music learners within the next 12 months
- Dimension 15. The system is being designed specifically so that norms can be determined by the community. Moreover, we are specifically including models of how trust and reputation arise and can be managed within such systems.
- Dimension 16. We are developing the Personal learning agent, and the Music Circle system in general (including norms, coordination, rest and reputation managements systems and so on) so that it will help the human users take a wide-variety of feedback on audio and suggested plans for practice from different agents in order to synthesise a particular plan of practice activity.

That is to say that that our system is central example of a *Crowd-based Socio-Cognitive System* and we are using norm-based MAS techniques in the specification,

design and implementation of our system. Next, we move to the specification of the personal learning agent which works with the human learning agent to facilitate music learning in our PRAISE system.

4 Requirements of the Personal Learning Agent

There are several issues about large online systems including (i) motivating the learner (high drop out rate, personalised learning pathways) (ii) connecting the learner (who can help me with this? and (iii) who is having the same problems, etc.) and giving the learner an individual pathway (how can I learn to do this?) One approach to combat these issues is to define a Personal Learning Agent we will use the specification techniques developed by Luck and d'Inverno over the last 20 years or so (e.g. [2, 9, 10]).

We will begin by framing the agent specification presented later with some requirements for the functionality of the agent. There are 4 key requirements: to store learner state, to report learner state, to find learning plans and to propose social connections. Each of these requirements has sub-requirements, as listed below:

1. Storing learner state:
   (a) Storing the goals of a person
   (b) Interpret the goals into desired skills and knowledge
   (c) Storing a person's current skills and knowledge
   (d) Storing a person's current and previous plans
2. Reporting learner state:
   (a) Reporting current state of goals and plans
   (b) Reporting current state of knowledge and skills
   (c) Reporting status of data/ content provided to and from the community i.e. plans, feedback, feedback agents, trust model
3. Plan finding:
   (a) Propose plans whose pre-conditions match current skills and knowledge
   (b) Propose plans whose post-conditions (goals) match a personal learning agent's goals
   (c) Generate evaluation data for plans based on users
   (d) Propose plans which are successful, i.e. verified post conditions
4. Agent finding:
   (a) Propose social relationships/ connections to people with similar goals/ skills/ knowledge (potential peers, potential as they must actively agree to connect to make a social relationship)
   (b) Propose connections to people with similar (musical/ geographical/ etc.) data
   (c) Propose connections to people who have related but superior skills and knowledge (potential tutors), or teaching goals. (I want to increase others' knowledge of scales on the guitar). These people might be able to assign plans, for example.

5 Formal specification of the Personal Learning Agent

In this section we will use the specification language Z to develop the models of our agents, following the methodology developed by Luck and d'Inverno [2, 9, 3].

Learner model

We begin our description by introducing our learner model. The purpose of the learner model is to represent various aspects of a person operating within our learning environment. There are two *types* which users of the system might want to learn about or teach about. The specification remains neutral about how they are encoded but this encoding might include free text descriptions or formulas in predicate calculous for example.

$$[Skill, Knowledge]$$

As an example a user might have the skill of playing the C major scale and the knowledge which includes being able to state which notes are in the scale of C major.

We then define *Proficiency* as the combination of skills and knowledge, representing all that a person would potentially wish to learn in music.

$$Proficiency ::= skills\langle\langle Skill\rangle\rangle \mid knowledge\langle\langle Knowledge\rangle\rangle$$

A particular person can be given a score which is an evaluation of their learning level regarding a particular skill or knowledge element:

$$Score == \mathbb{N}$$

Learning environment

We continue the description with some details about the learning environment which learners, teachers and agents will inhabit. For the purposes of our wider research, it is specialised for music education, and it is designed around a social, blended learning pedagogy wherein people upload recordings of themselves playing instruments and other media items. Discussion and feedback can occur around the uploaded items. Within the environment, people and agents can carry out tasks, where a task is something to be undertaken.

$$[Task]$$

We have identified 9 distinct tasks which can be carried out within our learning environment.

$$\begin{aligned} TaskType ::= {} & Practice \mid Listen \mid Makemusic \mid \\ & Upload \mid Share \mid Annotate \mid \\ & Question \mid Answer \mid Visualise \end{aligned}$$

Earlier, we mentioned that feedback might be provided about a media item. For the time being we define feedback as a given set. It is possible to define feedback in terms of constructive and evaluative praise and criticism. However, these are our first attempts at defining feedback we will remain neutral for the time being.

$$[Feedback]$$

We define evaluate to be a function which maps an proficiency to a natural number, e.g. I have evaluated the way you have played C major as scoring a 5.

$$evaluateproficiency : Proficiency \rightarrow \mathbb{N}$$

In the system the community may evaluate many different aspects. One of those is to evaluate feedback for example.

$$evaluatefeedback : Feedback \rightarrow \mathbb{N}$$

Goals, Beliefs and Plans

As with the definition of the SMART Agent Framework [2] we take a goal to be a state of affairs in the world that is to be achieved (by some agent).

$$[Goal]$$

The way that goals (or, equally, learning outcomes) are achieved is through a workflow of tasks: a sequence of tasks that have to be completed in order. We do not specify here who determines whether the tasks have been accomplished successfully or not because in general this could be a mixture of the system, the user themselves, the community and/or a teacher. Plans are typically specified in terms of what must be true before they can be adopted, what is true after they have been successfully completed, and the kinds of actions (or in our language tasks) that have to be completed in order Next we define a plan to be a set of preconditions (the skills and knowledge and agent must have before undertaking the plan) and a set of post conditions (which describe the new set of skills and knowledge the agent will have after the plan). The predicate part of the schema state that the intersection of the pre and post conditions are necessarily empty.

Plan

$$pre : \mathbb{P}\, Proficiency$$
$$post : \mathbb{P}\, Proficiency$$
$$workflow : \text{seq}\, Task$$

$$pre \cap post = \{\}$$

In specifying this system, it is useful to be able to assert that an element is optional. The following definitions provide for a new type, $optional[T]$, for any existing type, T, which consists of the empty set and singleton sets containing elements of T. The predicates, *defined* and *undefined* test whether an element of $optional[T]$ is defined (i.e. contains an element of type T) or not (i.e. is the empty set), and the function, *the*, extracts the element from a defined member of $optional[T]$.

$$optional[X] == \{xs : \mathbb{P}\, X \mid \# xs \leq 1\}$$

$$
\begin{array}{|l}
\hline
[X] \\
\textit{defined}\_, \textit{undefined}\_ : \mathbb{P}(\textit{optional}[X]) \\
\textit{the} : \textit{optional}[X] \pfun X \\
\hline
\forall\, \textit{xs} : \textit{optional}[X] \bullet \\
\quad \textit{defined xs} \Leftrightarrow \# \textit{xs} = 1 \land \\
\quad \textit{undefined xs} \Leftrightarrow \# \textit{xs} = 0 \\
\forall\, \textit{xs} : \textit{optional}[X] \mid \textit{defined xs} \bullet \\
\quad \textit{the xs} = (\mu\, x : X \mid x \in \textit{xs}) \\
\hline
\end{array}
$$

$$\textit{Bool} ::= \textit{True} \mid \textit{False}$$

Using this definition we can now specify the state of a plan. The state of a plan can be thought of as a running instance of a plan during the lifetime of a users activity. It means that the plan has been adopted to achieve a goal. In order to specify this we keep the information contained in the specification of a Plan using schema inclusion. We also state that if the plan has been started but not finished there will be a *current task* that the agent is currently undergoing. By also defining a flag called *finished* we can specify a plan state as follows. The predicate part states that the current task must have been defined in the workflow of the plan.

$$
\begin{array}{|l}
\hline
\textit{PlanInstance} \\
\textit{Plan} \\
\textit{current} : \textit{optional}[\textit{Task}] \\
\textit{finished} : \textit{Bool} \\
\hline
\textit{thecurrent} \in (\operatorname{ran} \textit{workflow}) \\
\hline
\end{array}
$$

The initial plan state (for any state schema the initial state should be specified in Z) is where the plan has just been proposed or adopted by a user.

$$
\begin{array}{|l}
\hline
\textit{InitialPlanInstance} \\
\textit{PlanInstance} \\
\hline
\textit{undefined current} \\
\textit{finished} = \textit{False} \\
\hline
\end{array}
$$

We are now in a position to define four specific sub-types of the plan state as follows.

1. Proposed Plan. A plan which has been selected to achieve a goal but which has not been started by the agent. As no task has been started the current task is set to undefined.

$$
\begin{array}{|l}
\hline
\textit{ProposedPlan} \\
\textit{InitialPlanInstance} \\
\hline
\end{array}
$$

2. Active Plan. A plan which is ongoing. It has not been completed and the current task is set to defined.

$$
\begin{array}{|l}
\hline
\textit{ActivePlan} \\
\textit{PlanInstance} \\
\hline
\textit{defined current} \\
\textit{finished} = \textit{False} \\
\hline
\end{array}
$$

3. FailedPlan. This is a plan which has a defined task but a flag set to finished. For example, this represents a situation one of the tasks in the workflow of a plan is too difficult for the user and the plan is discarded by the user.

$$
\begin{array}{|l}
\hline
\textit{FailedPlan} \\
\textit{PlanInstance} \\
\hline
\textit{definedcurrent} \\
\textit{finished} = \textit{True} \\
\hline
\end{array}
$$

4. Completed Plan. The flag *finished* is set to true and the current task becomes undefined.

$$
\begin{array}{|l}
\hline
\textit{CompletedPlan} \\
\textit{PlanInstance} \\
\hline
\textit{undefined current} \\
\textit{finished} = \textit{True} \\
\hline
\end{array}
$$

There are several operations that we could specify at the level of the plan but the key one is finish task. Either this leads to the plan being completed or the current place in the work flow moves to the next task.

In the first case the specification looks like this:

$$
\begin{array}{|l}
\hline
\textit{FinishTask}1 \\
\Delta \textit{PlanInstance} \\
\hline
\textit{current} = \{\textit{last}(\textit{workflow})\} \\
\textit{finished} = \textit{False} \\
\textit{undefined current}' \\
\textit{finished}' = \textit{False} \\
\hline
\end{array}
$$

In the second case like this:

$$
\begin{array}{|l}
\hline
\textit{FinishTask}2 \\
\Delta \textit{PlanInstance} \\
\hline
\textit{current} \neq \{\textit{last}(\textit{workflow})\} \\
\textit{finished} = \textit{False} \\
\textit{current}' = \{\textit{workflow}((\textit{workflow}^{\sim}(\textit{the current})) + 1)\} \\
\textit{finished}' = \textit{False} \\
\hline
\end{array}
$$

The other is to instantiate a plan which essentially means creating a PlanInstance in it is initial state from a Plan.

$$
\begin{array}{|l}
instantiateplan : Plan \rightarrow InitialPlanInstance \\
\hline
\forall p : Plan;\ ps : InitialPlanInstance \\
\quad \mid ps = instantiateplan(p) \bullet \\
\quad ps.pre = p.pre \land ps.post \\
\quad = p.post \land ps.workflow = p.workflow
\end{array}
$$

The (almost) inverse function of this is a function which takes any PlanInstance and returns the plan.

$$
\begin{array}{|l}
recoverplan : PlanInstance \rightarrow Plan \\
\hline
\forall p : Plan;\ ps : PlanInstance \mid p = recoverplan(ps) \bullet \\
\quad ps.pre = p.pre \land \\
\quad ps.post = p.post \land ps.workflow = p.workflow
\end{array}
$$

Beliefs

This is a representation of what the agent knows and what it can do. Again we remain neutral on the representation.

$$[Belief]$$

The Personal Learning Agent In the schema below we have the following definitions.

1. An agent has a set of goals at any stage which we call desires (typically these are associated with learning outcomes as described earlier in the document.)
2. An agent has a set of beliefs. These refer to the information which is stored about what the user knows or what the user can do (skills).
3. An agent has some interpret function which takes a goal and returns a set of proficiency (skills and knowledge). Note that the complexity of this function may vary as in some cases goals may be expressed as a set of proficiency directly and so this function becomes a simple identity function. However, in other situations this function has to take a free text description and turn it into a set proficiency. Clearly, in general no automatic process can do this and such an operation will often be left to the community. In which case we specify the agents interpret function as a *partial* function.
4. An agent has a similar interpret function for beliefs which maps its beliefs to a set of machine readable (skills and knowledge).
5. *intdesires* is a set of proficiencies which can then be used by the agent and the community to plan. Note then, that *interpreteddesires* is made up of the automatic function *interpret* of the agent, possibly the automatic interpretation of other agents, but also from human users in the music learning community.
6. *intbeliefs* is the analagous set of proficiencies which the agent has recorded as known or accompished by the agent.

7. It is not unreasonable to suggest that all tasks are not available to a user at all times and so the agent can record which tasks are currently available to a user. (If the internet is down, upload is not an available task. If a newcomer joins a community then possibly they do not feel like giving any feedback and so the agent can record that the user is currently not offering this task.).
8. Then we define the set of plans which the agent knows about (possibility learned from other agents). This is where the agent contains its *procedural knowledge* about what plans work in what situations to achieve which desired proficiency.
9. The agent maintains a record of all of the plans that have been completed and all of those which have failed.
10. There is a record of the intentions. This is a mapping from a set of proficiencies (this set may only have one proficiency in it of course) to the plan instance which the agent has adopted to attain those proficiencies.
11. Finally, we record all those interpreted desires for which the agent has no active plan.

    There are also two dummy variables that we can use (which can be calculated from the variables described so far but which aid us in the readability of the specification)
12. We define a variable record the tasks that the agent is current involved in (*currenttasks*) which can be calculated as the union of the tasks from the current plans.
13. We define a variable to records the current plan instances of the agent

Next we consider the constraints on the state of a personal learning agent

1. The interpreted desires are the result of applying the interpret desire function to the desires.
2. The interpreted beliefs are the result of applying the interpret desire function to the beliefs.
3. The intersection between interpreted desires and interpreted beliefs is an empty set, (in other words you can't desire a proficiency you already have).
4. If there is a plan for a subset of proficiencies then those proficiencies must be contained in the the interpreted desires.
5. If there is a plan for one subset of proficiencies and a plan for another distinct set pif proficiencies then their intersection is empty.
6. The unplanned desires are those interpreted desires for which there is no intention
7. The current tasks are calculated from taken the current plans and the current task from each.
8. The current plans are calculated from taking the range of the intentions.

$[X, Y]$

$$map : (X \to Y) \to (\operatorname{seq} X) \to (\operatorname{seq} Y)$$
$$mapset : (X \to Y) \to (\mathbb{P} X) \to (\mathbb{P} Y)$$

$$\forall f : X \to Y;\ x : X;\ xs, ys : \operatorname{seq} X \bullet$$
$$map\ f\ \langle\rangle = \langle\rangle \land$$
$$map\ f\ \langle x \rangle = \langle f\ x \rangle \land$$
$$map\ f\ (xs \frown ys) = map\ f\ xs \frown map\ f\ ys$$
$$\forall f : X \to Y;\ xs : \mathbb{P} X \bullet$$
$$mapset\ f\ xs = \{x : xs \bullet f\ x\}$$

PersonalLearningAgent

$$desires : \mathbb{P}\ Goal$$
$$beliefs : \mathbb{P}\ Belief$$
$$interpretdes : Goal \pfun \mathbb{P}\ Proficiency$$
$$interpretbel : Belief \pfun \mathbb{P}\ Proficiency$$
$$intdesires : \mathbb{P}\ Proficiency$$
$$intbeliefs : \mathbb{P}\ Proficiency$$
$$availabletasks : \mathbb{P}\ TaskType$$
$$plandatabase : \mathbb{P}\ Plan$$
$$completedplans, failedplans : \mathbb{P}\ Plan$$
$$intentions : (\mathbb{P}\ Proficiency) \pfun PlanInstance$$
$$unplannedintdesires : \mathbb{P}\ Proficiency$$
$$currenttasks : \mathbb{P}\ Task$$
$$currentplaninstances : \mathbb{P}\ PlanInstance$$

$$intdesires = \bigcup(mapset\ interpretdes\ desires)$$
$$intbeliefs = \bigcup(mapset\ interpretbel\ beliefs)$$
$$intdesires \cap intbeliefs = \varnothing$$
$$\bigcup(\operatorname{dom} intentions) \subseteq intdesires$$
$$\forall ps1, ps2 : \mathbb{P}\ Proficiency \mid$$
$$(ps1 \neq ps2) \land (\{ps1, ps2\} \subseteq$$
$$(\operatorname{dom} intentions)) \bullet$$
$$ps1 \cap ps2 = \{\}$$
$$unplannedintdesires = \bigcup(\operatorname{dom} intentions) \setminus intdesires$$
$$currenttasks = \{t : Task;\ ps : PlanInstance \mid$$
$$ps \in (\operatorname{ran} intentions) \bullet the\ ps.current\}$$
$$currentplaninstances = \operatorname{ran} intentions$$

Plan Finding Plan finding is the process of taking a set of candidate plans and selecting those whose preconditions are met and where at least some subset of the postconditions are desired. For this operation we assume the input of a set of candidate plans. Again we do not specify whether these comes from the agent (i.e. the agent's database of plans),

other agents in the community, or from the user or from other users and in general with be a *synthesis* of the users and the agents of users working together.

For now we will suppose that suitable plans have all preconditions satisfied and it is the case that both: (a) none of the postconditions are things which the user is already proficient in (b) all of the postconditions are current interpreted desires of the user. In the schema below *SuitablePlans* is generated which satisfy this constraint and from these one plan *adoptedplan* is selected. The state of the agent is then updated to include that the current plans now includes a mapping from the pre-conditions of the plan (which are necessarily interpreteddesires for which no plan exists.

Plan Completion The very simplest way this could happen is as follows:

1. Because of a successfully completed task a plan instance becomes an element of *CompletedPlan*
2. The post conditions are added to the interpreted beliefs (these may in turn be reverse interpreted into beliefs which can then be seen by the community)
3. Any post conditions that were formerly desires are now removed from interpreted desires (these may in turn be reverse inetreprered into beliefs which can then be seen by the community)
4. The completed plans function is updated with the plan that has just successfully completed.

CompletePlan

$$
\begin{array}{l}
completedplan? : CompletedPlan \\
\Delta PersonalLearningAgent \\
\hline
completedplan? \in (\mathrm{ran}\, intentions) \\
intentions' = intentions \rhd \{completedplan?\} \\
intdesires' = intdesires \setminus completedplan?.post \\
intbeliefs' = intbeliefs \cup completedplan?.post \\
completedplans' = completedplans \cup \\
\qquad \{recoverplan\ completedplan?\}
\end{array}
$$

However, this process will not be automatic in general within the system. In general, the user (or other users in the community) will be asked to evaluate the plan. There may be several ways in which this can happen. For example, a simple score could be given but in general each user who is evaluaring the plan considers each of the post conditions (or another member of the community does) to work out whether they are now proficiencies (intepreted beliefs), whether they have not been met and so are still interpreted desires, or whether they have not been met but are not desires. Indeed the evaluating user could rank each of the postconditions with a score and the agent may also wish to keep a snapshopt of the agent's state for future comparison by the community.

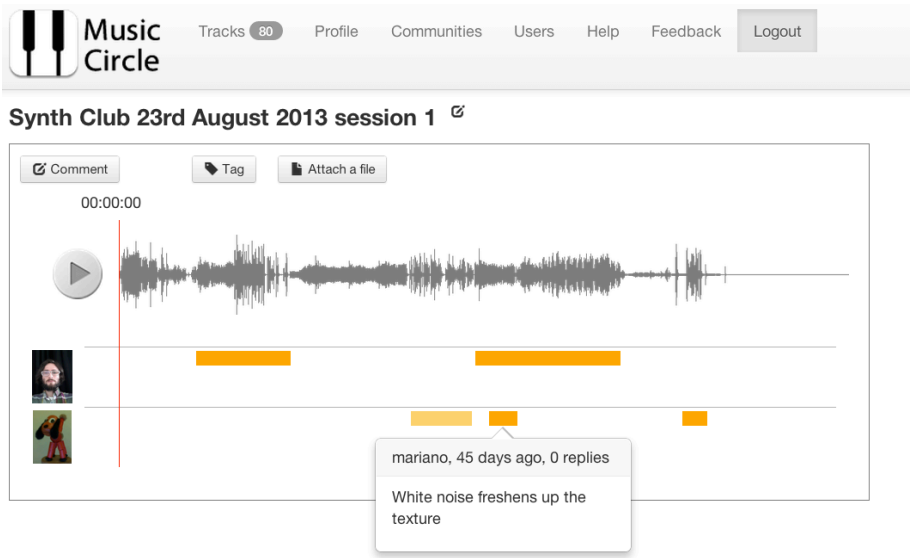

Fig. 1. The music discussion user interface

Finding and adopting a plan Plan finding is the process of taking a set of candidate plans and selecting those whose preconditions are met and where at least some subset of the postconditions are desired. For this operation we assume the input of a set of candidate plans. Again we do not specify whether these comes from the agent (i.e. the agent's database of plans), other agents in the community, or from the user or from other users and in general with be a *synthesis* of the users and the agents of users working together. For now we will suppose that suitable plans have all preconditions satisfied and it is the case that both: (a) none of the postconditions are things which the user is already proficient in (b) all of the postconditions are current interpreted desires of the user. In the schema below *SuitablePlans* is generated which satisfy this constraint and from these one plan *adoptedplan* is selected. The state of the agent is then updated to include that the current plans now includes a mapping from the pre-conditions of the plan (which are necessarily interpreteddesires for which no plan exists.

FindandAdoptPlan

$$PossiblePlans?, SuitablePlans! : \mathbb{P}\, Plan$$
$$adoptedplan : Plan$$
$$\Delta PersonalLearningAgent$$

$$SuitablePlans! = \{ps : PossiblePlans? \mid (ps.pre \subseteq intbeliefs) \wedge (ps.post \cap unplannedintdesires) = \{\} \bullet ps\}$$
$$adoptedplan \in SuitablePlans!$$
$$intentions' = intentions \cup \{(adoptedplan.post, instantiateplan(adoptedplan))\}$$

It would be a simple matter to add more detail to this schema including choosing the plan with the highest rating for example, or a plan which has completed successfully in the community the most number of times, or making sure the plan has not failed in the users history, or that the plan has not failed in the community with users which have similar profiles as defined by the personal learning agent. In general, the plan finding system requirements, and this specification alongside it, will develop as we gain experience of how the system is used.

Community of Music Learning

Agent finding Now we move to defining a community of learners each of which has one and only one personal learning agent. First we define the set of all users.

$$[User]$$

Community

$$community : \mathbb{P}\, User$$
$$agents : User \rightarrowtail PersonalLearningAgent$$

$$community = \mathrm{dom}\, agents$$

To this we can define all kinds of social relationships. For example, peer and teacher and others as they become useful. It is up to the designer of the system to state what the constraints are on any such relationships. To provide examples (not necessarily ones we would subscribe to) of how this is none we state that if user1 is a peer of user2 then user2 is a peer of user1 and, in addition, if user2 is a teacher of user1 then user1 cannot be a teacher of user2. Another example would be the idea of a fan who would always adopt the advice of another.

SocialRelationships

$$peer, teacher : User \leftrightarrow User$$
$$fans : User \leftrightarrow User$$

$$\forall u1, u2 : User \bullet (u1, u2) \in peer \Rightarrow (u2, u1) \in peer$$
$$\forall u1, u2 : User \bullet (u1, u2) \in teacher \Rightarrow (u2, u1) \notin teacher$$

Using these schemas it then becomes possible to ask agents to start to look for users who have similar profiles as stated in the requirements detailed earlier in this document. In order to refine the search to include (for example) looking for agents who have a motivation to teach, we will need to develop the specification to define ways in which agents can broadcast that they are able to teach certain plans. This will come in later versions of this specification.

6 Concluding remarks

We are developing a system for social music learning called MusicCircle that we hope will be populated by large numbers of learners. As part of the design of our system we will be incorporating personal learning agents to provide a more personalised, social and effective learning experience. In this paper have use a standard agent-based formal specification methodology for modelling social agent systems to specify the design of these agents. The website can be found at `musiccircleproject.com` and a screen shot of the system is given above.

At the heart of this is the design of a social system of personal learning agents that can enable users to have a stronger sense of their place in the social community of music learners. We have used a long-standing agent-based specification methodology for building models of these agent systems which we are using in the principled development of our system. In addition to using formal agent-based methodology for designing agent systems and on the other hand we are testing versions of our systems with users across a range of music learning cites in the UK (from school, to pre-consevertoire to the HE sector) so that our work clearly spans agent-based theory and the practice of building systems for large numbers of users. Furthermore, we have demonstrated that Music Circle is an example crowd-based socio-technical system as described by Noriega and d'Inverno in the paper.

In relating the theory and practice of sociological agent systems within the design of socio-technical system more generally also enables us in future work to consider a range of questions about how the scientific social multi-agent approach that our community has developed for 20 years or more can be applied to the analysis and design of crowd-based socio-cognitive systems. We need to understand to what extent a MAS approach to analysing and designing systems such as MusicCircle helps? Could we, for example, start to map out the space of such systems relating technology to sociality in a useful way using the multi-agent tradition? Then could we start to provide platforms and design methodologies for building such systems in the future using a regulated MAS approach?

One hope is that we will see a greater influence from the MAS community, applying the work developed over recent years to become mainstream in the analysis, design and specification of crowd-based socio-technical systems in the future.

Acknowledgements

This research was supported by the FP7 project in the Technology-Enhanced Learning Program called Practice and Performance Analysis Inspiring Social Education (PRAISE). The authors would like to thank Harry Brenton, Andreu Grimalt-Reynes, Maria Krivenski, Jonathan James and Francois Pachet who are colleagues on the PRAISE project at Goldsmiths.

References

1. A. L. Baylor and P. P. A. L. S. R. Group. The impact of three pedagogical agent roles. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, AAMAS '03, pages 928–929, New York, NY, USA, 2003. ACM.
2. M. d'Inverno and M. Luck. *Understanding agent systems*. Springer, 2003.
3. M. d'Inverno and M. Luck. Creativity through autonomy and interaction. *Cognitive Computation*, 4(3):332–346, 2012.
4. M. d'Inverno, M. Luck, P. Noriega, J. A. Rodriguez-Aguilar, and C. Sierra. Communicating open systems. *Artificial Intelligence*, 186(Complete):38–94, 2012.
5. W. L. Johnson, J. W. Rickel, and James C. Lester. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of ...*, pages 47–78, 2000.
6. A. Josephson. New Technology-based Models for Postsecondary Learning: Conceptual Frameworks and Research Agendas. Technical report, Computing Research Association, 2013.
7. D. Koller and A. Ng. The Online Revolution : Education at Scale. Technical report, Stanford University.
8. J. Lester, S. Converse, and S. Kahler. The persona effect: affective impact of animated pedagogical agents. In *CHI 97 Conference on Human Factors in Computing Systems*, Atlanta, 1997.
9. M. Luck and M. d'Inverno. A formal framework for agency and autonomy. *Proceedings of the first international conference on Multi-Agent Systems*, 254260, 1995.
10. M. Luck and M. d'Inverno. Structuring a Z specification to provide a formal framework for autonomous agent systems. *ZUM'95: The Z Formal Specification Notation*, pages 46–62, 1995.
11. H. Magnus, S. Annika, and S. Björn. Building a Social Conversational Pedagogical Agent-Design Challenges and Methodological Approaches. In D. Perez-Marin and I. Pascual-Nieto, editors, *Diana Perez-Marin (Editor), Ismael Pascual-Nieto (Editor)*, pages 128–155. IGI Global, 2010.
12. P. Noriega and M. d'Inverno. Crowd-based socio-cognitive systems. In *Crowd Intelligence: Foundations, Methods and Practices*. European Network for Social Intelligence, Barcelona, January 2104.
13. L. Pappano. The year of the MOOC. *The New York Times*, 2(12):2012, 2012.
14. O. Rodriguez. MOOCs and the AI-Stanford like Courses: two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance, and E-Learning*, 2012.
15. E. Sklar and D. Richards. The use of agents in human learning systems. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, AAMAS '06, pages 767–774, New York, NY, USA, 2006. ACM.
16. M. Soliman and C. Guetl. Intelligent pedagogical agents in immersive virtual learning environments: A review. In *MIPRO, 2010 Proceedings of the 33rd International Convention*. IEEE Computer Society Press, 2010.
17. M. Spoelstra and E. Sklar. Using simulation to model and understand group learning. In *Proc. AAMAS'07 Workshop on Agent Based Systems for Human Learning and Entertainment*, 2007.
18. J. Xiao, J. Stasko, and R. Catrambone. An Empirical Study of the Effect of Agent Competence on User Performance and Perception. In *Third International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 178–185, 2004.

Blind Arbitration

Proposal for Anonymous Crowdsourced Online Arbitration

Dusko Martic

Law, Science and Technology Joint Degree EM program,
IDT-Universitat Autonoma Barcelona, Bellaterra, Barcelona
dusko.martic@gmail.com
dusko.martic@unibo.it

Abstract .

Shifting the academic discourse from general term of ODR to more specific modalities like online arbitration is a clear sign of research advancements in the area of online dispute resolution. In this paper we explore the online arbitration in relation to crowdsourcing trends. The goal of this paper is to present new approach to online arbitration, based on several different priniciples and technologies. The proposal of anonymous/ privacy preserving online arbitration or "blind arbitration" is built upon technologies for online arbitration, crowdsourcing, blind bidding negotiation and founded upon privacy-by-design principles. We aim to propose higher-level of confidentiality, secrecy and privacy preservation along with leveraging "the wisdom of the crowds".

Keywords: Online dispute resolution, ODR, online arbitration, privacy-by-desing, crowdsourced arbitraton, anonymous arbitration, privacy preserving arbitration, blind arbitration, confindentiality,

1 Introduction.

While arbitration, in traditional ADR[1] terms, has been long developed and accepted as a suitable means to resolves conflicts, especially in international business arena, online arbitration[15] has merely embraced online communication as extension of its long established practices with appropriated case management systems and scheduling tools for automation and ease of processes. Full potential of online arbitration, through bigger role of technology still remains to be attained.

Online arbitration compared to traditional ADR approach bears promise of cheaper and efficient proceeding. However, technology also brings new risks to the protection

[1] Alternative dispute resolution

of integrity, confidentiality and security of e-proceedings[15][16]. Almost every arbitration institution guarantees communication security, confidentiality (if agreed), authenticity, and procedural integrity[3]. But even with all safeguards in place, some parties could feel reserved and seek for higher level of confidentiality, privacy or even anonymity.

At the same time we are witnessing the development of new online applications using "crowds" to facilitate faster, cheaper, collective work to respond to the demands of the markets or address some public or private need. Utility of "crowdsourcing"[14] in the field of dispute resolution has only been recently discussed and few initiatives have already emerged. This paper aims to further the discussion by proposing new approach to crowdsourced arbitration which provides with more private or anonymous model of adjudicative dispute resolution.

2 Crowdsourcing in online dispute resolution

The idea of using crowdsourcing for online dispute resolution has started with online juries. Cyberjuries appeared as online version or imitation of traditional juries, where they have similar role in representing values of a community. Marder points to their evolution from opinion polls to online mock juries[9]. Opinion polls model[2] may allow parties to express feeling about dispute without using legal language, anonymously and for free. Anonymity is achieved by representing parties with codes or numbers. Online mock trials are more specialized ODR tools design to give evaluation of the cases, usually used by lawyer for testing their argumentation and strategies before court.

Building on Marder's work, van der Herik and Dimov give more comprehensive overview of crowdsourced online dispute resolution [2] by offering three types: online opinion polls, online mock trials and crowdsourced ODR procedures rendering decisions that are enforced by private authorities. Adapting the ideas of Malone and Dellarocas[8], they also offer four building blocks for crowdsourced online dispute resolution (CODR): The crowd, incentives for motivating the crowd to participate in CODR, types of disputes which can be solved through CODR and CODR procedure.

Van der Herik and Dimov, display EBay's Community Court as most prominent example (up to now) of crowdsourced ODR procedures rendering decisions that are enforced by private authorities. It has been built to deal with negative review disputes between buyers and sellers on EBay. EBay's Community Court[13] had fairly straightforward procedure: after submissions from buyer and seller, a case is put before randomly selected panel of jurors. Jurors were experienced EBay community members fulfilling certain criteria. Upon reviewing submission, a juror needs to decide with

[2] Examples: iCourthouse (www.i-courthouse.com) SideTaker (www.sidetaker.com), People'sCourtRaw(www.peoplescourtraw.com), Truveli (www.truveli.org).

which party he/she agrees or feels that cannot make decision. Each case is reviewed by 21 jurors, on a voluntary basis. Rule and Nagarajan explained the motivation of jurors as a sense of service to the community. This system allowed fast resolution and private enforcement within EBay market.

3 Do we need privacy in online conflict resolution?

Since data about disputes could be sensitive, informational privacy[12] is of highest importance to the ODR system designers. However, as opposed to face-to-face ADR where we can trust that no information is being recorded, online interaction always leaves a trail. It is even more difficult to control or enforce professional and ethical standards of arbitrators in online arbitration. What the online arbitrator is doing with sensitive data and who can physically access his computer are just some of the issues that make parties question the integrity and security of online arbitration.

On the other hand, since dispute involve more than one party, sociologists[1] have identified three phases in the dispute: "naming" (internally recognizing that one has been harmed), "blaming" (confronting the wrongdoer) and "claiming" (pursuing legal remedies). Blaming and claiming are in front of the third party or public. Orna Rabinovich-Einy suggests that the more public a dispute, the less control over the information regarding the dispute the parties have, and, accordingly, the less room there is for secrecy and anonymity[12].

EBay's Community Court was not designed to maintain secrecy, but to effectively engage crowd in scrutinizing inappropriate reviews. Some providers, like iCourthouse, offer anonymity but disclose the facts and submissions of the case, which in the age of Google search could lead to easy discovery of identity. Jurors are also aware of the result of the dispute. In some cases, even the knowledge that dispute exists can lead to bad reputation to a party or a business. Spreading the information about disputes and their trail online can have counter-effect to attracting parties to crowdsourced online dispute resolution.

Hence the question: how can we leverage the wisdom of the crowds, but keep the high level of anonymity of parties and secrecy of the case?

4 "Blind" proposal

We propose "blind arbitration" or privacy preserving arbitration which could be simplified with following formula of principles and technologies that were combined in proposal for the conceptual model:

Blind arbitration = crowdsourced arbitration + arbitration management software + privacy by design[4] + (blind bidding negotiation).

The method for this anonymous crowdsourced arbitration can be described as dividing the integral, bigger text submissions of parties to small questions for arbitrators to answer. The principle idea that is being proposed is anonymous crowdsourced arbitration, through obfuscation of the general picture of dispute by focusing on small tasks or questions. In a way it is putting into practice famous big picture phrase "you can't see the forest for the trees".

In the following figure we illustrate broadly the model:

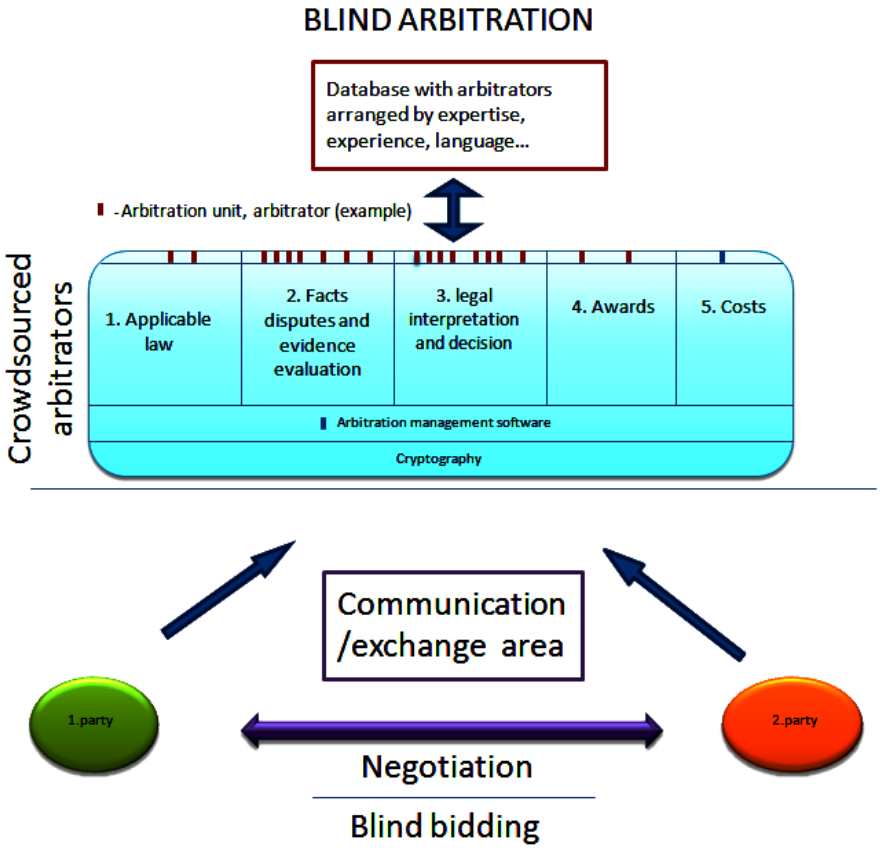

Fig. 1. Illustration of the model of blind arbitration

This proposal assumes that both parties agree to maintain the confidentiality/secrecy of the dispute and that it is in their best interest to keep it private, with as little information to give away about the dispute or about facts related or connected to the parties. Having in mind van der Herik and Dimov's four building blocks for crowdsourced

online dispute resolution[2], we will mostly focus on the fourth block- the CODR procedure. We will give only brief proposals for previous three building blocks.

5 Crowd, incentives and types of disputes in blind arbitration

Arbitrators will be selected by parties or software automatically (if decided by parties or consensus could not be reached) by certain criteria that are most relevant to the dispute, i.e. expertise in certain matter. Preferably, database of arbitrators on a global level should be in place as arbitration will be held online and the location of persons is not relevant.

Most importantly, arbitrators will be informed on need-to-know basis. They will not know who are the parties, what the dispute is about (except the question at their hand), what is the result of the process, are there any other dispute questions posed and the result of those issues. Arbitrators will be chosen to answer only one question and will not be connected to any other issue in the arbitration. This means that the person who deals with question of facts do not know what the dispute is about or what are the legal issues within case, as well as who are other arbitrators. They are all communicating only by online platform ignorant to who are other arbitrators in the database.

We propose incentives to be similar to the model of incentive/payment for microtasks offered at Amazon's Mechanical Turk[3], especially for answering simple questions with right-wrong answers. For more complicated answers with written explanations and reasoning, the payment would correspond to a price previously set by arbitrator (per question).

The range of types of disputes that theoretically could be solved by this method is wide, but for the moment (until fully tested) we would suggest focusing the disputes around single issue. It seems that the most appropriate would be e-commerce disputes or labor/contract disputes about the quality of produced work[4]. However, the model is flexible enough to be extended to more complicated disputes with several issues which could be handled simultaneously or subsequently.

6 The procedure

We will present the concept through three separated phases: confidentially agreement, submission and resolution. The third phase consists of four stages.

---

[3] https://www.mturk.com/mturk/

[4] i.e. disputes about the quality of outsourced work/results from the contract concluded on sites like Elance.com, Guru.com, Odesk.com. One party would claim that the work is not done according to the specifications and the other party disagrees

6.1 Phase 1 – confidentiality agreement.

From the beginning we need parties' agreement to this kind of arbitration, specifying it in more formal way, by arbitration agreement that clearly states the will of parties to be the part of the blind arbitration and the commitment to confidentiality agreement among themselves, with specific terms and high penalties in case of breach. If parties do not agree on these terms, the use of this model of arbitration would be rendered pointless since the other party would not care to keep the opposing party's privacy in check.

6.2 Phase 2 – submission of claims.

After registering for the platform, parties will be faced against each other during dispute with open communication channel through any messaging technology. This way they could directly negotiate and in any moment stop or stay the dispute proceeding, in order to have an agreement putting stop to a dispute.

For the same reason, but less revealing, we suggest simultaneous use of blind bidding technology[7][5][6], especially for the possible monetary disputes or about potential award issues. Blind bidding allows parties to hide their willingness to compromise, but to actively engage in blind offering with hope that the opposing party will be willing to compromise close to their offer. If at any moment offers overlap, the arbitration should halt automatically by software intervention.

The communication area should be also open to exchange of arguments of parties, submission of queries and for discovery purposes, but also to post question and submit issues and questions to arbitrators and to receive and respond on communications from arbitrators. However the communication area should not be revealed to arbitrators.

The conceptual model is founded upon idea that every dispute could be dissolved to its parts and disputed issues. We divide possible disputes in 4 general types: applicable law, disputes about facts (and connected to it-evidence evaluation procedure), interpretation of law disputes (about the rights and duties, interpretation of contract…), dispute about (appropriate) awards.

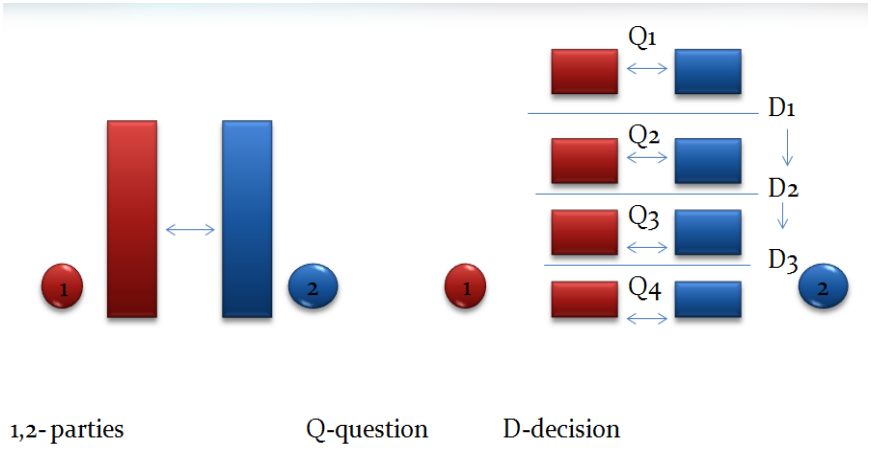

Fig. 2. Dissolution of submissions and argumentation

The electronic (fill-in) form, for posing disputed questions to arbitrators, corresponds to this division and organized in different brackets. When certain aspect is not disputed, it is stated in the bracket for that part of general dispute areas. For example, if applicable law is specified and not disputed by any party, it will be placed in the first bracket and clarified if any other regulation should be consulted.

If facts are not disputed, the recount of relevant facts has to be filled and confirmed by both parties (or just one party gives factual information and other agrees). If one party disputes any part then they are not in agreement and it should be decided by arbitrage. If facts are disputed, the parties will be given opportunity to submit evidence that should be properly anonymised when containing certain information connecting parties to it (if this is not possible because of the nature of the facts, evidence or question we will offer possible solution in later phase). Pictures of persons or any image or information that could identify parties should be scrambled or obfuscated at least.

If a dispute is about legal issues parties should support arguments citing legal references. Parties should be clear about their interpretations and about the legal question that they are submitting to the arbitrage. Arbitrators will not be giving decisions on the whole case at once but on each question separately. The questions, therefore, should be about essential issues and questions should be clear enough.

An award proposal should be clearly stated as well. If rules of procedure (and applicable law considers this lawful) insist that the award can be granted only based on claims of parties, software could by itself recognize and grant an award to a party who wins in dispute, so human involvement could be minimized in this aspect (in last phase).

This phase should be open so both parties can see final statements, claims, arguments of other party as it will be presented to the arbitrators. If they insist on their side of the

story and after at least one party finalizes its claim by pressing the submit option, the software will react by giving reasonable timeframe for other party to finish its argumentation, after which whatever is written in the form will be submitted.

The most important factor of their cooperation will be formulation of their submissions that do not reveal personal information, as previously agreed by their confidentiality agreement. Both parties will clearly state their arguments and ask appropriate question (legal or factual) to arbitrators to resolve a dispute. The whole phase will be structured so the argumentation of the parties could be separated in different brackets, and each bracket will end with specific question for arbitrators (about which there is a misunderstanding). The parties will be advised to formulate (collaboratively) single question for one issue. However, if they disagree on the question, the default question should be posed, appropriate to the stage of dispute. It would be also possible to ask the question in simplest manner: which party is right? If they do not reach consent about questions the default option should be viable so arbitrators could always choose one option. Possibilities of these questions will soon be further explained.

This phase is characterized also by collaboration in discovery phase and in anonymisation of data while submitting claims and questions. Not achieving any of these elements would lead to breach of agreement and parties should resort to some other form of dispute resolution like regular arbitration or judicial process. The agreement will clearly state that in case of an obstruction of any party, or simply by the will of one party, blind arbitration will seize and parties can resort to some other form of dispute resolution.

Submissions of claims and question will be handled by software built upon PET principles and cryptography for assigning either false name to parties or code name or any other type of hiding proper names of parties involved. Parties themselves will control submissions to prevent indirect discovery of their identities or any confidential information.

6.3 Phase 3 – answering questions.

The answering of individual question will be sequenced in stages:

Fig. 3. Sequence of answering questions.

6.3.1. Applicable law.

Applicable law is usually specified in contracts. Nevertheless, the issues sometimes occur and there is a need for clarification, especially in some predispute agreements. It would also be the first question raised in a proceeding both before an arbitration or in a judicial proceeding, therefore it is only natural to be the first answered. Depending on the answer, the following stages will be decided upon. Even for the factual dispute or even more likely in the case of evaluation of evidence, the applicable law could be sometimes essential precondition, especially if some evidences are to be evaluated in connection with certain standards set by specific law. Default question could be: what law should apply to the dispute?

6.3.2. Facts and evaluation of disputed evidence.

A dispute about facts and evidence evaluations is the most problematic since the general idea is to preserve confidentiality/privacy by withholding information about identity or any indirect information that could reveal the same. Arbitrators should give answer to which facts have occurred based on provided evidence or recounts by parties. The difficult question is how to keep an evaluator of facts and evidence in dark about the parties if they are pointing to parties and indicating them clearly.

At any moment at this phase, an arbitrator could pose questions to the parties, which they will answer, always having in mind not to breach confidentiality agreement. However, if there is more than one arbitrator answering the same question, answers of parties should be available to all of them, as it will serve as a basis for their decision.

If it is necessary to have someone's expertise or to attest that certain disputed facts have happened, an arbitrator would be limited to that aspect, not knowing a reason of one's testimony or deciding upon credibility. Arbitrator would not know the context of one's testimony or significance of his/hers decision for the rest of the process.

If facts are particularly revealing of some elements that parties would want to keep hidden there would still be options: the question could be broken down in several different aspects, the question could be misleading to the nature and answer that is sought, there could be possibility to hide certain elements of facts or to replace them with interchangeable things, at the same time together with these facts there could be offered additional false facts… We can imagine that parties offer 4 different set of supporting facts or stories, but only parties and software know accurate version. It would mean that arbitrators should decide on all of them and maybe their involvement would be 4 times higher, but that is again tradeoff between ultimate goal of confidentiality and costs of the process.

We could also imagine the situation where both parties have previously hired their expert to validate their claims. An arbitrator would give opinion based on their expert reports having in mind the reliability of their methods

In the end arbitrator(s) of facts would give the opinion/ answer to a question which would be taken as a ground on which the later phases would be decided, just as the facts were not disputed at all. If in his answer there would be some revealing element for the identity, it would be removed by parties before transitioning into the next stage. If a dispute involves larger number of facts to be verified, it is even easier to imagine involving bigger number of arbitrators to deal with this task simultaneously.

Default question in case of disagreement about question would be: which party's interpretation of facts is correct?

6.3.3. Legal expertise.

Many times facts will be undisputed and the only issues that parties will have is a legal interpretation of some situation. In the stage where different legal interpretations are resolved (if there are any), we emphasize two possibilities in previous phases: first, that there were no disagreements on facts of the case or they have been negotiated and settled upon; second, the disputed facts have been settled by the previous arbitrators and we can proceed with third stage.

The problem here is of similar nature to the previous stage: if the question is too revealing we must find a way for maintain the anonymity. One way would be for parties to ask general legal question based on false facts. Second way would be to break down legal question to the smallest possible unit, where answers could later be assembled into plausible complete legal answer (by award arbitrator). Third way would be, in addition to masking of names and identifiable information, to offer 4 different set of facts of similar but slightly different nature that would make it impossible to know the real facts of the case. Arbitrator would be required to give decision and reasoning on all 4 set of facts, not knowing which could be true.

Default question in case of disagreement would be: whose legal interpretation is correct?

Finally, if many arbitrators are used and give different decisions about same question, they could be compiled into one decision in favor of one party by majority. The party with decisions that have more than 50 percent of votes wins that question/mini case.

6.3.4. Awards.

The parties could specify the key question in one of the phases, where by answering in favor of one party, that party wins automatically the case. If there is more than one key question, they could set odd number of keys questions and agree that the party, whose interpretation is correct in more answers, wins. Previous questions would establish facts and build arguments for answering key question. However, if the parties do not agree about the key question(s), it remains for arbitrators in the last stage to decide

who has won based on question and accepted argumentation of wining parties in previous stage.

If parties agree, they could state that according to the final decision in a case an award is made in accordance with the claim of the winning party(automatically), or they can pose question about which award is appropriate for the dispute.

7 Future research

The aim of the paper is the initial presentation of the idea, with the hopes of stirring further discussion. Even though we tried to answer some question, inevitably there will be a lot more, relating to the legal, technical and economic aspects of such proposal. All of them will be subject of the future research. Some of the questions that instantly come in mind are: what would be the quality of justice offered by this model, what is the quality vs. secrecy relation, what are the average costs of the arbitration… As regard to some concerns about fair trial/hearing or ethical issues, we do not propose this model to be part of predispute mandatory arbitration clause, which could lead to abuse, but merely to serve as a possible tool if need ever occurs. We would argue that it is appropriate for certain disputes[10] therefore the need might occur in some cases.

8 Conclusion

In this paper we state our belief that it is possible to leverage the wisdom of the crowds for more efficient and cheaper dispute resolution but at the same time to keep the dispute a secret. For these reasons we proposed a counterintuitive idea to solve hidden disputes by opening the issues to unknown persons to decide about, but keeping them in "dark" as much as possible.

We have presented initial idea for the conceptual model that provides with certain confidentiality/privacy improvements in crowdsourced arbitration proceedings. Parties themselves will be bind by confidentiality set in the agreement on the start of the process, specifying their commitment to guard privacy of opposing party. To prevent arbitrators in online arbitration to know the parties or the subject of the dispute, we have proposed the way to break the dispute in smaller bits in order to distribute those bits to wider group on individual arbitrators (the crowd arbitrators), who work independently only on one bit. Thus, we achieve the effect that one single arbitrator cannot grasp the whole picture. We could say it puts into effect famous saying: "you cannot see the forest for the trees".

Our proposal is focused on enhancing impartiality of arbitrators, crowd engagement in dispute resolution and developing cheaper and more pleasant adjudicative form of ODR. Significant potential of blind arbitration could also lay on developing such model for mobile platforms[11], extending the possibility of massive, anonymous dispute resolution based on microworks. It could be one more tool in ODR arsenal giving a new

option to disputants who wish to remain anonymous, don't want to be subject of prejudice or receive bad reputation.

After all, lady Justice (Iustitia) is blindfolded for a reason.

9 Acknowledgements

I would like express gratitude to Erasmus Mundus Law, Science and Technology PhD (LAST-JD) program that allowed me to pursue my research.

10 References

1. Felstiner, W. et al.: Emergence and Transformation of Disputes: Naming, Blaming, Claiming..., The. Law Soc'y Rev. (1980).
2. Herik, J. Van Den, Dimov, D.: Towards Crowdsourced Online Dispute Resolution. J. Int. Commer. Law Technol. 2006, (2012).
3. Hörnle, J.: Cross-border Internet Dispute Resolution. Cambridge University Press (2009).
4. Langheinrich, M.: Privacy by design—principles of privacy-aware ubiquitous systems. Ubicomp 2001 Ubiquitous Comput. (2001).
5. Larson, D.A.: Artificial Intelligence: Robots, Avatars and the Demise of the Human Mediator. (2010).
6. Lodder, A.: The Third Party and Beyond. An analysis of the different parties, in particular The Fifth, involved in online dispute resolution. Inf. Commun. Technol. Law. 15, 2, 143–155 (2006).
7. Lodder, A., Zeleznikow, J.: Artificial Intelligence and Online Dispute Resolution. Online Dispute Resolution: Theory and Practice: A Treatise on Technology and Dispute Resolution. pp. 73–94 Eleven International Publishing (2012).
8. Malone, T.W. et al.: Harnessing Crowds: Mapping the Genome of Collective Intelligence. SSRN Electron. J. (2009).
9. Marder, N.: Cyberjuries: A new role as online mock juries. Univ. Toledo Law Rev. (2006).
10. Martic, D.: Online Dispute Resolution for Cloud Computing Services. ceur-ws.org. ceur-ws (2014).
11. Poblet, M. ed: Mobile Technologies for Conflict Management. Springer Netherlands, Dordrecht (2011).
12. Rabinovich-Einy, O.: Going public: Diminishing privacy in dispute resolution in the Internet age. Va. JL Tech. 4, 1–55 (2002).
13. Rule, C., Nagarajan, C.: Leveraging the Wisdom of the Crowds: the Ebay Community Court and the Future of online Dispute Resolution. ACResolution 2 4. (2010).
14. Surowiecki, J.: The Wisdom of Crowds. Anchor (2005).
15. Wahab, M.S.A.: ODR and e-Abritation. Online Dispute Resolution: Theory and Practice: A Treatise on Technology and Dispute Resolution. pp. 399–442 Eleven International Publishing (2012).
16. Wahab, M.S.A. et al.: Online Dispute Resolution: Theory and Practice: A Treatise on Technology and Dispute Resolution. Eleven International Publishing (2012).

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 delega-

tion 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)

Artificial Prediction Markets as a tool for Syndromic Surveillance

Fatemeh Jahedpari[1], Julian Padget[1], Marina De Vos[1], Benjamin Hirsch[2]

[1] Department of Computer Science, University of Bath, UK
{f.jahedpari,j.a.padget,m.d.vos}@bath.ac.uk
[2] EBTIC, Khalifa University, United Arab Emirates
Benjamin.Hirsch@kustar.ac.ae

Abstract. A range of data sources across the internet, such as google search terms, twitter topics and Facebook messages, amongst others, can be viewed as kinds of sensors from which information might be extractable about trends in the expression of matters of concern to people. We focus on the problem of how to identify emerging trends after the original textual data has been processed into a quantitative form suitable for the application of machine learning techniques. We present some preliminary ideas, including an agent-based implementation and some early results, about the application of artificial prediction markets to such data, taking the specific domain of syndromic surveillance (early stage recognition of epidemics) as an example, using publicly available data sets.

1 Introduction

This paper outlines some early stage research into the application of prediction markets to syndromic surveillance. Prediction markets are seen as a mechanism to forecast the outcome of future events by aggregating public opinion, in which market participants trade so-called securities that represent different probabilities about the (expected) outcome of a scenario. We describe prediction markets in more detail in section 2 and compare them with alternative approaches in section 5.

Syndromic surveillance monitors population health indicators which are apparent before confirmatory diagnostic tests become available, in order to predict a disease outbreak within a society at the earliest possible moment, with the aim of protecting community health. Clearly, the earlier a health threat within a population is detected, the lower the morbidity and the higher the number of lives that may be saved. Syndromic surveillance data sources include, but are not limited to, coding of diagnoses at admission or discharge emergency department, chief complaints, medical encounter prediagnostic data, absentee rates at schools and workplaces, over-the-counter pharmacy sales, Internet and open source information such as people post in social media. Each of these types of data can generate a signal during a disease development. Therefore, given

the vast amount of these data sources, a proper mechanism is necessitated to integrate them as soon as they become available.

In this research, we focus on developing a novel syndromic surveillance technique by integrating different data sources inspired by the crowd-sourcing behaviour of prediction markets. To achieve our goal, we train a multiagent system in an artificial prediction market in a semi-supervised manner.

2 Prediction Markets

Prediction markets have been used to forecast accurately the outcome of political contests, sporting events, and economic outcomes [19]. In this research, we use an artificial prediction market as a mechanism to integrate several syndromic surveillance data sources to predict a level of disease activity within a population on a specific date. This section briefly explains the preliminaries of prediction markets.

The prediction market, also known as an information market, originated at the Iowa Electronic Marketplace (IEM) in 1988 as a means to bet on presidential elections. A prediction market aims to utilise the aggregated wisdom of the crowd in order to predict the outcome of a future event [16]. In these markets, traders' behaviour has the effect of externalising their private information and beliefs about the possible outcomes, and can hence be used to forecast an event accurately [12]. Prediction markets are increasingly being considered as approaches for collecting, summarising and aggregating dispersed information by governments and corporations [10].

In prediction markets, traders bet on the outcome of future events by trading securities. A security is a financial instrument, like a financial stock, that pays a profit (or makes a loss) based on the outcome of the event. Each outcome of an event has a security associated with it. Traders can buy or sell any number of securities before the expiry time of the security. A security expires when the outcome of the event is realised. To illustrate a simple case, a prediction market can be used to predict "if candidate 'X' will win the election" by offering two securities of 'Yes' and 'No'. Assuming the market finally ends with candidate 'X' winning the election, all traders will receive $1 payoff for each 'Yes' security they own and $0 for their 'No' securities, losing the money they spent on buying them.

The aggregated monetary bets made by market traders dynamically determine the price of each security before the market ends. The market price of a security represents the price at which the security can be bought or sold. Also, it can be interpreted as representing the probability of that outcome occurring by fusing the beliefs of all the market participants. Arguably, the price that an agent would pay to buy a security indicates how confident s/he is in the outcome of the event. For example, if a trader believes that the chance of candidate 'X' winning is 80%, s/he then would be willing to buy a 'Yes' security at any price up to $0.80.

A prediction market is run by a market-maker who is the company or individual that interacts with traders to buy and sell securities. The market-maker determines the market price using a market trading protocol. The logarithmic market scoring rule (LMSR) designed by Hanson [9] is an automated market maker. Using LMSR, the price and cost of a security is calculated as follows:

$$C(q_i) = b * \log(\sum_{i=1}^{m} e^{q_i/b})$$

and

$$P(q_i) = \frac{\exp(q_i/b)}{\sum_{j=1}^{m} \exp(q_j/b)}$$

respectively, where m is number of securities that market offers, each for one possible outcome and $q_i \in (q_1, q_2, \ldots, q_m)$ represents the number of units of security i held by market traders The larger the value of b, the more money the market maker can lose. It also means that traders can purchase additional quantities of a security without causing significant price swings. Note that the price of a security only applies for buying a infinitesimal number of shares and the price of the security immediately changes as soon as a traders start trading. In order to calculate the cost of a trading X securities, the market makers must calculate $C(q + X) - C(q)$.

3 Agent-Based Simulation Architecture

In order to explore empirically the application of artificial prediction markets to syndromic surveillance, we have developed an agent-based simulation, which we now describe, followed by some preliminary results in section 4.

3.1 Agents and Strategies

Our model integrates data and beliefs of different data streams by simulating an artificial prediction market to predict the outcome of an event, which in this case is the disease activity level on a specific date. Each data stream includes the quantitative value of a particular disease activity level for a specific place for different period of time. Each agent is responsible for one data stream and trades securities in various prediction markets based on its capital and belief about the disease activity level of the market event date. Trading agents will in due course (see below) learn from each market based on the revenue they receive and the losses they make when the market closes. Consequently, they can update their strategy, beliefs and confidence for the future markets.

The system has a market-maker that uses a scoring rule to calculate the market price for each security in the market, and a data distributor that provide agents with the data stream for which they are responsible, and trading agents. The simulation mechanism,

specified in Algorithm 1, is as follows. At the beginning of the simulation, all the trading agents are awarded an equal amount of initial capital. For each training example, a prediction market (let us say prediction market for week T) is established. At this time, the Data Distributor Agent provides available data to the trading agents, according to their role. Then, the trading agents participate in the market according to their available capital, beliefs, and trading strategies. Agents can trade any number of securities before the deadline for the closing of the market. Once the market deadline is reached, the market-maker reveals the winning security and rewards the winning security holders with \$1 for each winning security they own. These revenues are added to their capital. However, the agents who own losing securities, lose capital equal to the amount spent on purchasing them.

During the simulation, agents with superior data, strategy and analysis algorithms are likely to accumulate greater capital and hence affect market prices and eventually the outcome. In other words, important – by these metrics – agents are identifiable and have greater influence in predicting the outcome of the event. This increased influence of the more successful agents should increase the accuracy of the system overall: the agents are not in competition *per se*, so we do not care which agents are better, but we do want the better ones to have more effect on the prediction mechanism.

The first agent strategy is based on zero intelligence [7], and so has no scope for learning: buy-sell and security choices are random subject to the constraint of not trading at a loss. For the second strategy, we add a basic learning mechanism, following the design of zero intelligence plus [3], in which agents update their trading strategy and beliefs based on the reward they received from the market in order to improve their reward in future markets. This is achieved by incorporates a simple machine learning mechanism (Widrow-Hoff) to adapt their individual behaviour to the market trend.

3.2 Market Instantiation

Different data sources have different timeliness in detecting a disease outbreak. For example, some data sources such as social media data can signal a disease activity level perhaps two weeks earlier than physician data. Therefore, for the system to be capable of forecasting the outcome at the earliest possible moment and not wait for all the agents' data to arrive to start prediction, we will run multiple concurrent markets for consecutive prediction weeks. For example, if the simulation week number is 1, then 4 further markets for weeks 2 to 5 will also be open. Once the deadline of the first market (week 2) is reached, then that market (week 2) closes and another market after the last market is opened (week 6). Consequently, the agents who have data for those further markets can start trading earlier in those markets and take advantage of cheaper prices, which will lead to the updating of market prices as early as possible. In addition, all data with different timeliness ranging from 4 weeks to one day before the event date will be incorporated in each market and at the same time agents can use their knowledge achieved from the previous market when predicting the outcome of a given market.

Algorithm 1: Agent-Based Simulation Architecture Algorithm

```
1  Give start up capital to each agent
2  Simulation-Current-Week C;
3  Market-Date T;
4  for T ← 1 to end do
5      Data Distributor disseminate data, which are accessible by week C, to each agent
       according to agent expertise;
6      Start Prediction Market for Week-T;
7      while Market deadline is not reached do
8          Wait();
9          In here, agents will decide the level of disease activity in week T and trade
           security according to their belief and strategy;
10     end
11     End Prediction Market;
12     Reveal the winning security (Based on the label of training examples);
13     Each agent new capital ← previous capital + revenue gained in this market − amount
       spent for purchasing securities;
14     Now, according to the utility received in this market, agents should update their
       trading strategy and beliefs ;
15     T=T+1;
16     C=C+1;
17 end
```

For the sake of simplicity, Algorithm 1 considers just the one market, as does our current implementation.

4 Preliminary Results

The configuration of the controlling parameters of our system need thorough investigation through running a large number of simulation experiments. These settings include but are not limited to constraints on number of market participants, required time for each market, initial capital for each agent, type of monitored diseases, constraints and requirements for agents to trade, and the minimum required number of training examples. We have only just begun to explore this parameter space.

As discussed in Section 1, there is a vast number of syndromic surveillance data sources. Much research has been done to compare these data sources with the actual value of disease activity level for a specific disease in a particular place. For example, Culotta [5]

stated that he could track influenza rates in the United States using Twitter messages with 95% correlation. Corley [4] could track flu rates in the United States with a correlation of 76% by examining the proportion of blogs containing the two keywords of "influenza" and "flu". Google Flu Trend [6] can predict 'flu activity level with a 97% correlation by analysing queries sent to the Google search engine.

In the first two batches of experiments, we have tried two well-known trading strategies for trading agents: Zero Intelligence (ZI) [7] and Zero Intelligence Plus (ZIP) [3], the former to provide a baseline behaviour and the latter to investigate the effect of a simple learning mechanism on the trading decision. We now discuss each of these in more detail.

4.1 Preliminary Results with ZI

In the first set of experiments, agents use a Zero Intelligence strategy (adapted from [7]) when trading securities in the market. In this model, agents consider a limit price, according to their data, for each security of the market. In each day of the market, they choose one security randomly and purchase a random quantity of that security, if its limit price is higher than its market price or they sell a random quantity of that security if its limit price is less than its market price. In both situations, the agents considers the maximum number of securities that can be traded, based on their available capital and the securities they own. LMSR, described in Section 2, is used as market scoring rule, as it provides infinite liquidity [8] and does not suffer in thin markets where the number of traders are small.

Each agent in the experiment is awarded $10 at start up and one market is established for each training example. The winning security is chosen based on the United State influenza-like illnesses rate from 30 September 2002 to 01 September 2003 [3]. Each market offers eight securities, corresponding to one security for each standard deviation from the mean, covering from -4 to $+4$ standard deviations.

Figure 1 shows how extending the period of a prediction market can help agents to predict the outcome of the event better. It demonstrates that accuracy goes up and the mis-classification rate falls as the duration of the market increases. As can be seen from the figure, accuracy increased from 82% for 10-day-long markets to 93% for 90-day-long markets, after which it is almost flat. From this, we conclude that 110 days seems a sufficient period for each market and hence our subsequent experiments at this stage use this market length. Longer duration markets provide agents with sufficient time to trade enough numbers of the desired securities and approach more closely the equilibrium price.

The purpose of the experiments reported in this paper is primarily to establish confidence in the behaviour of the simulation, by providing tailored data feeds with known

[3] `http://www.cdc.gov/flu/weekly/fluviewinteractive.htm`

properties and then observing whether the agents achieve their expected level of performance, given that data and their (known) strategy. These experiments have a population of 20 agents, each receiving data from data streams with a specified correlation with the United States influenza-like illnesses rate. The agent names in the following figures represents the type of data the agent is receiving. For instance, a95 denotes an agent that receives data with a 95% correlation with the United States influenza-like illnesses rate.

Since all the agents in these experiments are essentially identical in terms of strategy, the difference in their data sources should lead to them obtaining different amounts of revenue in each market. Figure 2 shows their revenue when each market ends. As the Figure shows, a100 agent, which has complete information about all the events of the experiment, earns high revenue in all markets with the exception of 31/07/2003. In this case, it was not making the wrong choice – it cannot – but was trying to purchase a large number of securities and since the price of a security increases as a result of its purchase, the agent (as explained in Section 2), did not have sufficient capital to complete the deal. In other words, for agents without perfect information, the enforced random choice of security to trade (the ZI strategy) means the agent cannot select the most appropriate one, but rather the one that chance dictates and hence it makes a loss.

Figure 3 shows the capital held by each agent at the end of each market in one experiment and Figure 4 shows the average capital of agents over 50 runs. The main observation from these figures is that, as expected, agents with higher quality data are able to achieve higher levels of revenue. As can be seen from these figures, agent a100 accumulates more capital than other agents even those with high quality data such as a99. The reason for this is that a100 never makes a mistake while the other agents do and as soon as one agent predicts an outcome, it dedicates most of its capital to purchase the corresponding security. Therefore, once an agent predicts a wrong outcome, it loses all its capital, while agent a100 keeps earning revenue in each market and accumulates more capital and hence invests more on upcoming markets and earns more revenue again. Also, agent a100 causes the price of the correct security to increase rapidly as it purchase a large quantity of it and, therefore, makes it difficult for other agents to buy significant quantities of that security due its high price.

Clearly, more comprehensive experimentation is necessary, backed up with appropriate statistical confidence tests. In this section, we have only used the most basic of strategies and one that has known flaws [3]. However, it provides both a useful baseline performance, as well as a setting in which initial hypotheses about the effectiveness of the prediction market model can be validated (such as the agent with 100% correlated data dominating the market and all others losing all their investments).

4.2 Preliminary Results with ZIP

In the second set of experiments, we have changed the agent strategy from ZI to ZIP, by which the agent uses data about trends in the market in order to adjust their behaviour to

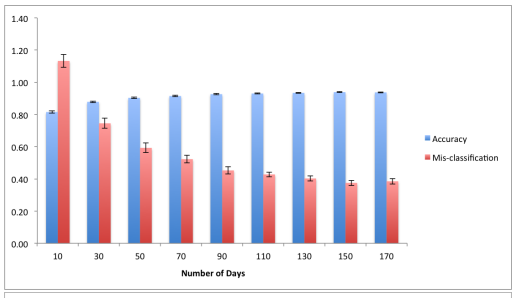

Fig. 1. ZI: Comparing accuracy (s.d.: 0.017–0.047) and mis-classification (s.d.: 0.101–0.289) on the y-axis vs. duration of the prediction market (x-axis). Each data point is the average of 50 experiments with same parameter settings

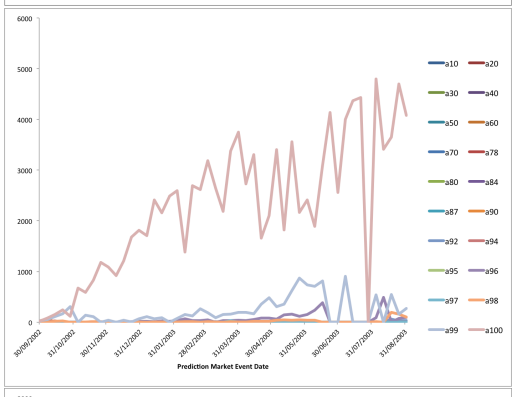

Fig. 2. ZI: Comparing revenue of agents (y-axis) at the end of each market (x-axis) for an example run chosen at random.

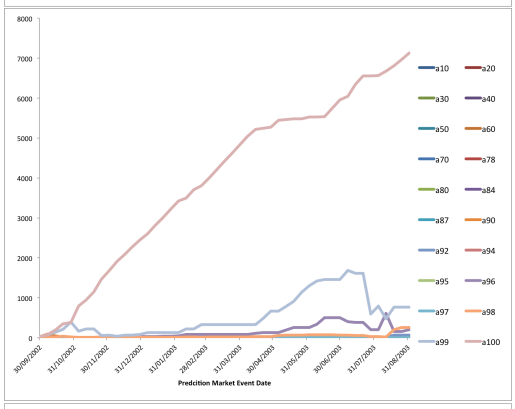

Fig. 3. ZI: Comparing capital of agents (y-axis) at the end of each market (x-axis), for an example run chosen at random.

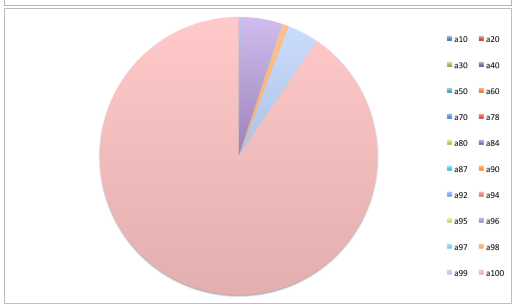

Fig. 4. ZI: Comparing capital of agents (y-axis) at the end of the experiment (averaged over 50 runs).

be less random (as in ZI) and more in line with the market valuation of a given security. Although ZIP and its variants have been shown to be effective strategies in terms of profit making, there are two reasons why this approach is likely to be ineffective in the context of prediction markets:

1. The ZIP strategy depends upon both buyer and seller employing this strategy, but in the prediction market, the two parties are the buyer and the market-maker, of which the latter has no interest in profit and which has no strategy as such. Consequently, only one party in the market is 'learning'. This still has a positive effect as discussed below, but starts to underline the difference between trading markets (with bilateral strategies) and prediction markets (with unilateral strategies).
2. The ZIP strategy aims at trading for profit regardless of the (financial) instrument being traded, leading to the establishment of an equilibrium price, whereas the point of a prediction market is to choose the *right* instrument, rather than the currently most profitable.

The experiments are run with the same data as for ZI. Thus Figures 6 and 7 show the results from a randomly chosen 110 day market, as was done for ZI. It is notable that ZIP achieved higher accuracy – nearly 97% – with market durations of > 30 days, than in the ZI experiment. However, as before a100 agent dominates the market.

The story in terms of revenue (Figure 6) is much the same as for ZI, although a99 stops making a profit much sooner with ZIP. As can be seen from the figure, the agent obtains revenue until 04/11/2002, which is the first time that it makes a mistake and loses the majority of its capital. Consequently, the agent has little remaining capital, but continues earning money for the following two markets, but its second mistake (on 25/11/2002) bankrupts it, after which it cannot invest further. This scenario applies to all other agents and causes a100 to dominate the market, as it never makes any mistake.

5 Related Work

Many syndromic surveillance systems exist worldwide, each designed for a specific country, region or state [11, 20, 15, 13, 17]. We refer to them as traditional, since they do not utilise internet based data. While these systems can detect an outbreak with high accuracy, they suffer from slow response times. For example, the Centers for Disease Control and Prevention (CDC) publishes USA national and regional data typically with a 1–2 week reporting lag. It monitors over 3,000 health providers nationwide to report the proportion of patients seen that exhibit influenza-like illnesses (ILI) [4] [5, 6].

On the other hand, modern syndromic surveillance systems appeal to internet based data such as search engine queries, health news, and peoples' posts on social networks

[4] `http://www.cdc.gov/flu/weekly/fluactivity.htm`

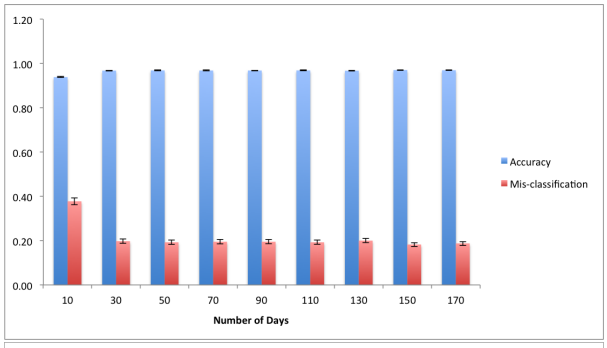

Fig. 5. ZIP: Comparing accuracy (s.d.: 0.010–0.018) and mis-classification (s.d.: 0.062–0.109) on the y-axis vs. duration of the prediction market (x-axis). Each data point is the average of 50 experiments with same parameter settings

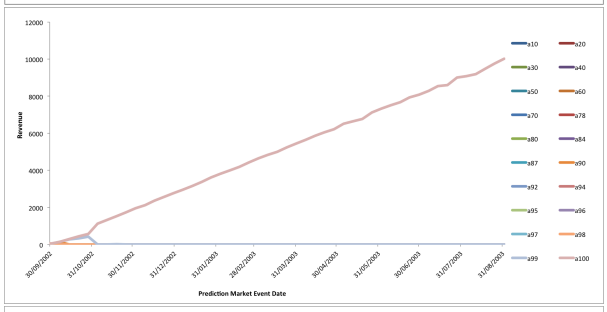

Fig. 6. ZIP: Comparing revenue of agents (y-axis) at the end of each market (x-axis) for an example run chosen at random.

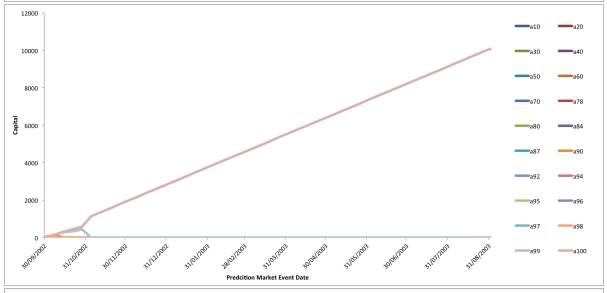

Fig. 7. ZIP: Comparing capital of agents (y-axis) at the end of each market (x-axis) for an example run chosen at random.

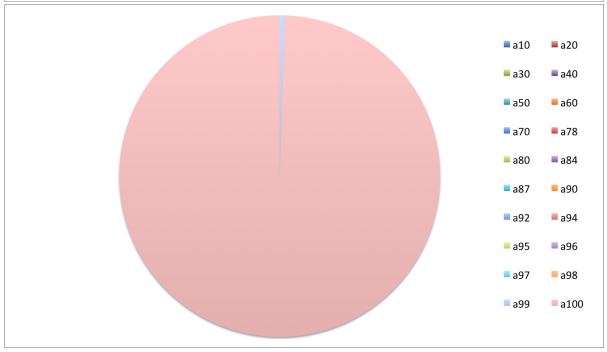

Fig. 8. ZIP: Comparing capital of agents (y-axis) at the end of the experiment (averaged over 50 runs).

to predict an outbreak earlier [18, 2, 4], albeit with necessarily lower precision. While some of them claim that they could achieve high accuracy, they are vulnerable to false alarms [6, 1] due to their dependence on a single data stream and disregarding the benefits from fusing different data sources. Ginsberg et al [6] state, regarding Google Flu Trends, that "Despite strong historical correlations, our system remains susceptible to false alerts caused by a sudden increase in ILI-related queries. An unusual event, such as a drug recall for a popular cold or flu remedy, could cause such a false alert".

To the best of our knowledge, there is no system that fuses both traditional and internet based data sources. This could be due to the different timescales that these data sources have and the consequent issues of appropriate synchronisation.Prediction markets can overcome this problem as traders can trade securities as soon as they receive new information and impact the price and consequent probability of an event outcome. It is interesting to note that Polgreen et al [14] report on the use a prediction market with human health care expert participants to forecast infectious disease activity 2-4 weeks in advance.

Moreover, internet based system are only suitable for places where sufficient source data is available. For example, twitter-based systems cannot have a high accuracy on places where using twitter is not very common, if even accessible. In addition, even if sufficient data is available, system accuracy cannot be guaranteed worldwide since peoples' behaviour changes from place to place, reflecting differing (digital) cultures. For example, people in a particular city may seek a physician as soon as they encounter the symptoms of a disease and do not trust online information, while people in another city may defer visiting a doctor and seek out online information in order to cure themselves at the early stages of their sickness. Furthermore, peoples' behaviour may change over time. For example, a particular social media may become less popular and cede its role to newer technology over the time.

6 Discussion

Since we are in an early stage of this research, a substantial part of the work is to come. We have numerous ideas that have yet to be implemented, including: (i) the learning capability of agents, (ii) consideration of the confidence of agents, (iii) of the different timeliness of data streams, and (iv) the effect of a heterogeneous population of agents with different trading strategy and risk prediction model –among other characteristics. The very preliminary results we have meet our broad expectations for the behaviour of prediction markets, but it is too early to say whether they can be a general-purpose tool with useful levels of precision and recall across a range of domains. The ZI strategy, being essentially random under the constraint of not making a loss, establishes a useful performance baseline, as well as a framework against which to validate the basic system hypotheses. The ZIP strategy, while appropriate for bilateral markets seeking to establish equilibrium prices, is inappropriate – at least, as conventionally formulated –

for prediction markets, although the dampening effect of the learning mechanism does lead to higher prediction rates and smoother overall behaviour.

We welcome feedback on the appropriateness of the approach and the above directions for development as well as alternative mechanisms that might be incorporated in the prediction market setting.

References

1. Harshavardhan Achrekar, Avinash Gandhe, Ross Lazarus, Ssu-Hsin Yu, and Benyuan Liu. Predicting flu trends using twitter data. In *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, pages 702–707. IEEE, 2011.
2. Herman Anthony Carneiro and Eleftherios Mylonakis. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases*, 49(10):1557–1564, 2009.
3. Dave Cliff and Janet Bruten. Minimal-intelligence agents for bargaining behaviours in market-based environments. Technical Report HPL-97-91, Hewlett-Packard Laboratories, 1997. Available via `http://www.hpl.hp.com/techreports/97/HPL-97-91.html`, retrieved 20071115.
4. Courtney D Corley, Diane J Cook, Armin R Mikler, and Karan P Singh. Text and structural data mining of influenza mentions in web and social media. *International journal of environmental research and public health*, 7(2):596–615, 2010.
5. Aron Culotta. Detecting influenza outbreaks by analyzing twitter messages. *arXiv preprint arXiv:1007.4748*, 2010.
6. Jeremy Ginsberg, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S Smolinski, and Larry Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2008.
7. Dhananjay K Gode and Shyam Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of political economy*, pages 119–137, 1993.
8. Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.
9. Robin Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1:3–15, 2007.
10. Robin Hanson, Ryan Oprea, and David Porter. Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior & Organization*, 60(4):449–459, 2006.
11. Michael D Lewis, Julie A Pavlin, Jay L Mansfield, Sheilah OBrien, Louis G Boomsma, Yevgeniy Elbert, and Patrick W Kelley. Disease outbreak detection system using syndromic data in the greater Washington DC area. *American journal of preventive medicine*, 23(3):180–186, 2002.
12. Evdokia Nikolova and Rahul Sami. A strategic model for information markets. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 316–325. ACM, 2007.
13. Yasushi Ohkusa, M Shigematsu, K Taniguchi, and N Okabe. Experimental surveillance using data on sales of over-the-counter medications-japan, November 2003–April 2004. *MMWR Morb Mortal Wkly Rep*, 54:47–52, 2005.
14. Philip M. Polgreen, Forrest D. Nelson, and George R. Neumann. Use of prediction markets to forecast infectious disease activity. *Clinical Infectious Diseases*, 44(2):pp. 272–279, 2007.

15. Emergency department syndromic surveillance system: England & Northern Ireland, September 2013.
16. Russ Ray. Prediction markets and the financial "wisdom of crowds". *Journal of Behavioral Finance*, 7(1):2–4, 2006.
17. Communicable and respiratory disease report for england & wales, October 2013. Available from: `http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317139997178`, retrieved 20131216.
18. Alessio Signorini, Alberto Maria Segre, and Philip M. Polgreen. The use of twitter to track levels of disease activity and public concern in the U.S. during the Influenza A H1N1 Pandemic. *PLoS ONE*, 6:e19467, 05 2011.
19. Erik Snowberg, Justin Wolfers, and Eric Zitzewitz. Prediction markets for economic forecasting. Working Paper 18222, National Bureau of Economic Research, July 2012. Available from: `http://www.nber.org/papers/w18222`.
20. Fu-Chiang Tsui, Jeremy U Espino, Virginia M Dato, Per H Gesteland, Judith Hutman, and Michael M Wagner. Technical description of rods: a real-time public health surveillance system. *Journal of the American Medical Informatics Association*, 10(5):399–408, 2003.

A Research Agenda for Prediction Markets

Patrick Buckley[1]

[1]Kemmy Business School, University of Limerick, Ireland
Patrick.Buckley@ul.ie

Abstract. This paper describes the use of prediction markets as tools for enabling collective intelligence. Their benefits are explored and current applications are elucidated. Moving on from this, key open research questions from the literature are identified, and a research agenda that can address these issues is introduced.

Keywords: Prediction Markets, Collective Intelligence

1 Introduction

Prediction markets are a relatively novel form of decision making. The core insight upon which they are based is that a market mechanism can be used to enable two processes which are crucial to effective decision making. First, the provision of individual rewards to participants prompts truthful information revelation. Second, asset price movement within a market provides a mechanism that can be adapted to support information aggregation. When deployed using Information Technology (IT), prediction markets can trivially scale to hundreds or even thousands of participants. This scalability enables collaborative decision making on a scale that many other group decision making mechanisms would find prohibitive. They provide a method of generating collective intelligence that can draw upon the wisdom of large, disparate crowds.

This paper is structured as follows. In section 2, we introduce the concept of prediction markets. We particularly focus on the theorised benefits of prediction markets from a decision making perspective and elucidate current applications of prediction markets. As befits a relatively novel innovation, there are many open research questions regarding prediction markets, which are discussed in section 3. In section 4 we present a brief description of a methodology which can provide data to investigate a research agenda in prediction market that can address some of the previously identified issues. We conclude in section 5 by calling for help in operationalizing this research agenda.

2 Prediction Markets

2.1 Prediction Markets

Prediction markets are "markets that are designed and run for the primary purpose of mining and aggregating information scattered among traders and subsequently using this information in the form of market values in order to make predictions about specific future events." [1, p. 75]. The theoretical roots of prediction markets can be found in Hayek's conceptualization of markets as near perfect transmitters of information [2]. This perspective on market operation led to the formulation of the efficient market hypothesis, which has been expressed as stating that stock "prices at any time 'fully reflect' all available information" [3, p. 383]. There are a number of forms of the efficient market hypothesis, including the weak, semi-strong and strong form, which make more or less demanding claims as to the efficiency of information transmission within markets [4], [5]. While it is relatively trivial to point to specific examples of market failure, in general, speculative markets such as those in stocks, commodities and future options do a credible, if imperfect job of aggregating relevant information into market prices [6]. This position is backed by a substantial body of empirical evidence [5], [7, 9].

A prediction market is created by offering a contract on the outcome of a future event of interest for sale to a group of participants. For example, suppose an organisation wishes to forecast whether or not a project will reach its next milestone on time. The organisation could create a contract PROJ, which will pay a holder €1 on the date of the milestone if the milestone is reached or €0 otherwise. The organisation would set the initial price of the contract at 50 cents and then offer it for sale to individuals participating in the project. Under these circumstances, if an individual believes that the project is likely to reach its milestone, they will buy the contract in the expectation of receiving a greater reward in the future. Equally, if a rational individual believes the project will not reach its milestone, then they will sell (or 'short') the contract, taking the profit immediately. Individuals buying or selling the contracts being offered will have the effect of moving the price of the contract.

This two-outcome model can be easily extended to allow for the creation of contracts across a range of disjoint outcomes. For example, a prediction market can be created which asks participants to forecast what will be the most successful product from a range of products. They can also be used to allow participants to forecast values rather than select from a particular set of options. As an example, participants may be asked to forecast the total sales of a particular product.

Prediction markets differ from traditional financial markets in two important ways. First, prediction markets operate by enabling participants to trade contracts, whose value is dependent upon the outcome of a future uncertain event [10]. In a prediction market, the trade of contracts in a market place allows participants to exchange information. The trade of contracts also acts as a decision mechanism, since the price of the contract at any point in time can be viewed as the consensus opinion of all the participants in the market as to the likelihood of an event occurring. In this way, the

trade of contracts enables the underlying processes of communication and decision making that is required to allow group decision making to occur [11].

The second distinguishing characteristic of a prediction market is that its primary concern is the elicitation of information [12]. In the modern world, many markets exist that allow participants to trade assets whose value is dependent upon an uncertain future event. While these markets can be viewed as prediction markets from a certain perspective, in general this paper will follow the guidelines proposed by Wolfers and Zitewitz [13]. It steers away from markets where the primary role is enhancing the enjoyment of an external event through taking on risk. Similarly, markets whose primary rationale for existence is that they enable the hedging of financial risk will not be considered prediction markets.

2.2 Benefits of Prediction Markets

Researchers have identified a number of theoretical benefits of prediction markets over comparable information aggregation mechanisms such as polls or expert groups [14]. First, prediction markets provide incentives for truthful information revelation. Second, they provide an algorithm for automatically communicating and aggregating information. Third, prediction markets implicitly weight the information supplied by participants. Fourth, prediction markets can scale efficiently to very large groups, a major advantage over other forms of group decision making, particularly where relevant information is widely dispersed. Fifth, prediction markets can operate in real-time over a long period of time. Finally, prediction markets can be designed in such a way as to allow for trader anonymity.

Prediction markets are instantiated by offering contracts for trade whose value is dependent upon the outcome of a future event. Contracts are specified in the format, "Pay €X if event Y occurs". Individual participants buy and sell these contracts. Rewards for correct forecasts accrue to the individual who holds the contract. This individualization of reward creates an incentive for individuals to hold contracts in events they believe are likely to occur [15]. By providing an individualized incentive some of the challenges associated with information revelation in other domains can be ameliorated [10]. In a deliberative group, individuals may have little incentive to reveal private information, since any benefits will accrue to the group as whole. By providing information to the group, they bestow benefits on others without any reward to themselves, and possibly facing high private costs [16]. The provision of a direct financial incentive to an individual can serve as a counter weight to the emotional, political and professional factors that may inhibit truthful information revelation in a group setting. Since participants are rewarded for accurate decisions, all other things being equal, the provision of individualized incentives should promote information search [17]–[19].

The second characteristic of prediction markets is that they implicitly contain an algorithm for information aggregation. The operation of the market in contracts, and the trading it facilitates automatically creates the equilibrium price which is used as a proxy for estimates about the event of interest [10]. By allowing experts to trade with each other, prediction markets allow disparate opinions and beliefs to be aggregated

into a coherent, consistent whole [18]. As well as providing a mechanism for aggregating the private beliefs of individuals, prediction markets can also enable individual participants to extract information from observing market estimates [20], and correct biases in publicly available information [21].

Several authors point out that prediction markets implicitly weight the information supplied by participants [18], [22], [23]. If participants are more confident of their beliefs in a particular topic, they will be willing to buy more of the relevant contracts, and vice versa. The ability of participants to choose the level of their investment allows them to indicate their confidence in their information in a manner which is automatically accommodated by the aggregation algorithm.

The nature of the market structure also means that prediction markets can scale to very large groups [15]. When considering a market that utilizes information technology to enable trading, the only real limits on the number of participants are computational. This means that prediction markets potentially have lower running costs, particular if they are in operation over a period of time [24]. Most of the overheads in deploying prediction markets are involved in setting up the market and attracting participants. It also means that prediction markets can be created that can utilize participants from outside traditional organizational boundaries, recruiting participants from suppliers, customers and other stakeholders in order to improve the decision making process.

Prediction markets can operate in real-time [10], [24]. This gives them a significant advantage over other comparable information aggregation methods such as polls. Finally, prediction markets can be designed in such a way as to allow trader anonymity [25]. Power relationships and social interactions in group decision making are often seen as responsible for some of the weaknesses of group decision making [11]. The utility of this attribute can vary, but the ability to enable it demonstrates the flexibility of prediction markets.

2.3 Applications of Prediction Markets

Markets which share the defining characteristics of prediction markets have existed for hundreds of years. Specific examples from the literature include markets on Papal elections in 16th century Italy, parliamentary elections in 18th and 19th century Britain and American presidential elections [26], [27]. Modern interest in prediction markets is generally held to have begun with the establishment of the Iowa Electronic Market (IEM) in 1988, which is often seen as the first implementation of a prediction market. (Joyce Berg et al. 2008a). Since then academic and practitioner interest in prediction markets has continued to grow [1].

Modern operational prediction markets can be broadly divided into three categories. The first subdivision is that between public and private prediction markets. A public prediction market is one which invites participation from the general public. A private prediction market is one created by a sponsor which seeks to recruit participants from a specific, albeit potentially very large population. Within public prediction markets, some prediction markets operate using real currency. Participants invest their own money in the market, and gain or lose according to their performance.

Other public prediction markets use virtual currency to enable trading. Table 1 lists some exemplars of these prediction markets.

Type	Example
Public (real currency)	Intrade, http://www.intrade.com Betfair, http://www.betfair.com Iowa Electronic Market, http://tippie.uiowa.edu/
Public (virtual currency)	The Hollywood Stock Exchange, Hub-dub, http://www.hubdub.com Newsfutures, http://www.lumenogic.com Foresight Exchange, http://www.ideosphere.com
Private	Qmarkets, http://www.qmarkets.com Inkling markets, http://inkling.com Crowdcast, http://www.crowdcast.com Prokons, http://www.prokons.com

Table 1. Selected operational prediction markets

Private prediction markets are most pertinent to this discussion. Organizations are interested in using prediction markets to tap the valuable private information held by employees and other stakeholders in the organization [21]. Academic references and analyses on the use of prediction markets as internal decision support tools for various organizational functions is still limited, although increasing all the time. Ortner [28] describes the use of prediction markets in a project management process in Siemens in Austria, with another example of prediction markets use in project management offered by Remidez and Joslin [25]. A number of papers discuss the use of prediction markets as sales forecasting tools [29], [30]. A similar case study, forecasting market share in the Austrian mobile phone market is described by Waitz and Mild [31]. Hopman [32] describes the use of prediction markets for demand forecasting in Intel, with other authors offering examples from the medical domain [24], [33]. Hahn and Tetlock report Eli Lilly have used prediction markets to evaluate what drugs will be successful, while Microsoft have used them to forecast sales of software [18]. Other organizations that are reported in the literature as having used prediction markets include Motorola, Qualcomm, Infoworld, MGM, Chiron, TNT, EA Games, Yahoo, Corning, Masterfoods, Pfizers, Abbott, Chrysler, General Mills, O'Reilly and TNT [34].

Other authors have focused on providing theoretical descriptions of the applications of prediction markets in organizations. Passmore et al. [35] describe how prediction markets can be used to support the Human Resource function in organizations. Other authors have suggested prediction markets can have applications in the domain of risk management [36]–[39]. Sunstein [16] offers a list of possible applications of prediction markets, while other authors point out the power of prediction markets as communication tools in an organizational setting [40,41].

3 Open Research Questions

Much of the academic work on prediction markets to date has focused on assessing their accuracy, both relative to comparable methods and in absolute terms. Academic research to date suggests that prediction markets "can provide more accurate forecasting and effective aggregation than other predictive technologies" [10, p. 45]. Empirical work has demonstrated their effectiveness versus competing mechanisms [42]. Other authors caution against drawing definitive conclusions, but summarise the existing empirical evidence as cautiously optimistic [43]–[45].

The establishment of the basic credibility of prediction markets as information aggregation tools has to lead to calls for studies which move beyond assessing predictive accuracy [46]. A number of research questions emerge from the literature. Much of the predictive power of prediction markets is derived from having large numbers of traders. A key question that emerges from the literature is how traders can be attracted to participate in prediction markets [45]. Related to this question is the concern that as a group decision making tool, prediction markets may be more attractive to individuals who possess certain personality traits. If prediction markets only attract individuals with a high risk tolerance, this may potentially limit their usefulness, particularly in organisational decision making contexts.

Another major concern noted in the literature is the theoretical possibility that prediction markets can be adversely affected by manipulation [10], [45], [47]. In this context, manipulation is an attempt by an individual or group of traders to affect the outcome of the prediction market in a manner which contradicts their own privately held information. Individuals may be motivated to manipulate a prediction market if their utility for determining the outcome of a prediction market outweighs the incentives offered for truthful information revelation.

4 Proposed Research Agenda

Addressing the research question outline above requires a research methodology which can provide data which has a number of properties. A research agenda investigating these concerns requires the use of psychometric instruments to measure personality traits of individuals and the correlation of those measurements with observed behaviours in a prediction market. It would be necessary to collect data across a temporal period. This would allow gathering data on how trading patterns and behaviours such as attempted manipulations impact upon the market as a whole. By collecting data across a temporal window and correlating that data with measurements of individual participant's personalities, it becomes possible to investigate how different individuals respond to different types of feedback.

One potential source of data that can be used to investigate these issues is prediction markets which are used in a pedagogical setting. This application of prediction markets has recently begun to receive academic interest [39], [48], [49]. The benefits of prediction market participation to learners in the cognitive and affective domains of

learning makes a powerful case for their inclusion in curricula, particularly in large group teaching environments. This in turn opens up the possibility of prediction markets that are run in an educational setting being used as research tools.

A prediction market that is ran as part of a large course will have a stable pool of participants. The participants can be accessed directly by researchers, and prompted to complete psychometric instruments. The prediction market will be a relatively controlled environment in that the questions asked, the duration of the trading periods and the incentives offered are all under the control of the researcher. The data that is collected by the market, including price movements and trading decision are all captured by the market and can be correlated with data on specific individuals through the use of an identifier. Of courses it is necessary to ensure that the pedagogical justification for using prediction markets in an educational design is not undermined by the research programme, but with careful design it should be possible to both provide learners with a valuable educational experience and at the same time drive a research agenda forward.

5 Conclusions

Prediction markets have been positioned in the literature as tool for enabling collective intelligence and group decision making. Their potential has led to calls in the literature for more nuanced research programmes which move beyond evaluating their accuracy to investigate issues such as participant behaviour and the effect of manipulation.

We have pioneered the use of prediction markets as pedagogical tools, and have published extensively in the area. In this paper, we propose that this specific application of prediction market is a potentially useful research methodologies that can be used in investigate a number of issues of concern to prediction markets researchers. We believe that we can make a contribution to the larger study of collective intelligence by providing a more nuanced understanding of the strengths, weaknesses and characteristics of prediction markets. We would welcome collaborators and partners who would be interested in developing this research agenda. We would be delighted to offer our expertise in deploying prediction markets in an educational setting to partners, with a view to developing a research agenda that could both investigate the research questions outlined above and also investigate the effect of culture on prediction market performance.

6 References

[1] G. Tziralis and I. Tatsiopoulos, “Prediction Markets: An Extended Literature Review,” *J. Predict. Mark.*, vol. 1, no. 1, pp. 75–91, Feb. 2007.

[2] F. A. Hayek, “The Use of Knowledge In Society,” *Am. Econ. Rev.*, vol. 35, no. 4, pp. 519–530, 1945.

[3] E. F. Fama, “Efficient Capital Markets: A Review of Theory and Empirical Work,” *J. Finance*, vol. 25, no. 2, pp. 383–417, 1970.
[4] G. Kamp and P. A. Koen, “Improving the Idea Screening Process within Organizations using Prediction
Markets: A Theoretical Perspective,” *J. Predict. Mark.*, vol. 3, no. 2, pp. 39–64, Aug. 2009.
[5] R. Kolb, *Understanding futures markets*, 5th ed. Malden MA: Blackwell Publishers, 1997.
[6] R. Hanson, “Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation,” *J. Predict. Mark.*, vol. 1, no. 1, pp. 3–15, Feb. 2007.[7] B.. G. Malkiel, “The Efficient Market Hypothesis and Its Critics,” *J. Econ. Perspect.*, vol. 17, no. 1, pp. 59–82, Winter 2003.
[8] B. G. Malkiel, “Reflections on the Efficient Market Hypothesis: 30 Years Later,” *Financ. Rev.*, vol. 40, no. 1, pp. 1–9, Feb. 2005.
[9] R. Roll, “Orange Juice and Weather,” *Am. Econ. Rev.*, vol. 74, no. 5, pp. 861–880, 1984.
[10] C. Hall, “Prediction Markets: Issues and Applications,” *J. Predict. Mark.*, vol. 4, no. 1, pp. 27–58, May 2010.
[11] D. G. Ellis and B. A. Fisher, *Small Group Decision Making: Communication and the Group Process*, 4th ed. New York: McGraw-Hill Book Company, 1994.
[12] T. W. Bell, “Private Prediction Markets and the Law,” *J. Predict. Mark.*, vol. 3, no. 1, pp. 89–110, Apr. 2009.
[13] J. Wolfers and E. Zitzewitz, “Prediction markets,” *J. Econ. Perspect.*, vol. 18, no. 2, pp. 107–126, 2004.
[14] E. Servan-Schreiber, J. Wolfers, D. M. Pennock, and B. Galebach, “Prediction Markets: Does Money Matter?,” *Electron. Mark.*, vol. 14, no. 3, pp. 243–251, 2004.
[15] R. W. Hahn and P. C. Tetlock, “Introduction to Information Markets,” in *Information Markets: A New Way of Making Decisions*, Washington D.C: AEI-Brookings Joint Center for Regulatory Studies, 2006, pp. 1–12.
[16] C. R. Sunstein, “Deliberation and Information Markets,” in *Information Markets: A New Way of Making Decisions*, Washington D.C: AEI-Brookings Joint Center for Regulatory Studies, 2006, pp. 66–100.
[17] J. E. Berg and T. A. Rietz, “Prediction Markets as Decision Support Systems,” *Inf. Syst. Front.*, vol. 5, no. 1, pp. 79–93, Jan. 2003.
[18] R. W. Hahn and P. C. Tetlock, *Information Markets: A New Way of Making Decisions*. Washington D.C: AEI-Brookings Joint Center for Regulatory Studies, 2006.
[19] C. R. Sunstein, “Deliberating Groups vs. Prediction Markets (or Hayek’s Challenge to Habermas),” *Episteme J. Soc. Epistemol.*, vol. 3, no. 3, pp. 192–213, 2006.
[20] K. Kálovcová and A. Ortmann, “Understanding the Plott-Wit-Yang Paradox,” *J. Predict. Mark.*, vol. 3, no. 3, pp. 33–44, Dec. 2009.
[21] T. S. Gruca and J. E. Berg, “Public Information Bias and Prediction Market Accuracy,” *J. Predict. Mark.*, vol. 1, no. 3, pp. 219–231, Dec. 2007.
[22] J. E. Berg and T. A. Rietz, “The Iowa Electronic Markets: Stylized Facts and Open Issues,” in *Information Markets: A New Way of Making Decisions*, R. W. Hahn and P. C. Tetlock, Eds. Washington D.C: AEI- Brookings Joint Center for Regulatory Studies, 2006, pp. 142–169.
[23] A. Graefe and C. Weinhardt, “Long-Term Forecasting with Prediction Markets A Field Experiment on Applicability and Expert Confidence,” *J. Predict. Mark.*, vol. 2, no. 2, pp. 71–91, Sep. 2008.

[24] P. M. Polgreen, F. Nelson, and G. Neumann, "Using Prediction Markets to Forecast Trends in Infectious Diseases," *Microbe*, vol. 1, no. 10, pp. 459–465, 2006.
[25] H. Remidez and C. Joslin, "Using Prediction Markets to Support IT Project Management - Research in Progress," in *International Research Workshop on IT Project Management 2007*, Montreal, Quebec, Canada, 2007.
[26] P. W. Rhode and K. S. Strumpf, "Historical Presidential Betting Markets," *J. Econ. Perspect.*, vol. 18, pp. 127–141, May 2004.
[27] P. W. Rhode and K. S. Strumpf, "Historical Political Futures Markets: An International Perspective," Universtiy of Arizona, Oct-2008.
[28] G. Ortner, "Forecasting Markets - An Industrial Application," Technical University of Vienna, 1997.
[29] K.-Y. Chen and C. R. Plott, "Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem," Hewlett Packard Labs, California, 2002.
[30] K.-Y. Chen, L. R. Fine, and B. A. Huberman, "Predicting the Future," *Inf. Syst. Front.*, vol. 5, no. 1, pp. 47–61, Jan. 2003.
[31] M. Waitz and A. Mild, "Improving Forecasting Accuracy in Corporate Prediction Markets A Case Study in the Austrian Mobile Communication Industry," *J. Predict. Mark.*, vol. 3, no. 3, pp. 49–62, Dec. 2009.
[32] J. Hopman, "Using Forecasting Markets to Manage Demand Risk," *Intel Technol. J.*, vol. 11, no. 2, pp. 127–136, 2007.
[33] D. Rajakovich and V. Vladimirov, "Prediction Markets as a Medical Forecasting Tool: Demand for Hospital Services," *J. Predict. Mark.*, vol. 3, no. 2, pp. 78–106, Aug. 2009.
[34] G. Tziralis, G. Vagenas, and S. Ponis, "Prediction Markets, an Emerging Web 2.0 Business Model: Towards the Competitive Intelligent Enterprise," in *Web 2.0*, Boston, MA: Springer US, 2009, pp. 271–291.[35] D. L. Passmore, E. D. Cebeci, and R. M. Baker, "Market-Based Information for Decision Support in Human Resource Development," *Hum. Resour. Dev. Rev.*, vol. 4, no. 1, pp. 33 –48, Mar. 2005.
[36] O. Bergfjord, "Prediction Markets as a Tool for Management of Political Risk," *J. Predict. Mark.*, vol. 2, no. 2, pp. 1–12, Sep. 2008.
[37] M. Berlemann and C. Schmidt, "Predictive Accuracy of Political Stock Markets Empirical Evidence from a European Perspective," Humboldt Universitat Berlin, 2001.
[38] G. Bruggelambert, "Information and efficiency in political stock markets: using computerized markets to predict election results," *Appl. Econ.*, vol. 36, no. 7, pp. 753–768, 2004.
[39] J. Garvey and P. Buckley, "Implementing control mutuality using prediction markets: a new mechanism for risk communication," *J. Risk Res.*, vol. 13, no. 7, pp. 951–960, 2010.
[40] P. Buckley and F. McGrath, "Managing prediction markets," in *Proceedings of the special interest group on management information system's 47th annual conference on Computer personnel research*, New York, NY, USA, 2009, pp. 217–220.
[41] M. Gebauer and D. Glienke, "Communicating Strategy with Information Markets," in *Analysis*, San Diego, CA., 2007.
[42] G. H. Van Bruggen, M. Spann, G. L. Lilien, and B. Skiera, "Prediction Markets as institutional forecasting support systems," *Decis. Support Syst.*, vol. 49, no. 4, pp. 404–416, Nov. 2010.
[43] T. S. Gruca, J. E. Berg, and M. Cipriano, "Incentive and Accuracy Issues in Movie Prediction Markets," *J. Predict. Mark.*, vol. 2, no. 1, pp. 29–43, May 2008.

[44] J. Ledyard, "Designing Information Markets for Policy Analysis," in *Information Markets: A New Way of Making Decisions*, Washington D.C: AEI-Brookings Joint Center for Regulatory Studies, 2006, pp. 37–66.
[45] J. Wolfers and E. Zitzewitz, "Five Open Questions About Prediction Markets," in *Information Markets: A New Way of Making Decisions*, Washington D.C: AEI-Brookings Joint Center for Regulatory Studies, 2006, pp. 13–36.
[46] T. S. Gruca, J. E. Berg, and M. Cipriano, "Consensus and Differences of Opinion in Electronic Prediction Markets," *Electron. Mark.*, vol. 15, no. 1, p. 13, 2005.
[47] C. Deck and D. Porter, "Prediction Markets in the Laboratory," *J. Econ. Surv.*, p. n/a–n/a, 2013.
[48] P. Buckley, J. Garvey, and F. McGrath, "A case study on using prediction markets as a rich environment for active learning," *Comput. Educ.*, vol. 56, no. 2, pp. 418–428, Feb. 2011.
[49] D. Raban and D. Geifman, "Designing Online Information Aggregation and Prediction Markets for MBA Courses," *Interdiscip. J. E-Learn. Learn. Objects*, vol. 5, pp. 247–262, 2010.