Twitter Thread by Maxim Raginsky

Maxim Raginsky
@mraginsky

In this thread, I will argue that the conventional wisdom that Shannon's information theory is all about syntax and not about semantics stems from superficial reading. On the contrary, even his 1948 BSTJ paper is already concerned with syntax, semantics, *and* pragmatics. 1/14

Of course, it does not help matters that Shannon himself states that "semantic aspects of communication are irrelevant to the engineering problem" of communication on the first page of his paper. But you have to read the entire paper, not just the first page. 2/14

Let's first pin down the notions of syntax, semantics, and pragmatics in the sense of Charles Morris: syntax is about relation of symbols to other symbols, semantics is about their relation to objects, and pragmatics is about their relation to subjects. 3/14

Let's start with pragmatics. The full formulation of Shannon's theory is concerned with the problem of reliable communication of some state of the world over a noisy medium, subject to a fidelity criterion. This is an operationalist, pragmatist concern. 4/14

We may envision the problem where we take an action U contingent on the state of the world X and incur the cost c(X,U). We cannot access the state of the world directly, but must mediate the act of selecting the action through transmission of symbols over a noisy channel. 5/14

It is crucial to note that X and U themselves may be nonsymbolic, analog objects. The simplest set-up in Shannon's theory then involves the following steps: source encoding X -> W, where W is a discrete message to be transmitted over the channel; 6/14

channel encoding W -> S, where S is a string of symbols that will actually be sent over channel; channel decoding S' -> W', where S' is the corrupted string of symbols at the channel output and W' is the decoded version of W'; 7/14

and, finally, the source decoding W' -> U. The purely syntactic part of this arrangement is W -> S -> S' -> W'; indeed, at this point we do not care about the meaning of X or U or about their pragmatic relation, we only want to ensure that W = W' with high probability. 8/14

The channel code W -> S is a one-to-one map that will introduce redundancy into S for the purposes of error correction (parity checks, etc.). The channel map S -> S' is stochastic. The channel decoder S' -> W' attempts to recover the message W despite the presence of errors. 9/14

The problem of designing the channel encoder/decoder is purely computational and symbolic -- this is syntax par excellence (see, e.g., algebraic theory of error correcting codes). This is where we care about entropy not exceeding capacity of the channel. 10/14

However, the source encoder X -> W and source decoder W' -> U are of a different variety. The encoder maps a nonsymbolic state X into a discrete message W, and the decoder maps the decoded discrete message W' into a nonsymbolic action U. 11/14

These two maps implement, respectively, measurement and control. Following Howard Pattee and Robert Rosen, this linkage of nonsymbolic external aspects of the environment to symbols that can be manipulated computationally is precisely semantic. 12/14

Finally, the overall purpose of the entire interconnection is about minimizing expected cost. Since we are explicitly talking about usefulness and purpose, this is manifestly pragmatic. 13/14

It is important to keep in mind that quantities like entropy, mutual information, etc. are only meaningful insofar as they relate to operationalist, pragmatic concerns like expected cost, probability of decoding error, rate and minimum distance of an error-correcting code. 14/14