

Truth, Lies and BS; distinguishing classes of dishonesty

Martin Caminada

University of Luxembourg

In this paper we distinguish three classes of dishonesty: lies, bullshit and deception. We show how these classes can be modelled in classical epistemic logic (S5) as well as in formal argumentation.

A lie, in its most simple form, can be defined as the utterance of a statement which the speaker knows not to be true. That is, an agent X is lying on proposition p iff the following holds: $utters_A(p) \wedge K_A(\neg p)$

Another form of dishonesty is to make statements about things one has 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. 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 [5, 6] and to allow the reader to easily relate the phenomena described in this paper to his every day life experiences. As described in [5], 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 [5] 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, 3].

In its simplest form, BS can be characterized as follows: $utters_A(p) \wedge \neg K_A(p) \wedge \neg K_A(\neg p)$

As with lies, there is also an intensional aspect related to BS. Although one intends the hearer to believe that p , it is often more important that the hearer will believe that A is knowledgeable about p . While a liar has a very distinct purpose of wanting the hearer to believe p (because such a belief would have consequences that would suit the liar’s goal), a bullshitter might be equally well off by telling the hearer that $\neg p$, as long as he appears knowledgeable in doing so.

The third form of dishonesty to be discussed is that of deception. Although deception can be described in a very broad way, for current purposes we are interested in a more focussed concept of deception, as applied in [1]. The basic idea of deception is to provide the hearer with correct information, which the hearer is most likely to use to make an incorrect inference. As an example, suppose one wants to persuade a friend to come over for the weekend. One could try to persuade him by claiming the newspaper predicts good weather this weekend, even though one knows that the local newspaper weather forecast is notoriously unreliable, and that the much more reliable TV-news predicts rain all weekend. In this case, one did not tell anything untrue, or lacking sufficient backing. The newspaper really *does* predict good weather. But by telling this to one’s friend, he will make an inference that one knows to be incorrect, namely that this weekend the weather will probably be good. Thus, deception is a particular form of dishonesty that one can apply even without speaking anything else than the truth.

One of the interesting things about deception is that it depends on nonmonotonic reasoning. Deception basically functions by providing some pieces of information and withholding other pieces of information in order to lead the victim to wrong conclusions. If we would tell that Tweety is a bird, without telling that Tweety is a penguin, the hearer would most probably derive that Tweety can fly, which we know to be

wrong. With classical (monotonic) logic, this would not be possible. Withholding information in a classical formalism will result in inferences that are *missing*, whereas in a nonmonotonic formalism it results in inferences that are *wrong*. With deception, one makes use of the nonmonotonic inference capabilities of the other person in order to implant wrong beliefs, without having to resort to lying ourselves.

In standard epistemic logic (S5), the possession of knowledge is basically a binary phenomenon. One either has knowledge about p or one does not. It is also possible to characterize the concept of knowledge 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* [4], 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 described in [2], argumentation gives rise to a more subtle concept of knowledge. An agent X is said to be more knowledgeable 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. The party that wins the discussion is 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 climate change being interviewed on television. If this “expert” is not able to reply to the interviewer’s objections against his theory of climate change it would be hard to claim he has real knowledge 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.

Using the thus described concept of knowledge, we can re-examine the classes of dishonesty distinguished earlier. A lie can be characterized as making a statement that is not in line with the arguments that one has at one’s disposal (that is, either declaring an argument or associated proposition as justified while it follows out of one’s argumentation framework that it is not, or vice versa). BS can be characterized as making statements about which one has no proper knowledge, that is, based on a very small set of relevant arguments one has at one’s disposal, compared to the set of arguments one could have at one’s disposal. Deception can be characterized as “feeding” the other party with a carefully selected subset of arguments, in order to change the status of the other party’s argument (justified or not justified) to something one does not maintain oneself. The existence of the aforementioned classes of dishonesty has significant implications w.r.t. agent strategies and mechanism design, as is explained in [2].

References

- [1] J.E. Adler. Lying, deceiving, or falsely implicating. *The Journal of Philosophy*, 94(9):435–452, 1997.
- [2] M.W.A. Caminada. Truth, lies and bullshit; distinguishing classes of dishonesty. In *Social Simulation Workshop at the International Joint Conference on Artificial Intelligence (SIJCAI)*, 2009.
- [3] Nick Davies. *Flat Earth News; an award-winning reporter exposes falsehood, distortion and propaganda in the global media*. Vintage Books, London, 2008.
- [4] 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.
- [5] Harry G. Frankfurt. *On Bullshit*. Princeton University Press, 2005.
- [6] Harry G. Frankfurt. *On Truth*. Alfred A. Knopf, 2006.
- [7] Nassim Nicholas Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.