

Task 1: Claude Shannon and Zipf in Generative Models for Languages

Claude Shannon viewed the process of communication as a mathematical process from one point to another through different components, which are described below:

1. Source: Is the person who produces the message. It could also be a computer sensing data just as Artificial Intelligence does.
2. Transmitter: It is the object that transforms the message into signals, such as a cellphone or a keyboard.
3. Channel: Is the mean through which the information travels. It could be air, wires, fiber optics, etc.
4. Receiver: Is the one that converts the signal into a message, like the speaker of a cellphone or a screen.
5. Destination: Is the person that receives the message.

Claude Shannon not only saw communication as the meaning of the words, but as a whole process that has been being developed through the years with a complexity we can't even imagine and we don't really think of, which involves codification.

He also considered redundancy and randomness. He explains both extremes are completely useless for successful communication since they don't carry enough information in any way. This is because if someone started saying words that make no sense between them, and are just deliberately spoken, we would just have units of information, but we wouldn't have any message. On the other hand, if someone repeated the same word they would probably not convey any message since it is not enough information to understand what they are saying. If we just say the same thing over and over again while trying to express something, we will not get to any point.

For that reason, Claude Shannon also implies that messages that contain the maximum amount of information would be the ones that are found in between these two points (randomness and redundancy). It would be a balance point in which ideas could be transmitted to other people since the units of information would be coherently structured, and there wouldn't be any lack of information.

Claude Shannon's point of view could also be connected to Zipf's law, who states there is probability and language since in any large collection of words, there will be common words that repeat throughout the text.

This may be related because, the less predictable a message is, the more information it carries, because that would be new information. But if we just repeat the same ideas, our statements become predictable. And that should be the reason language keeps evolving through the days and years. If we want to develop as a society, and make a change in our world, in who we are, and where we are heading as humanity, already existant words may be not enough. And I think that's what makes language and the ability we have to communicate extremely beautiful, because we are always finding new ways to understand world and understand each other through words.

An example where we can see all of this applied is in famous Chat GPT, which is a generative model for languages based on probability and prediction of words. When you speak to Chat GPT, it analyzes the words you have used and based on that it creates a sentence using the most likely word to be used next. By analyzing how we use language, it will say, for example, I study Chinese at college or I study Chinese at the library rather than saying I study Chinese at the se language and which is the following most predictable word.