Storytelling and the Science of Information and Decision Systems

Kush R. Varshney
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
krv@mit.edu

Many stories were told and many different issues were raised and discussed at the November 12–14 symposium on paths ahead in the science of information and decision systems. Prof. K. J. Åström and others gave a picture of control theory residing between computing and communication, between mathematics and engineering/biology/ economics/physics. Until the last session, however, no one had tried to define what the science of information and decision systems is. In the last session, Prof. M. I. Jordan succinctly said that what we, researchers in the science of information and decision systems, are all after is *intelligence*. I agree that intelligence is the essence of the science and the destination for the paths ahead. I would further argue that *storytelling* is what we are after.

A recurrent theme in honoring Prof. S. K. Mitter during the symposium was that his thinking occurs in all different planes. Storytelling may seem like a field afar from information and decision systems at first, but it is just a view of the field from a different plane of thought. Storytelling is the intelligence that distinguishes humans from others. The oral storyteller, the balladeer, the singer of tales, is a human information and decision system.

As discussed by A. B. Lord [1], the bards that were a part of the living tradition of oral narrative in traditional societies heavily relied on formulaic language—prefabricated phrases that are retrieved from memory at the time of performance. When reciting epic poetry, the storyteller accessed an appropriate description from his or her mental collection of descriptions to generate the story. There could not be any delay in this process. No two performances were exactly the same; the decision of what to include incorporated feedback from the audience. The objective was not only to convey the basic story, but also to keep the audience enraptured for as long as possible. The best storytellers had learned the largest and best collections of formulaic language, but were also able to deploy that language to keep their audiences completely under their control.

The following question often arose during the symposium, including in the debate on computation and insight by Prof. P. A. Parrilo and Prof. S. P. Boyd. What is a solution? It used to be a formula, like the Wiener filter. Then it became an algorithm, like the Kalman filter. More than one panelist suggested that a solution is really a narrative in whatever form it takes. The oral storytellers drew from their banks of formulas using their own algorithms to produce a narrative that controlled the audience. The story produced on the fly by the storyteller was the solution to a difficult control problem. If we can develop an artificial storytelling machine, we will have developed a very general solution in the science of information and decision systems.

A research area promoted by Prof. Jordan was that of nonparametric Bayesian topic models based on hierarchical Dirichlet and beta processes. These generative models grow in complexity when presented with more and more training data. Structures and representations that are learned involve the sharing and reusing of features and parts from different training samples. Such a trained model is very much akin to a storyteller who has picked up formulaic language from hearing many stories. The key is to learn this model or structured collection of formulaic structures in a way that it can be used for specific tasks such as story generation for control. Although if you have seen or heard enough, even very simple models and retrieval techniques are effective [2].

Another topic discussed during the symposium was of simplifying and condensing what is taught to students of the science of information and decision systems. This is also a storytelling problem: finding a narrative that is interesting, conveys a basic story, and uses formulaic components prudently. The compilation and redaction of oral narratives into written works often involves finding and combining commonalities. Redaction of the theory of information and decision systems ought to be organized around formulaic parts which reoccur in various guises, one example being orthogonal projection and the Pythagorean theorem. The same formula need not be repeated several times, just as the tale of the tortoise and the hare need not be repeated as the tale of the snail and the cockroach, or as the tale of the sloth and the spider monkey.

Oral storytelling is part of the foundation of humanity as much as anything else. Many researchers in our field take inspiration from biological and human systems, including examples presented during the symposium. Traditional storytelling is a type of human intelligence that could and should be an inspiration for the development of solutions in the science of information and decision systems.

References

[1] Albert B. Lord. *The Singer of Tales*. Cambridge, MA: Harvard University Press, 1960.

[2] Antonio Torralba, Rob Fergus, and William T. Freeman. 80 Million Tiny Images: A Large Dataset for Nonparametric Object and Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11):1958–1970, November 2008.