Frames, conditional probabilities & diagnostic information

ALEX TILLAS
(Heinrich Heine Universität Düsseldorf)

According to Frame Theory (FT), all knowledge is structured in frames – recursive attribute-value structures. Despite the various advantages of FT though, frames seem to capture mere correlations between category traits without distinguishing between different kinds of statistical/probabilistic information contained in a given concept. For instance, frames do not distinguish between conditional probability information (e.g. x will have f (e.g. a heart) given that x falls under c (e.g. ANIMAL)) and diagnostic information about the conditional probability that x will fall under c given that x has f. In turn, it is not clear how FT could account for necessary and contingent characteristics/attributes falling under a given category-concept and the respective frame. Given that diagnostic information plays a crucial role in category detection, accounting for such information would also grant added cognitive adequacy to FT.

In this paper, I suggest a way to account for how conditional probability information and diagnostic information could be represented in frames by appealing to the principles of Hebbian learning and informational semantics. Even though the suggestions made here are of a theoretical nature, they do enjoy significant empirical support from independent evidence, e.g. Associative Long-Term-Potentiation. More specifically, the underlying idea is that concepts are built in virtue of an abstraction process selecting similarities across representations formed during encounters with instances of a given category. Similarities are understood here in terms of frequencies of occurrences to the extent that the more frequently occurring features across members of a given set would naturally capture the stronger similarities amongst them. Adopting a light-hearted reductionist view and slightly paraphrasing Hebb, the more frequently a certain feature occurs across instances of a given kind, the stronger the connections between neurons that ground perception of that feature will grow. What is key is that concepts are built out of representations carried by stronger connections, i.e. representations of features/properties that most, if not all, members of a given category bear. It is for this reason that concepts exhibit nomological covariance properties with members of that category (counterfactually supporting covariance). Given that the target of FT is to represent type-frames, FTs should aim at modeling abstracted representations, or the output of the abstraction process, (rather than representations of particulars). In turn, abstraction could be seen as a frames-formation process.

Stronger connections carry statistical information about feature occurrences across instances of a given kind. Information of features carried by stronger connections could be used to form a 'feature-hierarchy' for the appropriate frame in virtue of a process that sits on top of the aforementioned abstraction process. Essentially the latter process ascribes coefficients to representations of features. Simply put, the stronger the connection carrying information about a certain feature in a given set, the greater the coefficient ascribed to this information/representation will be; in turn the higher up in the feature-hierarchy this feature will appear. It should be clear that the present suggestions do not describe feature-hierarchies in FT. Rather, they aim at shedding light onto how feature-hierarchies are formed and how frames are built.

Ascribing a coefficient to representations allows for differentiating between the 'significance' of different features in a given category/concept and in turn for specifying their position in the appropriate feature-hierarchy. For example, 'having a heart' will appear more prominently in the ANIMAL feature-hierarchy, in comparison to 'having legs'. For 'having a heart' correlates more strongly with instances of animals than 'having legs' (to the extent that fish and certain reptiles have hearts but no legs). In this simplified scenario 'having legs' will yield smaller conditional probability, and in turn diagnostic information about it, in comparison to 'having a heart' in the animal concept and the respective frame.

As an offshoot of this paper, it is argued that the set of representations with the greatest coefficients ascribed to them constitutes category-features that are crucial for category detection. In turn representations with greatest-ascribed-coefficients approximate the 'necessary' attributes of the respective frame.

Finally, by ascribing coefficients to representations and forming feature-hierarchies would allow frames to capture differences about a given property featuring in different categories and in turn frames. In this way, one could account for 'White' having a greater probability to appear in SNOW than in FLOWER, and that 'White' has in turn a bigger diagnostic value for SNOW than for FLOWER. The fewer times a feature is given the greater possible coefficient, call it coefficient z, the greater its diagnosticity for the category, for which it bears coefficient z, will be. For instance, if only snow is white, and x is white, then x is (or in any case has the greatest probability of being) snow (ceteris paribus). In more complex cases, it seems intuitive that maximum diagnosticity is ultimately captured in terms of combining all representations to which greater-than-a-certain-level coefficients have been ascribed.