Report on "The Shape of Cooperative Communication"

August 10, 2016

1 Summary

This paper studies communication in cheap talk settings without incentive problems where the receiver observes a noisy version of the sender's message and the value of information may differ across states (e.g., information may be more valuable in good states than bad states). The paper consider settings in which there is little noise, and noise is uniformly distributed. The main result characterizes the message strategy used by the sender to convey his information to minimize error. The authors argue that *Variation in language should be saved for instances in which it is most important to accurately communicate, and we show that this is equivalent to saving variation for instances that arise frequently. If a given set of events occurs often, then using variation to distinguish among those events is valuable.* The authors show that this theory has the potential to explain data patterns in a host of applications, ranging from mutual fund ratings, journal reviewers, financial analyst recommendations.

1.1 Comments

The paper is well written and motivated. It provides some elegant results that can explain puzzling data paterns in communication settings.

1. I particularly liked the setting considered in Proposition 1 and would suggest the authors focus and explore in more depth the setting with both continuous message space, possibly extending it to settings with large noise and non-uniform (possibly unbounded) noise.

2. Assuming exogenous noise (i.e., noise that is independent of the language chosen by the sender) in an economic communication setting is a bit strange in most applications because it amounts to assuming that by changing the language the sender can change the informativeness of the communication, as if when say a manager reports the firm's earnings in euros they become more informative than when he reports them in dollars.

3. I am a bit confused about the following: in a continuous symmetric Gaussian setting the sender would be able to achieve almost perfect communication via a linear strategy –regardless of the amount of noise– by using a large slope. The basic prediction of that example would be that the presence of noise leads the sender to exaggerate the state relative to the case without noise so to make his message more informative. I presume that if the support of the message strategy is restricted to be compact (or the signal is restricted to have finite variance) then one gets the result that noise induces exageration (i.e., slope greater than 1) where it matters, i.e., everywhere but at the tails of the distribution.

4. There are too many models in the paper. There are three models: a) continuous model with noise, b) discrete messages model without noise, and c) discrete actions with noise and uncertainty about payoff. I think the paper would improve if the authors committed to a single model (i.e, the first model) rather than three. To my taste the model with discrete messages is redundant and seems to be there only to justify the motivating examples.

5. I would suggest the authors focus on the standard symmetric setting where $I = 1$, unless they develop applications where $I(\cdot)$ is endogeneized. It's well known that the optimal information system (i.e., reporting strategy) will depend crucially on the loss function one uses, but explaining data patterns based on changes in $I(\cdot)$ is like assuming those patterns. Since this is not even needed for the main results, then why have it.

6. Information theory. The authors reference the economics literature. But communication without incentives is a classic problem in information theory. In fact, it seems to me that a similar problem has been studied by the Information Theory literature under the label of "communication with a noisy channel." This literature focuses on characterizing how much information can be transmitted when the channel is noisy. It would be worth knowing how this paper contributes relative to the information theory literature.

7. What about uniqueness? As is well known, standard cheap talk models without noise do not have clear predictions about the distribution of messages/ratings, since there is a large multiplicity of languages that are consistent with a given equilibrium. The case with noise is different because then the language itself affects the amount of communication; so the only multiplicity should be about the level of the reporting function.