Klaus Krippendorff 2022.2.18

Dear Loet and others discussing Chapter 4 of Loet's book:

I am sorry, I had planned to be part of the discussion but I messed up the time zones. I had hoped to make five comments.

1. I hate to say, I disagree with the definition of information which you said you took from Shannon's theory of communication. Shannon's definition of information is the difference in the entropy of a receiver before and after receiving connecter to a sender.

   $$T(Sender{:}Receiver) = H(Receiver) - H_{Sender}(Receiver)$$

   Where: $$H_{Sender}(Receiver) = H(Sender, Receiver) - H(Sender)$$

   and $H(Sender,Receiver)$ is the joint entropy of sender and receiver, your H_{12} . The amount of information transmitted is symmetrical, time does not enter the calculus:

   $$T(Sender{:}Receiver) = H(Sender) + H(Receiver) - H(Sender,Receiver)$$

   This is consistent with your formula (4.5).

   Weaver's addition to Shannon's general formulation is a simple mapping from the set of signals to the set of corresponding meanings. This is the most primitive conception of meaning I can think of.

   Bateson's definition: "Information is the difference that makes a difference" talks of significance instead of meaning with is not as specific as a particular meanings but it generalizes meaning as Larry Richards pointed out in his contribution to the discussion. Bateson's 1st "difference" refers to uncertainty reduction: When one ask a genuine question, one envision several possible answers. The information provided by the answer reduces the uncertainty of many answers to one. His 2nd use of the word "difference" appears in the context "making a difference" which means "importance" or "significance". His formulation acknowledges that not all variations in messages make a difference. Only what is relevant to a receiver can be informative.

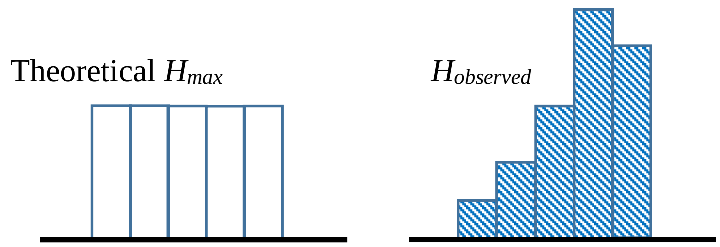
   Admittedly, unless a semantic space is finite and enumerable, Bateson's definition of information is difficult to quantify. However, Bateson's conception is often enough.

2. Your definition (4.1) is mathematically clear. However, to me calling that proportion "redundancy" is not. In ordinary English, redundancy is the inclusion of unnecessary variation like repeating a statement when one was already sufficient. In human communication, redundancy is important because we may not attend to all the nuances of a communication. For another example, we would not be able to correct spelling errors if English wouldn't have redundancy. Bateson's reference to "making a difference" correctly excludes redundant aspects of messages from his concept of information.

   In cryptography, where Shannon developed his conception of information, redundancy is an important aid to decode encoded messages. If every signal of an encrypted message has a unique meaning, a cryptographer would have difficulty decoding it without a known code. If a cryptographer knows that the original of an encoded message was written in English, it is the redundancy of English spelling that facilitates helps figuring out what an encoded message meant.

Your definition (4.1) equates redundancy with what I would call channel capacity. While I think (4.1) is far removed everyday, common dictionary, and Shannon's definitions, you are of course free to define anything you like but it would have helped any reader to get a clearer sense of how your definition relates to your overall mission.

Specifically, (4.3), H_{max} , is a measure of the entropy of equally likely alternatives whereas $H_{observed}$ is the entropy of unequally distributed alternatives. Graphically your redundancy is the relative difference between the following two distributions:



Your Figure 4.5 is pretty clear showing that the number of alternatives currently available, H_{max} , is increasing exponentially due to the digitalization of technology. One question I wanted to raise is why you think that your definition of redundancy was "historically excluded". Do you think that ancient humans had no idea of the alternatives available to them?

I believe human beings face some biological limits on the number of alternative actions they can engage in. However, considering the combinatorial nature of language and vocal variations humans can produce and recognize these are huge numbers, possibly exceeding the alternatives made available by digital technologies.

To me, your measure of redundancy as the difference between the entropies of the two distributions depicted above explain very little: channel capacities for sure, unused variety perhaps, possibilities worthy of exploration, maybe. It would surely help knowing where H_{observed} resides and for whom. The reality of our redundancy measure would be difficult to locate, as H_{max} is an analyst's choice.

3. The idea of decomposing complex systems into related parts is standard systems analytical practice. W. Ross Ashby sought to generalize Shannon's theory into a calculus with the aim of tracing the flow of information quantities within and between the parts of complex systems. In 1962-3, I was one of his students and witnessed even participated in his effort. In 1969, he published the information theoretical equations he had worked out. I continued this work for a while but encountered a problem that Ashby had not solved – nor did you.

   To recognize this problem, keep in mind that Shannon's original theory accounted for linear flows of information from senders to receivers. It turned out that systems consisting of three (or more) parts can exhibit interactions that are irreducible to the sum of information flows between any two parts. Algebraically, this is handled easily by defining interaction quantities that preserved the accounting equations. Ashby in 1962 and you in

2018 were ok with that. I had reasons to lose confidence in these accounting equations a couple of years before I published my solution in a 1980 paper. http://repository.upenn.edu/asc_papers/237.

As you acknowledged and quoted me for observing: The algebraically calculated information quantities for interactions among even numbers of variables are positive, but negative for uneven numbers. As all entropies are defined in terms of logarithms of observable probabilities, negative quantities did not make much sense to me. I found out that they occurred because the logarithms of interaction terms were not probabilities but mere artifacts of the accounting equations whose parts were to sum to the total. Moreover, it turned out that these unreal terms occurred largely when the interactions were circular, not explainable by information flows involving a smaller numbers of variables. I developed an algorithm that rectified these oddities of accounts for interactions among three or more variables, effectively bypassing the artificiality of algebraic accounting equations while preserving the aim set out by Ashby of accounting for all the complexities that multi-valued probabilities could exhibit.

I dared saying that Ashby's and my own accounting equations before my correction to which I now need to add 2018 equations fail to measure anything real when the complexities of data involves three or more variables, parts of systems, or people in organizations. The algorithm I devised to account for the information quantities in complex interactions preserved the idea of decomposing complex systems into information quantities between and within parts avoided the artificial nature of the algebraically obtained quantities – convenient for the analyst but with questionable relationships to data.

4. Your section 4.8 concerns the interaction quantities $T_{123\ldots n}$ a measure that I consider flawed for the above-mentioned reasons. This section introduced directional relationships between variables. As entropies are measures of probability distributions, arrows of time, causality, or dependency can only come from the definition of the variables in the data. As your $T_{123\ldots n}$ does not differentiate definitional dependencies, I have no clue about what justifies the arrows between the three components in your Figure 4.8, where they come from, what they represent, and most important, what could make them change them from clockwise to anti-clockwise.

   In this section, you refer to the principle of triadic closure. If I understand it correctly, this construct of network analysis is defined by three nodes being maximally connected by three links, constituting a unity absent otherwise. Networks of binary relationships between variables are relatively flat description of systems by comparison to the top-down decompositions of complexities that underlie the idea of accounting equations. My question concerns the connections you are making. To be clear, network analysis could not depict a three variable system whose entropy equals that of the tertiary interaction, all binary communication measuring zero (except by lumping the three nodes into one). It would be capable of representing a three variable system whose entropy equals the sum of the three information quantities between any pair of variables, the tertiary interaction measuring zero (except that there would be no recognizable triadic closure) – see the above link for numerical examples. Even if a network analysis would consider quantities associated with each link, it could not cope with any case between these two extremes. So, where do these arrows come from?

I suspect, please show me if I am wrong, section 4.8 is your effort to frame the positive or negative quantities of the algebraically calculated higher-order interaction quantities due to whether the sum of the lower-order interactions over- or under-determines the higher-order quantities in terms of feed forward or backward. Besides, you would need time to talk of feed forward or backward. I am unable to relate the algebraic oddity of positive or negative quantities to what you claim your simulation reveals.

5. While Ashby's and my subsequently corrected accounts of complex probabilistic system into smaller more manageable parts continues to make sense to me, I lost confidence in the ability to describe social systems in terms of probability distributions. My question to you is what your calculus of redundancy contributes to an understanding of social phenomena. Yes, alternatives are important but they are not given. They evolve in human interactions, with the technology we create, and are encouraged by the vocabularies in a population – see Figure 4.5.

   I applaud your opposition to Luhmann's simplistic notion of meaning but wonder how meaning fears in your calculus of redundancies.

   I appreciate your insistence that social systems do not exist as biological organisms do. However, I would say that social systems are not merely social scientific abstractions. They reside in parts in their human constituents' mind and become real when enacted into networks of conversations. To me conversations are evolutionary processes that generate alternatives. From this perspective, social organizations should serve their constituents well-being and amplify the human agency of their members for the benefit of all. Focusing on the button-up generating of alternatives is radically opposed to the customary use of biological metaphors in sociology according to which all parts need to serve the viability of the whole, in effect supporting authoritarian conceptions of society. I think we agree on that.

   Also, to me, social organizations need to be reconstitutable during consensually agreed periods in time, for special occasions, or when members leave; and they need to provide contexts in which possibilities grow that individual members are unable to create on their own. Unfortunately, accounting equations explain what was observed. I think it is more important to keep track of where and under which conditions alternative actions are created or prohibited, and which fears of expected uncertainties are reduced by information made availble.

It indeed was a pleasure to read your chapter. It brought to life so many issues I almost forgot.

Klaus