Argumentation Logic as a Formal Foundation for Collective Intelligence Systems: a Roadmap

Anna Zamansky, University of Haifa

Ofer Arazy, University of Alberta and University of Haifa

Collective intelligence is shared (or group) intelligence that emerges when aggregating the collective efforts of many individuals (Woolley et al., 2010). The success of collective intelligence (CI) systems rely heavily on the effectiveness of the aggregation mechanism (Surowiecki, 2005). Individuals may contribute a simple data element (for instance, a number representing an estimated or predicted value), whereas the aggregation mechanisms is quite simple (group's statistical average, voting, or a prediction market); alternatively, some CI systems require individuals to contribute much more complex knowledge (for example when posting content to Wikipedia or when reporting a sightseeing to a citizen science environmental monitoring system (Bonney et al., 2009)), requiring more sophisticated aggregation mechanisms (Lévy and Bonomo, 1999). The focus of the current investigation are those CI systems that require the contribution of rich (often multi-dimensional) knowledge elements and therefore must employ complex aggregation mechanisms, often acting as a socio-technical system that involves both human actors and automated tools.

For an example of a collective intelligence system, consider a citizen science system where individuals use mobile devices to report on wildlife sightseeing in the African savanna; a report includes – in addition to a photo or video - attributes such as GPS coordinates, date & time, current weather, and supports the contributor in identifying the exact species. Suppose that contributor A provides a blurry photo of wildlife and identifies the species as 'tiger'; contributor B responds to A's post by challenging the species identification and by providing a link to a reliable information source which states that tigers do not live in that habitat. Contributor C, a known expert, makes a correction and identifies the species in the photo as 'Chita'. Another example is a participatory planning platform, where citizens work together (often in collaboration with the authorities) to develop plans that tackle various issues (e.g. global warming). They may share different ideas, refine other's suggestions, and collaboratively develop a proposal. Despite these examples representing very different applications of collective intelligence, they share some fundamental interaction patterns: a contributor posts a piece of information (reports an observation, proposes an idea), while others comment on that information element (e.g. expressing support, adding missing information, or challenging its accuracy), creating a complex network of information elements (nodes) and their relations (arcs). In such applications, the goal of a CI platform is to aggregate and synthesize individual's contributions in a *coherent* way. However, determine the coherence of a system of inter-linked knowledge elements is not straightforward. To date, the arduous and error-prone task of aggregation and synthesis of individual knowledge elements within CI is primarily performed manually. The sheer size and complexity of these systems require automated support tools and such automation calls for formal foundations of CI.

Our thesis is that formal foundations can be utilized in CI systems for: (I) constructing tools and interfaces to support contributors' interactions and (II) developing mechanisms for aggregating individuals' information elements, determining the qualities of the system as a whole**.** In particular, we argue that the logic of argumentation naturally lends itself to interaction patterns typical of CI.

Argumentation theory (Van Eemeren et al., 2013) is an interdisciplinary field concerned with the construction of evaluation of arguments and their interactions. Prakken and Sartor (2002) list three layers in an argumentation process: (i) *Logical layer*: concerned with the rules for constructing arguments; defines what arguments are and how pieces of information can be combined to provide basic support for a claim, (ii) *Dialectical layer*: focuses on conflicting arguments and introduces notions such as counter-argument, attack, rebuttal, etc., (iii) *Procedural layer*: regulates how an actual dispute can be conducted; defines the possible speech acts, and the discourse rules governing them.

There are clear connections between the layers of argumentation theory and the domain of collective intelligence. Indeed, the *procedural* and *dialectical* layers have received quite a bit of attention within the research community in areas related to CI. However, to the best of our knowledge, the *logical* layer, which is associated with formal rules for sense-making and is the focus of this short paper, has received less attention. In logic-based approaches, argumentation is expressed as a network of arguments (i.e. propositions, statements) and relations between them. As the baseline for our investigation, we start with Dung's abstract argumentation framework (AF) (1993, 1995), which is one of the most well-known logical approaches. In this framework *arguments* are treated as abstract entities related only by a binary *attack* relation, capturing the situation where one argument undermines the credibility of another. Many extensions of this framework have been proposed, utilising concepts such as preferences, as well positive interactions (support) between arguments in addition to the negative ones (attack). The aim of this position paper is to propose some basic principles for the adaptation of logic-based approaches in the spirit of argumentation frameworks to the context of CI, paying particular attention to the nature of contributions in collective intelligence systems and to the aggregation mechanisms employed.

Requirements for Argumentation Logic to Support Collective Intelligence Systems

Our analysis of existing collective intelligence systems has identified the following key requirements:

- **Basic entities.** Basic entities in CI are an '*information elements*' (*IEs*), i.e., the knowledge elements posted by contributors. IEs are often multi-dimensional data structures, containing several *attributes*, where each attribute may be assigned a value (e.g. date, geographical location).
- **Basic binary relations between IEs.** Commenting on another contribution, rating or placing a \`thumbs up' next to it puts the original IE in a binary relation with the new feedback. We identify the following properties of relations as necessary for the CI context:
    - Different *types of relations*, such as '*attack*' or '*support*' (in contrast to having only one *attack* relation in Dung's framework).
    - Different *attributes of relations* (e.g., *capacity*, or strength, weight associated with the authority of the source or his confidence level).
    - Possibility to define *relations at the attribute level,* and not only IE level (e.g. user A makes a contribution reporting seeing elephants in a particular site and time, and B attacks the time attribute arguing that elephants are not active at this time of day).
- **Derived binary relations between IEs.**
    - Some *IEs could be seen as related due their 'proximity'*, which could be determined based on differences in attributes' values. E.g.,, observations taken minutes apart at close places may be deemed as related .
    - Various options for the *propagation of IE relations* should be considered. In contrast to Dung's framework, an attack of an attack may not entail a support (e.g., A originally reports of seeing a tiger, B 'attacks' by A, stating the observed animal is a cheetah, and C 'attacks' B identifying the species as a leopard; clearly, C does not support A). Other examples: support of a support entails a support, support of an attack entails an attack.
- M**ulti-ary relations between IEs.** Relations linking several IEs, to be associated with higher-level decision-making. In the participatory journalism example, an editor may point to several postings associated with the same news story (IEs that complement and support one another).
- **System-level properties**. By processing the interrelated elements described above, system –level properties such as coherence emerge, representing the qualities of the "collective Intelligence".

Moving Forward: a Roadmap

Below briefly we provide a roadmap for further research, identifying relevant prior studies:

Delineating the nature of IEs and their attributes. While keeping a high level of abstraction in the spirit of Dung, for developing useful CI foundations, we need to incorporate domain-specific structure into observations and the relations between them. First, *attributes* of observations play an important role in CI and need to be explicitly incorporated. Some related works discuss attributes in this context (for example Chesñevar et al. (2006)), yet we are not aware of works that explicitly incorporate them.

Delineating types of relations. A criticism often advanced against Dung frameworks is their restricted expressive capability of allowing only attacks between arguments. This had led to quite a number of extensions, such as attacks from sets of arguments (Nielsen and Parsons, 2007), attacks on attacks (Modgil, 2009) and meta-argumentation (Boella et al., 2009). A more general model of abstract dialectical frameworks (Brewka and Woltran, 2010) unifies the various relations between arguments while remaining on the same abstraction level as Dung's framework. Less formal frameworks could also provide some insights regarding the possible types of relations. Argumentation schemes (Walton, 2013), an approach emerging from informal logic, and could be utilized to identify the ways in which arguments can be attacked and defended. Also relevant to the analysis of relation types are ontologies of argumentation (Chesñevar et al., 2006), studies that have applied Speech Act Theory to categorize interactions in online collaboration (Convertino et al., 2008, Borge et al., 2012, Bender et al., 2011, Maryam, 2013), and tools developed to support the interactions on collective deliberations (Heras et al., 2013, Wyner et al., 2012, Iandoli et al., 2014, Kirschner et al., 2003)

Another aspect that should be explored in this context is degree-based relations. Works exploring fuzzy argumentation frameworks that extend Dung' (1993) approach allow the representation of the relative strength of the (attack) relationships between arguments (Janssen et al., 2008) or the degree of trustworthiness of information sources (da Costa Pereira et al., 2011).

Deriving system level properties. The key to automatic support for aggregating collective intelligence is the ability to fit together a large collection of often conflicting pieces of information into a *coherent* "big picture". While the study of coherence is in the center of many disciplines, such as psychology, law, AI and cognitive science, a widely accepted formal definition has yet to be provided. Several computational models of coherence have been proposed. For example, Thagard and Verbeurgt (1998) define coherence in terms of maximal satisfaction of competing constraints. Elements in this model can cohere (or fit together) or incohere in various ways, imposing positive and negative constraints. The coherence problem then consists of dividing a collection of elements into accepted and rejected sets in a way that maximizes coherence. To date, the fields of logic-based argumentation and cognitive theories of coherence have so far progressed in parallel with very little interaction (see (Pasquier et al., 2006) for an exception). CI may be a particularly suitable domain to explore the interconnections between them, as semi-automatic tools, bridging between cognitive and algorithmic approaches to coherence may be particularly useful in CI platforms.

In conclusion, we have made the case for developing argumentation-based formal foundations to support CI platforms. Moreover, based on our analysis of several existing CI systems, we have articulated a set of requirements for such foundations. Beyond their usefulness in structuring interactions and informing the design of user interfaces, such foundations could facilitate (semi-)automatic reasoning about system-level properties (such as coherence), serving as a "reasoning engine" of generic CI platforms.

References:

Bender, E. M., et al. Annotating social acts: Authority claims and alignment moves in Wikipedia talk pages. Proc. of the Workshop on Languages in Social Media, 2011. ACL, 48-57.

Boella, G., Gabbay, D. M., Van Der Torre, L. & Villata, S. 2009. Meta-argumentation modelling I: Methodology and techniques. *Studia Logica,* 93**,** 297-355.

Bonney, R. et al. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience,* 59**,** 977-984.

Borge, M., Ganoe, C. H., Shih, S. I. & Carroll, J. M. Patterns of team processes and breakdowns in information analysis tasks. Proc. of CSCW2012, Seattle, WA. ACM, 1105-1114.

Brewka, G. & Woltran, S. Abstract dialectical frameworks. Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, 2010.

Chesñevar, C., Modgil, S., Rahwan, I., Reed, C., Simari, G., South, M., Vreeswijk, G. & Willmott, S. 2006. Towards an argument interchange format. *The Knowledge Engineering Review,* 21**,** 293-316.

Convertino, G., Mentis, H. M., Rosson, M. B., Carroll, J. M., Slavkovic, A. & Ganoe, C. H. Articulating common ground in cooperative work. Proc. of CHI2008. ACM, 1637-1646.

Da Costa Pereira, C., Tettamanzi, A. G. & Villata, S. Changing one's mind: Erase or rewind? possibilistic belief revision with fuzzy argumentation based on trust. Proc. of AI2011. 164-171.

Dung, P. M. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning and Logic Programming. IJCAI, 1993. 852-857.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence,* 77**,** 321-357.

Heras, S., Atkinson, K., Botti, V., Grasso, F., Julián, V. & Mcburney, P. 2013. Research opportunities for argumentation in social networks. *Artificial Intelligence Review,* 39**,** 39-62.

Iandoli, L., Quinto, I., De Liddo, A. & Buckingham Shum, S. 2014. Socially augmented argumentation tools. *International Journal of Human-Computer Studies,* 72**,** 298-319.

Janssen, J., De Cock, M. & Vermeir, D. Fuzzy argumentation frameworks. Information Processing and Management of Uncertainty in Knowledge-based Systems, 2008. 513-520.

Kirschner, P. A., Buckingham-Shum, S. J. & Carr, C. S. 2003. *Visualizing argumentation: Software tools for collaborative and educational sense-making*, Springer.

Lévy, P. & Bonomo, R. 1999. *Collective intelligence: Mankind's emerging world in cyberspace*, Perseus.

Maryam, T. 2013. Dialogue act recognition in synchronous and asynchronous conversations.

Modgil, S. 2009. Reasoning about preferences in argumentation frameworks. *AI,* 173**,** 901-934.

Nielsen, S. H. & Parsons, S. 2007. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. *Argumentation in multi-agent systems.* Springer.

Pasquier, P., Rahwan, I., Dignum, F. & Sonenberg, L. 2006. Argumentation and persuasion in the cognitive coherence theory. *Frontiers in artificial intelligence and applications,* 144**,** 223.

Prakken, H. & Sartor, G. 2002. The role of logic in computational models of legal argument: a critical survey. *Computational logic: Logic programming and beyond.* Springer.

Surowiecki, J. 2005. *The Wisdom of Crowds*, Random House Digital, Inc.

Thagard, P. & Verbeurgt, K. 1998. Coherence as constraint satisfaction. *Cognitive Science,* 22**,** 1-24.

Walton, D. 2013. *Argumentation schemes for presumptive reasoning*, Routledge.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. & Malone, T. W. 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science,* 330**,** 686-688.

Wyner, A., Wardeh, M., Bench-Capon, T. J. & Atkinson, K. A Model-Based Critique Tool for Policy Deliberation. JURIX, 2012. 167-176.