Available online at www.sciencedirect.com
ScienceDirect
Artificial Intelligence 171 (2007) 619–641
Artificial Intelligence
www.elsevier.com/locate/artint

Argumentation in artificial intelligence

T.J.M. Bench-Capon, Paul E. Dunne *

Department of Computer Science, University of Liverpool, Liverpool, United Kingdom

Received 27 April 2007; received in revised form 27 April 2007; accepted 1 May 2007
Available online 10 May 2007

Abstract

Over the last ten years, argumentation has come to be increasingly central as a core study within Artificial Intelligence (AI). The articles forming this volume reflect a variety of important trends, developments, and applications covering a range of current topics relating to the theory and applications of argumentation. Our aims in this introduction are, firstly, to place these contributions in the context of the historical foundations of argumentation in AI and, subsequently, to discuss a number of themes that have emerged in recent years resulting in a significant broadening of the areas in which argumentation based methods are used. We begin by presenting a brief overview of the issues of interest within the classical study of argumentation: in particular, its relationship—in terms of both similarities and important differences—to traditional concepts of logical reasoning and mathematical proof. We continue by outlining how a number of foundational contributions provided the basis for the formulation of argumentation models and their promotion in AI related settings and then consider a number of new themes that have emerged in recent years, many of which provide the principal topics of the research presented in this volume.

Keywords: Argumentation models; Dialogue processes; Argument diagrams and schemes; Agent-based negotiation; Practical reasoning

1. Introduction

In its classical treatment within philosophy, the study of argumentation may, informally, be considered as concerned with how assertions are proposed, discussed, and resolved in the context of issues upon which several diverging opinions may be held. Thus philosophical investigations of argumentation, from Aristotle to the present day, have addressed such themes as: the mechanisms by which "legitimate" argumentation in support of a claim may be distinguished from "flawed" argumentation; analyses of the typical structures that constitute argument components and argumentation development; the processes by which participants engaging in debate may advance their respective positions and undermine contrary stances and arguments, etc; and the contexts in which these questions are decided. The importance of such philosophical theories to so-called everyday reasoning has a long and distinguished history in AI, and contributions from contemporary philosophical analyses continue to play a major role in the evolution of effective computational exploitation of argumentation technology.

Within the simplified overview of argumentation outlined in the preceding paragraph, one can, already, identify a number of themes whose elements embody issues of a computational nature in the following:

* Corresponding author.
E-mail address: ped@csc.liv.ac.uk (P.E. Dunne).

0004-3702/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.artint.2007.05.001

- Defining the component parts of an argument and their interaction.
- Identifying rules and protocols describing argumentation processes.
- Distinguishing legitimate from invalid arguments.
- Determining conditions under which further discussion is redundant.

It is, of course, the case that similar issues underpin one well-established and highly-developed theory: that of formal logic and mathematical proof. It is no coincidence that much of the formal computational treatment of argumentation has its roots in ideas developed from AI inspired contributions to logic and deductive reasoning. So one finds in mathematical proof theory core concepts such as: precisely defined means for expressing assertions (e.g. formulae in a given logical language); accepted bases on which to build theorems (e.g. collections of axioms); procedures prescribing the means by which further theorems may be derived from existing theorems and axioms (e.g. templates for inference rules); and precise concepts of termination (e.g. a sentential form is derivable as a theorem, "true"; or is logically invalid, "false").

While the structural elements presented in this view of mathematical reasoning have proven to be a useful basis in the development of argumentation-based models in AI, the formal apparatus and methods of mathematical reasoning are, ultimately, radically different in nature to those of importance when considering the concept of argumentation as it is familiar from everyday contexts, e.g. as it might occur in political debate, the discussion of ethical principles, deliberation in judicial settings, etc. While there are, of course, parallels that can be made,—e.g. that those engaged in debate have some collection of accepted premises on which there is agreement, possibly, even, some recognition of when contributions to a discussion are "unreasonable" or flawed, etc.—there are, however, a number of fundamental distinctions between the concepts "P is a formal proof that T holds" and "P is a persuasive argument for accepting T". Thus, in mathematical reasoning,

(a) The premises can, ultimately, be explicitly defined in terms of closed concepts, e.g. the axioms of Euclidean geometry, the Zermelo–Frankel basis for set theory (ZF). Furthermore classical mathematical reasoning is based on an assumption that such premises are, collectively, *consistent*.[1]
(b) Reasoning and analysis takes place within a closed, tightly defined context, i.e. there is no notion of "incomplete" or "uncertain" information.
(c) Conclusions are final and definite: if P *is* a correct proof that T, then T is, *ipso facto* valid and this status does not admit subsequent qualification or amendment, let alone retraction.
(d) Reasoning and conclusions are entirely *objective*, not susceptible to *rational* dispute on the basis of subjective views and prejudices.[2] Proof is demonstration whereas argument is persuasion.

In argument and discussion as encountered in everyday contexts, it is rare that any, let alone all, of these apply: the premises upon which debates may build are often presupposed as forming part of the background assumptions common to all parties involved; the information and knowledge brought to bear in the course of discussion will often be incomplete, vague, or uncertain. The remaining two aspects, in many ways, highlight the most significant differences between "logical proof" and "persuasive argument". Arguments are *defeasible*: the reasoning that formed a persuasive case for T, in the light of changes in viewpoint or awareness of information not previously available, may subsequently fail to convince. This defeasibility is never removed: an argument may cease to be challenged and so accepted, but the *possibility* of challenge remains. Finally, the extent to which an argued case is accepted is *subjective*, dependent on the views, attitudes, and prejudices of the audiences to which it is directed. The same case may convince some people but, equally, fail to convince others.

[1] We note that in a number of systems, consistency *cannot* be formally proven, cf. [95] and so, in such cases, consistency is, indeed, an *assumption*.

[2] Some clarification of this claim may be in order. Suppose Δ is a derivation of φ within a theory $\langle A, R\rangle$ (with axioms A and inference rules R). Within the *same* theory, the proof Δ admits no rational, objective basis for dispute: criticisms that "φ is 'inconvenient' or 'counter-intuitive'" are subjective, and entirely irrelevant to the status of φ within the theory $\langle A, R\rangle$. In order to give rational grounds for not accepting φ it is necessary to endorse an alternative theory within which φ *cannot* be derived. As a concrete example, consider the axiomatic basis ZF extended by the so-called "Axiom of Choice" (ZF + AC): although widely adopted in modern theory this conflicts with Intuitionist principles which disqualify AC as an axiom so that theorems dependent on AC are (rationally) not accepted by Intuitionists.