# On Scientific Enquiry and Computational Argumentation

**Federico Cerutti**[1]

**Abstract.** In this speculative paper we discuss how existing work in formal argumentation can support the creation of a *Regulæ Philosophandi Ratiocinator*, i.e. a machinery implementing general principles of formal science. In particular, we review two research projects in this light, one aimed at supporting intelligence analysis—CISpaces.org—and one aimed at assessing natural language interfaces to formal argumentation.

## 1 INTRODUCTION

Epistemology is central to western philosophy: the pessimistic cave story of Plato as well as the optimistic of *anamnesis* are examples of it. When it comes to using computer science to support epistemology we cannot avoid to look at Leibniz and his *Caluculs Ratiocinator*—e.g. [3, p. 654]—as a precursor of several approaches aimed at creating a language for representing every piece of available knowledge and then applying logical reasoning to infer new knowledge. While extremely powerful in specific contexts, those approaches are not widely adopted in scientific enquiry due to their general lack of robustness against highly uncertain and only partially observable phenomena.

This paper is in favour of a *Regulæ Philosophandi Ratiocinator* (cf. [9, p. 387]) that implements modern and widely adopted theories of epistemology. In particular, according to Popper, the advancement of scientific knowledge is based upon a process of conjecture and refutation [10], an inherently argumentative process. In Section 2 we summarise the CISpaces.org project [13] and show how existing theories of computational argumentation can already provide (limited) support for scientific enquiry in real domains. While we abstain from discussing approaches to argument mining [8], language clearly plays a role in formulation of theories. However, "although clarity is valuable in itself, exactness or precision is not [. . . ]. Words are significant only as instruments for the formulation of theories, and verbal problems are tiresome: they should be avoided at all cost" [10, p. 28]. In Section 3 we summarise an experiment assessing natural language interfaces to formal argumentation [1] and criticise it.

## 2 SCIENTIFIC METHODS OF INTELLIGENCE ANALYSIS: CISPACES

In [13] we discuss how the process of representing reasoning lines using argumentation schemes [15] and structured argumentation techniques—in particular a simplified version of ASPIC [7]—supports, with the help of efficient algorithms for computing semantics extensions [2], the process of sensemaking, by complementing human expertise in the generation of intelligence products.

[1]Cardiff university, UK, CeruttiF@cardiff.ac.uk

CISpaces.org, based upon [13], facilitates the core phase of sensemaking within the intelligence analysis process in a declarative format. Intelligence analysis is an iterative process of foraging for information and sensemaking in which the analysis structure increases incrementally from a shoebox of information, through evidence files, to the generation and evaluation of hypotheses.

CISpaces.org therefore supports—yet not guide—the analyst in a process of conjecture—of hypotheses—and refutation, based on critical questions and other known arguments/facts through an assessment of their acceptability status. Although CISpaces.org has been developed for addressing tactical situational understanding problems—in particular answering questions associated to *who* did *what*, *when*, *where*, and *to which purpose*, hence linking causes to effects and evaluating competing hypotheses—in [5] we showed its flexibility by analysing the case of Prosecutor v. Karadžić (MICT-13-55-A) in front of the UN International Criminal Tribunal.

Let us consider here a simpler case. In [14] (now retracted), Wakefield et al. present an early report investigating the case of 12 children experiencing a loss of acquired skills, including language, where "onset of behavioural symptoms was associated, by the parents, with measles, mumps, and rubella [MMR ed.] vaccination in eight of the 12 children" (conjecture).

This paper triggered a larger study (a cohort of 537,303 children) summarised in [4] where it it is shown that "There was no association between the age at the time of vaccination, the time since vaccination, or the date of [MMR ed.] vaccination and the development of autistic disorder" (refutation).

While in principle those arguments can be formalised and handled by CISpaces.org, thus supporting in part the claim that computational argumentation can be of benefit for scientific enquiry, they also highlight the need for further studies. The class of argumentation schemes used in scientific enquiry is only partially overlapping with those analysed in [15]. For instance, the findings discussed in [4] heavily rely on results from statistics that should be further represented in form of arguments. While in some cases arguing on the basis of the results of statistical tests or probabilistic inferences without further discussions can be acceptable [16], in other cases a deeper analysis [12] might be necessary. This in general depends on the audience of the analysis, which manifests the need for communicating arguments.

## 3 NATURAL LANGUAGE INTERFACES TO FORMAL ARGUMENTATION

The use of graphical models is the most common approach used in the formal argumentation community to capture argumentative structures. However they require a significant levels of training and resource to be produced and consumed. Instead of training users on another (graphical) language for representing argumentative structures, we can leverage our societal model, through which we are trained in reading and writing; that is, using natural language.

It is germane to consider one specific communicative goal, namely explaining the acceptability status of arguments in a given argumentation framework. In an experiment described in [1], we investigate this communication goal with a Wizard of Oz approach—hence manually generating different pieces of texts—albeit not in a scientific enquiry context.

The experiment consists of presenting participants with texts, written in natural language, followed by a questionnaire. Texts provide natural language interfaces to the following knowledge base formalised using [11]: $\Gamma = \langle S, D \rangle$ with $S = \{s_1 :\Rightarrow s_{AAA};\ s_2 :\Rightarrow s_{BBB}\}$ and $D = \{r_1 : s_{AAA} \wedge \sim ex_{AAA} \Rightarrow c;\ r_2 : s_{BBB} \wedge \sim ex_{BBB} \Rightarrow \neg c;\ r_3 :\sim ex_{pref} \Rightarrow r_1 \prec r_2\}$. According to [11] $\Gamma$ gives rise to the following set of arguments: $\mathcal{A}rgs = \{\mathbf{a}_1 = \langle s_1, r_1 \rangle, \mathbf{a}_2 = \langle s_2, r_2 \rangle, \mathbf{a}_3 = \langle r_3 \rangle\}$ where $\mathbf{a}_2$ $\mathcal{A}rgs$-defeats $\mathbf{a}_1$, hence the set of justified arguments is $\{\mathbf{a}_2, \mathbf{a}_3\}$.

We generate texts for four different domains, including weather forecast (derived from an example discussed in [6]):

> The weather forecasting service of the broadcasting company AAA says that it will rain tomorrow. Meanwhile, the forecast service of the broadcasting company BBB says that it will be cloudy tomorrow but that it will not rain. It is also well known that the forecasting service of BBB is more accurate than the one of AAA.

Participants then are asked to determine which of the following positions they think is accurate: *Tomorrow will rain* ($\mathcal{P}_{\mathbf{A}}$); *Tomorrow will not rain* ($\mathcal{P}_{\mathbf{B}}$); *I cannot conclude anything tomorrow weather* ($\mathcal{P}_{\mathbf{U}}$).

Our hypothesis is that the majority of participants would agree with BBB's statement (*position* $\mathcal{P}_{\mathbf{B}}$). While such an hypothesis is supported in general, in the case of weather forecast the actual percentages of agreement are: 5% for $\mathcal{P}_{\mathbf{A}}$; 50% for $\mathcal{P}_{\mathbf{B}}$; and 45% for $\mathcal{P}_{\mathbf{U}}$. When asked to justify their position, the majority of participants supportive of $\mathcal{P}_{\mathbf{U}}$ refers to the fact that weather forecast naturally carries high uncertainty.

An interesting question—that we hope to answer soon—is whether significantly different results could be obtained using modified textual representations, such as the following:

> Tomorrow will be cloudy but not rainy according to BBB; opposite forecast is provided by AAA, known to be untrustworthy.

Indeed, critiques to this experiment can be: (1) different formulations might very well serve the same communicative goal; (2) different audience will require different level of explanation.

Regarding the first aspect, in CISpaces.org we adopt a rather pragmatic approach. The output of its natural language generation system in the given case would be[2] the following:

> We have reasons to believe that:
> - tomorrow should be cloudy but not rainy
>
> Here the pieces of information we received:
> - The weather forecasting service of the broadcasting company AAA says that it will rain tomorrow.
> - The forecast service of the broadcasting company BBB says that it will be cloudy tomorrow but that it will not rain.
> - It is well known that the forecasting service of BBB is more accurate than the one of AAA.

[2]Currently CISpaces.org does not handle preferences (yet) hence the hypothetical sentence.

Regarding the second aspect, it is worth noticing that the above three natural language interfaces expose neither the reasoning lines linking premises to conclusion, e.g. "The weather forecasting service of the broadcasting company AAA says that it will rain tomorrow, *therefore tomorrow should rain*;" nor the reasoning lines leading to computing the acceptability status of arguments, e.g. "*Since AAAs and BBBs conclusions are incompatible, and since BBB is more accurate than AAA, it is reasonable to tentatively conclude that tomorrow should not rain*."

## 4 CONCLUSION

In this speculative paper we argue that scientific enquiry can be supported by formal argumentation, that is uniquely equipped to implement the process of conjecture and refutation discussed by Popper [10]. While we base our speculation only on anecdotal evidence, notably adaptations of CISpaces.org and of previous experiments on natural language interfaces to formal argumentation, they seem convincing enough to suggest that we can soon be equipped to build a *Regulæ Philosophandi Ratiocinator*, a machinery implementing general principles of formal science.

## REFERENCES


[1] F. Cerutti, N. Tintarev, and N. Oren, 'Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation', in *ECAI 2014*, pp. 207–212, (2014).

[2] F. Cerutti, M. Vallati, and M. Giacomin, 'jArgSemSAT: An Efficient Off-the-shelf Solver for Abstract Argumentation Frameworks', in *KR 2016*, pp. 541–544, (2016).

[3] G. W. Leibniz, *Philosophical Papers and Letters*, Springer Netherlands, Dordrecht, 1976.

[4] K. M. Madsen, A. Hviid, M. Vestergaard, D. Schendel, J. Wohlfahrt, P. Thorsen, J. Olsen, and M. Melbye, 'A Population-Based Study of Measles, Mumps, and Rubella Vaccination and Autism', *New England Journal of Medicine*, **347**(19), 1477–1482, (2002).

[5] Y. McDermott Rees and F. Cerutti. Request for leave to make submissions as amicus curiae. `https://goo.gl/PKHRHM`, 2018.

[6] S. Modgil, 'An argumentation based semantics for agent reasoning', in *LADS2008*, pp. 37–53, (2008).

[7] S. Modgil and H. Prakken, 'A general account of argumentation with preferences', *Artificial Intelligence*, **195**, 361–397, (2013).

[8] M.-F. Moens, 'Argumentation mining: Where are we now, where do we want to be and how do we get there?', in *FIRE '13*, (2013).

[9] I. Newton, *Philosophiæ naturalis principia mathematica*, Apud G. & J. Innys, 3rd edn., 1726.

[10] K. R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge, 5th edn., 1989.

[11] H. Prakken and G. Sartor, 'Argument-based extended logic programming with defeasible priorities', *Journal of Applied Non-Classical Logics*, **7**(1-2), 25–75, (1997).

[12] S. T. Timmer, J.-J. C. Meyer, H. Prakken, S. Renooij, and B. Verheij, 'A two-phase method for extracting explanatory arguments from bayesian networks', *International Journal of Approximate Reasoning*, **80**, 475 – 494, (2017).

[13] A. Toniolo, T. J. Norman, A. Etuk, F. Cerutti, R. W. Ouyang, M. Srivastava, N. Oren, T. Dropps, J. A. Allen, and P. Sullivan, 'Agent Support to Reasoning with Different Types of Evidence in Intelligence Analysis', in *AAMAS 2015*, pp. 781—-789, (2015).

[14] A. Wakefield, S. Murch, A. Anthony, J. Linnell, D. Casson, M. Malik, M. Berelowitz, A. Dhillon, M. Thomson, P. Harvey, A. Valentine, S. Davies, and J. Walker-Smith, 'Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children', *The Lancet*, **351**(9103), 637–641, (1998).

[15] D. Walton, C. Reed, and F. Macagno, *Argumentation schemes*, Cambridge University Press, NY, 2008.

[16] M. Williams, Z. W. Liu, A. Hunter, and F. Macbeth, 'An updated systematic review of lung chemo-radiotherapy using a new evidence aggregation method', *Lung Cancer*, **87**(3), 290 – 295, (2015).