Argument Understanding and Argument Choice
A Case Study

Helmut Horacek[1]

Abstract. Several categories of discourse moves and sequences have been studied in formal models of disputes. However, most of these models make two simplifications that neglect important factors in argumentation: (1) raising an argument is typically done by introducing one or several new facts in the dispute, assuming that the associated warrant is self-evident, and (2) variants of arguments addressing the same issue are rarely assessed in terms of their benefits and drawbacks. In this paper, we illustrate these two points by studying the role of alternative arguments in explaining the solution to a rather simple, but not so easily understandable problem. Arguments may differ in terms of the effort needed to communicate them, the confidence they achieve, and requirements on knowledge of the audience, which makes their relative benefit task- and context-dependent.

1 INTRODUCTION AND MOTIVATION

In the literature, several categories of argumentative moves have been studied in formal models of disputes, including arguments based on perception, statistics, and causality (see the sources of *prima facie* reasons in [7]). Arguments are examined in terms of their logical grounding [7], their role and contribution to progress in the discourse [6], and their potential to defend against attacks as raised by critical questions in argumentation schemata [10]. However, most models of argumentation include simplifications concerning the comprehensibility and variation of arguments. On the one hand, raising an argument is typically done by introducing one or several new facts in the dispute, assuming that the associated warrant is self-evident. Making the underlying reasoning more precise and explicit aims at uncovering implicit assumptions and potential sources for critical questions rather than addressing the comprehensibility of an argument. On the other hand, alternatives in arguments addressing the same issue are rarely considered, although benefits and drawbacks may vary significantly among possible tasks and contexts. We are convinced that studying these factors is likely to improve the understanding of driving forces underlying natural argumentation and associated skills significantly.

In this paper, we address the role of knowledge and purpose in argument choice in a case study, by examining the role of several categories of arguments in explaining the solution to the so-called *goat problem*. This problem constitutes a superficially simple task, but this task is not easily understandable at first, so that it gives rise to a variety of arguments providing sources of explanations. Arguments may differ in terms of the effort needed to communicate them, the confidence they achieve, and requi-

rements on knowledge of the audience. Typical scenarios where the choice among such arguments and their presentation plays a prominent role include teaching reasoning in tutorial systems and argumentation within qualitative economic models.

This paper is organized as follows. First, we introduce the goat problem and its solution. Then we describe variants of arguments justifying that solution over the typically occurring misconception and discuss benefits and drawbacks. Finally, we sketch an operationalization of these concepts.

2 RUNNING EXAMPLE – THE *GOAT PROBLEM*

The goat problem is a superficially simple problem that originates from a game show. The problem comprises two consecutive guesses to be made by a candidate, with an apparently hidden dependency. The scenario consists of three doors, a car, and two goats. Behind each of the doors there is either the car or one of the goats, and the goal of the candidate is to guess where the car is (see Figure 1). In the starting position, the candidate makes an – apparently arbitrary – guess and picks one of the doors behind which he hopes the car being located. Then the showmaster opens one of the other two doors, unveiling one of the goats behind this door. Then the candidate is to make the second and final choice, in which he can stick to his original guess or alter it. The crucial question in the whole problem is whether one of these alternatives is superior to the other – and why – or whether the second choice offered is also a pure guess.

When confronting people with this problem, it turns out that not only finding but even understanding the solution is surprisingly difficult. The overwhelming majority of people unfamiliar with the problem believes that both alternatives in the second choice have the same likelihood to win, but this view is simply wrong. In contrast, changing the original choice is superior by a significant margin, winning two out of three times per average. The reason basically lies in the difference between the situation when the candidate first picks the door with the car behind it and the complementing situations when the candidate first picks a door with a goat behind it. In the second case, the showmaster has *no* choice, since he must present the only remaining goat and open the door in front of it. In the first case, however, the showmaster can pick any of the two remaining doors, and we can assume that he takes one or the other with equal likelihood. Hence, the second case occurs twice as often as the first case, so that altering the original choice is significantly superior.

---
[1] Universität des Saarlandes, FB 14 Informatik, Postfach 1150, D-66041 Saarbrücken, B.R.D., email: horacek@ags.uni-sb.de
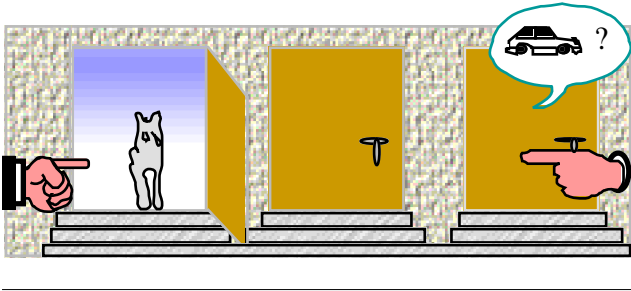
Figure 1. Running example scenario – the goat problem

3 CATEGORIES OF ARGUMENTS

Despite its superficial simplicity, the goat problem proves to be difficult to understand for humans. Therefore, several attempts have been undertaken to find illustrative explanations for the reasoning required to solve the problem. In addition, arguing in favor of the correct solution in a dispute may be of interest. Achieving a concession may not necessarily involve complete understanding on behalf of the other person, although this may also be beneficial for related prupuses, such as strengthening confidence. Consequently, there are several ways of arguing in favor of the solution, including the categories illustrated in Table 1:

1. As in many other situations, the simplest argument type is appeal to expert opinion. While this type of argument appears to be convincing to some extent, the confidence in it is limited, since the goat problem is fully accessible to well-justified logical reasoning.

2. Justification by statistics is probably the most convincing argument – in the given problem, this argument is not defeasible, since the task is to find out about the better chances in general, and not in an individual situation. There is an anecdote, that even a famous researcher in probability theory failed to understand the rationale behind the goat problem and did not believe in the solution. It was only due to simulations carried out by his students, which made his mind change – he knows perfectly well about the likelyhood of deviations from expected outcomes. The contribution to understanding the underlying rationale, however, is also not present in this argument category.

3. An extremely suitable argument is reference to an analogous problem, since a good deal of prior understanding can be exploited in this manner. The goat problem has a perfect counterpart in the game of bridge, namely the problem of restricted choice, where one of the opponents is in a situation that is isomorphic to the situation of the quizmaster in the goat problem. Unfortunately, referring to this analogy requires quite specific expertise, that is, being acquainted with the game of bridge at a non-trivial level. Whenever this argument is meaningfully applicable, its explanatory effect is very high.

	North	
	♠ A 10 x x x	
West ♠ x x ?		♠ J/Q ? East
	♠ K x x x	
	South	

	South	West	North	East
First trick:	♠ K	♠ x	♠ x	♠ J/Q
Second trick:	♠ x	♠ x	♠ ? (A or 10)	

Other things being equal, a singleton jack or queen in the East hand is twice as likely as queen and jack doubleton.

Figure 2. The principle of restricted choice in the game of bridge

4. The most commonly occurring argument is the exposition of a causal reason. In contrast to the other ones, an argument in this category provides a perspective on the rationale behind it, at least on some coarse-gained level. As the example texts in Table 1 demonstrate, even a short version is significantly longer than the arguments in the other three categories.

Since the rationale behind the solution to the goat problem is not easily to understand for humans, most of these arguments, specifically the reference to analogy and the causality exposition can be given in varying degrees of detail, the texts in Table 1 being on the short end of the scale. The analogy reference can also be formulated as a hint ("Consider how the problem of restricted choice in bridge can be related to the goat problem"), with a variety of adds-ons about the relation between the two problems. Moreover, the correspondence between the problems can be elaborated explicitly in an explanation, identifying the quizmaster with the defender playing the honor card in the first round, and mapping the associated ocurrences and decision preferences. Versions of the causality argument can differ even more in terms of detail and perspective, making the significance of the first choice evident, elaborating its consequences.

Like varying degrees of detail in the associated exposition, the suitability of categories of arguments justifying the solution to the goat problem depends on a number of contextual factors. One crucial factor is presence of specific knowledge that is required for using the analogy argument in a meaningful manner. Occasionally, testing the expertise of the audience prior to choosing an argument category may be beneficial to check the applicability of an efficient argument. Another factor is the goal of the discourse, which may range somewhere between the aim of just winning a dispute to the goal of enhancing the experience of the audience, as in a tutorial setting. If "winning" is the primary concern, a "hard" and comparably short argument such as appeal to expert opinion or reference to statistical results is probably preferable. When explanation is the primary concern, such references can only be accompanying arguments to a causally-based exposition. Moreover, this exposition needs to be tailored in an appropriate degree of detail according to the knowledge of the audience. Finally, even when winning a dispute is of some interest, this may be associated with a long-term goal of being

Table 1. Argument categories instantiated for the goat problem

1. Expert assessment
Informed experts recommend to change the original choice.

2. Justification by statistics
Simulations strongly favor changing the original choice.

3. Reference to analogical situation
The choice among the remaining doors works analogously to the problem of restricted choice in the game of bridge.

4. Causal reasoning
Altering the original choice is superior to staying with the original one. When the car is behind the door not previously pointed at, the quizmaster was forced to open the door he did, whereas he had a choice when the car is behind the door the candidate pointed at in his first guess.

Table 2. Argument categories and understanding and confidence

	Understanding	*Confidence*
Expert assessment:	low	reasonable, but limited
Statistical justification:	mediocre	depending on the task
Analogy reference:	depends on related knowledge	
Causal assessment:	depends on thematic knowledge	

assessed as a reliable arguer who deserves confidence. Under such circumstances, investigating in explanations that do not only convince the audience to some degree, but also enhance its understanding of the underlying rationale is likely to bear secondary benefits.

4 TOWARDS AN OPERATIONALIZATION

In most approaches to formal models of natural argumentation, a warrant justifying the inference $p \rightarrow q$ (or, more general $P \vdash q$) is treated as a "unit". When it is introduced in the dispute, it is provisionally accepted, and may be attacked later. The assumption is that the inference itself is understood, otherwise accepting or attacking it is not meaningful. In contrast, we make a crucial distinction between degrees of *understanding* and degrees of *confidence*, to assess the effectiveness of an argument. Sufficient degrees of both components are required to make the argument acceptable.

The confidence in an inference depends primarily on the category of the underlying warrant. For some categories, degrees of understanding are also relevant. In order to address the understanding component in argumentation, we require arguments to be modeled in varying degrees of detail, for use in communication. While it is normally assumed that an argument $P \vdash q$ is also raised in precisely that form, we introduce expansions of arguments that make the underlying derivation more explicit. Thus, communicating an argument can either be done directly by $Say(P \rightarrow q)$, or an expanded form is introduced in the dispute, through $Say(P \nabla_q)$ where $P \nabla_q$ is a derivation tree underlying the argument $P \vdash q$ that makes some of its intermediate results explicit.

Exposing arguments in appropriate degrees of detail to meet the mental capabilities of an audience is a common topic at the intersection of the areas of deductive system and natural language presentation. Arguments in communication are frequently much more concise than in a mechanical proof [1], exploiting discourse expectations and background knowledge [3], which also holds for everyday discourse in comparison to underlying

logical patterns [2, 8]. In contrast, some cognitively difficult reasoning patterns, such as modus tollens and disjunction elimination need to be exposed in more detail in order to support proper understanding [5, 9]. Hence, there are significant variations in terms of degrees of detail, which strongly influence degrees of comprehension, in accordance with the purpose of an expository explanation (full-depth, summary, sketchy idea [4]).

Based on these options, there are several factors which contribute to assessing the effectiveness of an argument, when raised in some chosen degree of detail:

• Degrees of *confidence* in the argument
• Degrees of *understanding* of the argument
• *Communicative effort* needed to expose the argument
• *Learning* of inferences through a detailed exposition

The last factor constitutes a kind of "investment" in subsequent sections of the dispute, with the idea that increasing the understanding of the other conversant may enable the beneficial use of causal or even analogical arguments with less communicative effort. The communicative effort is proportional to the size of the derivation tree that corresponds to the degree of detail in which the argument is to be presented. The degrees of understanding and confidence depend on the argument category, as sketched in Table 2. For an argument appealing to experts opinion, the degree of understanding is generally low, since a deeper understanding would require expertise. Moreover, a certain, but limited degree of confidence is present, in comparison to easier understandable arguments. Moreover, the degree of confidence depends on whether there is general agreement among experts about the issue at stake, or whether the expert opinion referred to is challenged by others. For an argument relying on statistics, the degree of understanding is similar, but it can be increased when more details are given about how the statistical procedure is used. The degree of confidence, in turn, may be increased when details about the strength of the statistical results are exposed. For the remaining argument categories, reference to analogy and causal assessment, the knowledge accessible to follow the causality in enough detail is the decisive factor. For analogy reference, that knowledge refers to the issue related through the analogy. In contrast to the other categories, the possible range in the degrees of understanding and confidence may vary significantly – they are virtually zero, if the causality (or analogy) is not understood, and maximal in case of full understanding.

In order to select among competing arguments from different categories and with varying degrees of detail, the domain in which the dispute takes place must be elaborated in two ways. Firstly, arguments must be made available in several version

distinguished in their degrees of detail, or a mechanism must be provided which allows for such a construction. Secondly, a user model must be elaborated which allows assessing the knowledge of the other conversant in terms of the items appearing in different versions of arguments. Moreover, on the side of proper argumentation, the benefits of argument categories must be put in a precise relation to each other, including partial success, when arguments are not exposed to the degree of detail needed, as well as some contributions for the communicative effort and for the "learning component". Once these prerequisites are fulfilled, argument selection can proceed according to the following lines: for each argument candidate, the most compact version is picked and evaluated. Those arguments which are assumed not being fully understood by the addressee are successively expanded in relevant aspects according to the variations available. This process is continued for each argument until one of the following holds: (1) no more expansions are possible, (2) the argument is considered comprehensible in the degree of detail considered, or (3) the communicative effort is considered to be on its limit. From all argument versions generated this way, the one that scores best is chosen.

In an advanced version, such a system requires a full-fledged natural language generation approach, at least for text planning, when abstracting from surface realization. The task is then to express a communicative intention – here, making an argument, given a repertoire of alternatives in varying details, to meet assumptions about the intended audience, which in some sense appears to be a classical text planning task. The only extension in terms of assessing the relative merits of the alternatives available lies in judging the role of making an 'investment' through providing detailed expositions, which may make subsequent atrgumentation easier or which may even be necessary to pursue some future line of argumentation. Similar considerations proved to be problematic in dialog systems when playing the role of an agent with certain interest.

5   CONCLUSION

In this paper, we have studied the role of competitive arguments and requirements on knowledge to understand these arguments. In a case study, we have discussed the benefit of arguments in terms of their context and task-dependency, including tutorial purposes, dispute winning, and long-term goals aiming at establishing confidence. In the preliminary state of this work, the associated formalization is still on an abstract level only, that requires task- and domain-specific interpretation for an operational application.

Apparently, the example chosen for our case study is idealized in comparison to real argumentative scenarios. The available choice and variations in detail may be more limited in several realistic situations and, most importantly, arguments might be defeasible or, at least, it may be possible to weaken their strength. Apart from tutorial applications, scenarios where the considerations raised in the paper are important, are discussions with unbalanced levels of expertise, specifically when the role of a referee is more prominent than in most formal models of

dispute. A typical application would be an expert discussion in television, arguing in favor or disfavor of competing strategies, such as economic models to improve the emploiment situation. In formal reconstructions of argumentative situations, such as cases at the court, benefits consist in uncovering implicit assumptions through raising critical questions. In addition to that, formal reconstruction of argumentation in more knowledge-intensive scenarios may also uncover missing knowledge required for following the course of the argumentation, through focusing on warrants that require a more detailed exposition. These additions, in turn, may lead to uncovering more deeply hidden implicit assumptions which improves not only the under-standing, but also the reliabilty of the argumentation.

REFERENCES

[1] R. Cohen. 'Analyzing the Structure of Argumentative Discourse'. *Computational Linguistics* **13**(1-2): 11-24, (1987).

[2] H. Horacek. 'Generating Inference-Rich Discourse Through Revisions of RST-Trees', in *Proc. of AAAI-98,* pp. 814-820, (1998).

[3] H. Horacek. 'Presenting Proofs in a Human-Oriented Way', in *Proc. of CADE-99*, pp. 142-156, Trento, Italy, (1999).

[4] H. Horacek. 'Tailoring Inference-Rich Descriptions Through Making Compromises Between Conflicting Principles'. *International Journal on Human Computer Studies* **53**:1117-1146, (2000).

[5] P. Johnson-Laird, R. Byrne. Deduction. Ablex Publishing, (1990).

[6] H. Prakken. 'On Dialogue Systems with Speech Acts, Arguments, and Counterarguments', in Proc. of *7th European Workshop on Logic for Artificial Intelligence (JELIA'2000),* Springer Lecture Notes in AI (LNAI) 1919, 224-238, Springer, Berlin, (2000).

[7] J. Pollock. 'Defeasible Reasoning'. *Cognitive Science* **11**:481-518, (1987).

[8] M. Thüring, K. Wender. 'Über kausale Inferenzen beim Lesen'. In *Sprache und Kognition* **2**:76-86, (1985).

[9] Marilyn Walker. 'The Effect of Resource Limits and Task Complexity on Collaborative Planning in Dialogue'. In *Artificial Intelligence* **85**:181-243, (1996).

[10] D. Walton. *Argumentation Schemes for Presumptive Reasoning,* Mahwah, N.J., Erlbaum, (1996).