Identity of Words in Natural Languages

by Sven Nilsen, 2019

Inspired by the philosophy of Ludwig Wittgenstein, in this paper I suggest a version of the path semantical axiom, adapted to identity of words in natural language.

The path semantical axiom states that, given two symbols `$F_0$` and `$F_1$` and two associated collections of symbols `$X_0$` and `$X_1$` respectively, if the two symbols `$F_0$` and `$F_1$` are identical, then the two associated collections of symbols `$X_0$` and `$X_1$` are identical. From this simple axiom, one can construct atomic functions, which can be bootstrapped into Type Theory and normal functions, which can be bootstrapped into Boolean Algebra, Probability Theory and so on.

Ludwig Wittgenstein advocated the following philosophical perspective of natural language:

> To understand language, one must observe how it is used.

Therefore, to interpret path semantics in the context of natural languages, one can define:

> If two words `$W_0$` and `$W_1$` are identical, then the way they are used are identical.

This immediately leads to a theory of agents that interpret, process and executes. One can image a word as a proposition about the existence of a symbol that is being interpreted by an agent. If two words are identical, then one can expect identical behavior.

However, behavior can be probabilistic. To be more precise, a collective of agents display roughly the same probability distribution in behavior when a symbol is being interpreted.

For example, if I use the word "left" to direct the attention of an English person, then I might use the Norwegian word "venstre" to do the same thing to a Norwegian person. The words "left" and "venstre" are identical except for the choice of language, because they are used the same way.

As a thought experiment, one can create a program that outputs "left" or "right" randomly and a human is instructed to click two buttons placed on the same horizontal line. The human clicks the left button when the word "left" is shown on the screen. The same thing is done for "right". By measuring which buttons that are clicked, one can learn to understand the difference between these two words. Afterwards, the human can type in "left" or "right" and the computer can click the buttons.

If the human in the thought experiment swapped "left" or "right" consistently, or chose a random button, the computer would learn to behave differently. The program could also output "choose your main hand" and different people would react differently, but consistently. It is also undefined how the computer should react if a human told it "choose your main hand". However, if the computer is just mimicking the behavior of humans, then it is up to humans to observe what the computer does, in order to understand what the computer means. After all, when the identity of words is determined by how they are used, it is also possible for computers to influence what they mean.