Simulating Knowledge and Dishonesty in a Client-Consultant Setting

Eugen Staab Martin Caminada

University of Luxembourg

In every day life, it is quite common for people to make statements about things they have no proper knowledge of. This is often done out of the desire to appear knowledgeable, even if one in fact is not. The situation here is different from the liar, who tells things he knows to be incorrect. Clearly, lying is not the right word to describe the basic concept here. In the remainder of this paper, statements made without the speaker having sufficient knowledge about their validity will be referred to as "bullshit", sometimes abbreviated to "BS". We use this somewhat provocative term not only for its conciseness, but also to be in line with existing literature [4, 5] and to allow the reader to easily relate the phenomena described in this paper to his every day life experiences. As described in [4], the difference between lies and BS is that with lies, there exists a negative relation to the truth, whereas with BS, there is from the perspective of the speaker no relationship at all between his statements and the truth.

Frankfurt [4] claims that the problem of BS is to some extent caused by the fact that in modern democratic society everyone is supposed to have an opinion about the current social and political issues, even if one does not have the time and means to be properly informed on all relevant aspects. In our view, however, there also exists a more mundane reason. The point is that more and more people started to make a living in professions that aim at generating, processing and providing information. Examples of this are journalists, business consultants, lawyers, financial analysts and even scientists. In these professions, it is vital to appear knowledgeable, even in situations where this is actually not the case. The phenomenal extent to which this happens, as well as its impact on society has been described in [7, 2].

In standard epistemic logic (S5), BS can be characterized as follows:[1] $utters_X(p) \wedge \neg K_X(p) \wedge \neg K_X(\neg p)$ One of the disadvantages of doing so, however, is that the possession of knowledge becomes basically a binary phenomenon. One either has knowledge about p or one does not. An alternative way to characterize the concept of knowledge is using formal (abstract) argumentation. One of the principles of abstract argumentation is the existence of a graph (Ar, att) where the set of arguments Ar provides the nodes, and the attack-relation att provides the arrows. Given such an *argumentation framework* [3], one can distinguish different ways (like complete, grounded, preferred, stable or semi-stable semantics) of identifying the set(s) of arguments which can collectively be accepted. Moreover, many of these principles (also called *argumentation semantics*) have associated proof procedures in the form of discussion games, in which two players (proponent and opponent) exchange arguments, each of which attacks the previous argument. Thus, whether an argument is justified depends on whether it can be defended in the associated discussion game.

As is explained in [1], argumentation gives rise to a more subtle concept of knowledge. An agent X can be said to have more knowledge w.r.t. a proposition p than an agent Y if it has at its disposal a strict superset of arguments relevant to p. More particularly, we can distinguish two different situations. If X and Y disagree about the status of p, then let them do the formal discussion game against each other. The party that wins the discussion can be said to be more knowledgeable w.r.t. p. If, at the other hand, X and Y agree on the status of p then let them discuss with other agents who disagree with X and Y. If X can maintain its position in a strict superset of situations where Y can maintain its position, then X is said to be more knowledgeable about p than Y.

The thus described notion of knowledge is not too far from everyday practice. Imagine an expert on, say, climate change being interviewed on television. If this "expert" is not able to reply to the interviewers objections against his theory of climate change it would be hard to maintain that he has real knowledge

[1] In contrast, lies could be described as: $utters_X(p) \wedge K_X(\neg p)$

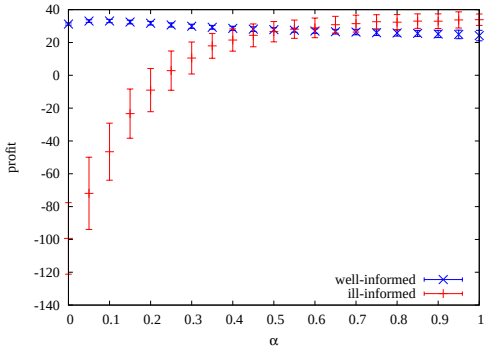

Figure 1: Profit (10% of consult. are ill-informed).

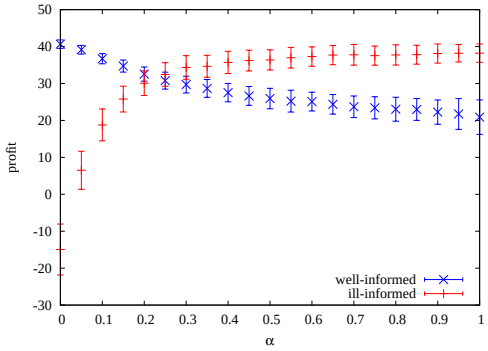

Figure 2: Profit (50% of consult. are ill-informed)

on this topic. Having knowledge implies the ability to defend one's position. Moreover, in the example of climate change, it is problematic to define knowledge simply as "justified true belief", since this assumes access to the objective truth, which in this case will only reveal itself in the medium to long term future. Similar observations can also be made in fields like investment strategies, macro economic planning and development aid. One cannot determine whether someone's position is "true"; one can only determine whether it is well-informed.

In order to better understand why the concept of BS is so common in society, we developed a software simulator (see [6] for a detailed description). We consider a client-consultant scenario. The information that consultants try to advise their clients on is modeled in a simple argumentation structure as follows (arrows represent the attack relation):

$$A_1 \leftarrow A_2 \leftarrow \cdots \leftarrow A_n \ ,$$

At the outset of a simulation run, argument A_1 is known to all consultants. Then, following the order of the indices, a certain number of new arguments becomes available in each time step, and can be acquired by the consultants. Some of the consultants, we call them *well-informed*, buy arguments as soon as these become available, since they want to be always up-to-date. The other consultants, called *ill-informed*, follow the idea of BS and only buy as many new arguments as needed to appear knowledgeable to the clients. Consultants of the first class are more expensive in their consultations (they invest more in their knowledge); but in contrast to the second class of consultants they never encounter clients that are actually better informed than they are (and so their reputation can be expected to be higher). Now, a client can rate a consultant with respect to the consultant's cheapness c and reputation r (we normalize both to $[0, 1]$). In our simulations, we let clients choose the next consultant with a probability proportional to $\alpha \cdot c + (1 - \alpha) \cdot r$, where parameter α allows clients to balance between the importance of a consultant's price and reputation. Finally, to compare the efficiency of the two classes of consultants we compute their *profit*, which depends on price, frequency of consultation and costs of the acquired arguments. Figures 1 and 2 show results of our simulator for specific settings (we show mean and standard deviation for 2^{10} runs). We leave it to the reader to verify that in fact there are settings where ill-informedness actually yields a higher profit than being well-informed – increasingly when the price of consultants takes priority over their reputation (α increases).

References

[1] M.W.A. Caminada. Truth, lies and bullshit; distinguishing classes of dishonesty. In *Proc. of the 1st IJCAI Workshop on Social Simulation (SS@IJCAI)*, 2009.

[2] Nick Davies. *Flat Earth News; an award-winning reporter exposes falsehood, distortion and propaganda in the global media*. Vintage Books, London, 2008.

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

[4] Harry G. Frankfurt. *On Bullshit*. Princeton University Press, 2005.

[5] Harry G. Frankfurt. *On Truth*. Alfred A. Knopf, 2006.

[6] Eugen Staab and M.W.A. Caminada. Assessing the impact of informedness on a consultants profit. Technical report, University of Luxembourg, September 2009. Online: `http://arxiv.org/abs/0909.0901`.

[7] Nassim Nicholas Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.