Liron Karpati
PHIL318G

Classical != L.O.T.

Fodor and Pylyshyn successfully argue that the mind is structured according to a classical architecture and not a connectionist architecture. They fail, however, to show that the mind is organized as a language of thought. I will demonstrate these assertions as follows. First, I will define crucial concepts. Then I will consider Fodor and Pylyshyn's systematicity argument, which concludes that the mind is not connectionist. Following the argument from systematicity, I will present how a connectionist might object and then, in turn, present a plausible classical counter-objection. After articulating why it is hard to refute that the mind is classical, I will highlight that Fodor and Pylyshyn mistakenly derive the existence of a language of thought from the conclusion that the mind is classical. To conclude, I will exit the dialectic between the classical and connectionist approaches and argue that deciding whether there exists a language of thought is crucial to future AI research.

The definitions Fodor and Pylyshyn give for classical and connectionist architectures are essential to their argument. By *architecture* Fodor and Pylyshyn mean a way to organize mental representations. Mental representations are best understood as the states in your brain that encode your understanding. For example, when you see a chair, something in your brain encodes the image of that chair, which is what you see. Whatever is the brain state that encodes the chair, is the mental representation. Fodor and Pylyshyn define a *connectionist* architecture as one in which mental representations are related *only* by association. A *classical* architecture is one in which mental representations are related by association, but also structural relation. The distinction is that in connectionist architectures there are no structural relations between mental representations. In classical architectures, if there are relations between mental representations, then there are structural relations between them.

Fodor and Pylyshyn argue that a classical architecture better accounts for the relationships between thoughts. This is the argument from systematicity. Their argument proceeds as follows. Observe that understanding 'Phil loves Mary' is clearly related to your understanding of 'Mary loves Phil,' or for that matter 'X loves Y' for arbitrary people X and Y. In a connectionist account of the mind, the only reason your understanding of 'Phil loves Mary' relates to your understanding of 'Mary loves Phil' is because the thought entokened by 'Phil loves Mary' is associated with 'Mary loves Phil.' Perhaps this is because every time you heard Phil tell Mary he loves her, she said it back. But now suppose every time Phil tells Mary he loves her, their dog barks. If the connectionist account is to hold, then your understanding of 'Phil loves Mary' is not any more connected to your understanding of 'Mary loves Phil' than it is to your understanding of 'dogs bark'. But surely this is wrong! Of course we would want to say that the thought that Mary loves Phil is more related to the thought that Phil loves Mary. Phil, Mary, and the predicate loves are all composite elements in both thoughts, which is why they are similar. As Fodor and

Pylyshyn go on to say, “if this [latter] explanation [for similarity] is right, then mental representations have an internal structure and there is a language of thought. So the architecture of the mind is not a connectionist network” (Fodor & Pylyshyn, 1988, p. 26-27). Implicit in this statement is that “there is a language of thought” is a stand in for saying that the mind is classical. As will be shown later, this is mistaken. For now, however, we let it stand and consider a connectionist’s response.

A connectionist might retort this conclusion with their own example. Consider that I went for a run and, as I ran, I listened to a podcast. In a few days, I went for another run, but with no podcast. Yet, on this second run, I was having thoughts (mental representations) about the podcast because I associated running with the podcast. Clearly my representation of me running and my representation of the podcast have little to no structural relation, yet there is an association. Thus, we have a relationship between mental representations that are not structurally related, so the mind is not classical.

While this example seems to highlight the associationist nature of mental representations, the classical approach can still account for it. Perhaps, on the initial run, you did not form separate representations of running and podcast. Instead, you formed a composite representation ‘running with a podcast’, built out of the sub-representations ‘running’ and ‘podcast’. It is not far-fetched that, on the second run, the representation for ‘running’, because it is structurally embedded in the representation of ‘running with a podcast’, signified the representation of ‘podcast.’ In this account, the representation ‘running with a podcast’ mediated the associative link. Thus structural relation accounts for the relationship between the representations, so the mind is still classical.

In general, whether or not one finds the classical reply to the running example plausible, the original example of ‘Phil loves Mary’ does show that structural relation plays an important part in explaining the relationships between mental representations. Appealing to the definition of classical architecture Fodor and Pylyshyn present, accepting that structure plays an important role in explaining relationships between mental representations entails accepting that the mind is classical. Although this conclusion closes the dialectic between connectionist and classical architectures, the story is not over.

Fodor and Pylyshyn did not intend to merely show that mental representations have internal structure, this would be a weak conclusion. They intended to conclude the stronger claim that there exists a language of thought. This is evidenced in their quote “mental representations have an internal structure and [so] there is a language of thought”[1] (Fodor & Pylyshyn, 1988, p. 26-27). It was a sleight-of-hand to throw in the “and [so] there is a language of thought,” because having internal structure is not sufficient for something to constitute a language. Surely no one

[1] Contextually, it can be inferred that Fodor and Pylyshyn meant “and” as “and so.”

would claim that the set of castles in Europe is a language, even though castles have rich internal structures and even mini castles (towers and keeps) inside of them.

Why should one care whether the organization of mental representations takes the form of a language? Take the case of natural language. We call natural language a language (in the technical sense) because there is sufficient recursive structure for a large part of it to be formalized. To declare that the set of mental representations constitutes a language (to posit a language of thought) is to say that the mind has sufficient structure to largely be formalized. This implicitly validates the formal language approach to AI. If, on the other hand, we decide that the set of mental representations are not sufficiently structured to constitute a language, then AI researchers would be motivated to use more 'black box' methods like neural networks. Thus, deciding on the existence of a language of thought is instrumental in guiding future research on AI. On this question, modern connectionists and classical theorists can continue their dialectic.

To summarize, it has been shown that Fodor and Pylyshyn successfully demonstrate that the mind obeys a classical architecture. However, their definition of classical architecture is too weak to infer the stronger claim that there exists a language of thought. Finally, it is suggested that deciding on the existence of a language of thought is important to guide future AI research.

References

Fodor, J. A., Pylyshyn Z. W. (1988). Connectionism and Cognitive Architecture: A Critical Analysis. *Cognition*, 28, 3-71.