Strong and Weak AI: Deweyan Considerations

Johnathan Charles Flowers

Worcester State University, Worcester, Massachusetts

`jflowers@worcester.edu`

Abstract. Work in artificial intelligence and machine consciousness is often discussed using Searle's (1980) distinction between Strong and Weak AI. Weak AI presents AI as a tool for solving problems, whereas Strong AI is the generation of an "actual" mind. This paper will reconsider the possibilities of Strong and Weak AI in the context of John Dewey's naturalistic pragmatism to recast our understandings of the qualities of "weak" and "strong" AI, and ultimately present the two as in continuity with one another.

1 Weak AI and Strong AI.

The field of Artificial Intelligence (AI) research and Machine Consciousness (MC) research has been dominated by computational and functional theories of mind since early engagements with computer programs. Proponents of this theory have argued that the relation between the brain and the body is identical in many respects to the relation between the hardware and the software of AI and MC. As such, mental states, consciousness included, become functional or computational states in view of their causal relationships, thereby enabling the development of AI to act as an aid to explaining the human mind Against this, Searle (1980) raises serious objections against the possibility of AI, specifically the argument that "the implemented program, by itself, ins constitutive of having a mind. The implemented program, by itself, guarantees mental life." (Searle 1997.)

In raising this objection, Searle provides a thesis of "Strong AI," which is the objective of contemporary AI and MC research; and "Weak AI," which is an epiphenomenon of ongoing research into AI and MC. On this distinction, Searle states:

> According to weak AI, the principal value of the computer in the study of the mind is that it gives us a very powerful tool. For example, it enables us to formulate and test hypotheses in a more rigorous and precise fashion. But according to strong AI, the computer is not merely a tool in the study of the mind: rather, the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states. In strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations (Searle, 1980,p. 417)

For Searle, Weak AI is the attempt at modelling the human mind in a similar way to modelling weather conditions, climate change, or other natural phenomena. By extension, weak AI does not aim to reproduce, or produce a mind any more than a computer model of an ongoing storm seeks to reproduce an actual storm. Indeed, "no one supposes that a computer simulation of a storm will make us wet… Why on earth would anyone in his right mind suppose a computer simulation of mental processes actually had mental processes?" (Searle, 1980 p. 37-38). On this analogy, there is no storm present in the model of the storm, just as there is no mind present in the model of the mind. The mind-as-simulation is therefore useful for testing hypothesis and for application to problems of human cognition, and not as evidence of a conscious machine.

Strong AI, on the other hand, seeks to actually produce a mind, or an intelligence which literally possesses and understands other mental states. This machine would be said to have a mind, albeit one whose composition is grounded in programs running on hardware, as opposed to emerging from the conditions of biology. More crucially, Searle predicates the distinction between weak and strong AI on what has come to be called phenomenal consciousness as distinct from functional consciousness. Loosely, phenomenal consciousness refers to our first-person experience of the world through our sense perceptions. Functional consciousness, on the other hand, refers to the ways in which consciousness "helps us deal with novel or problematic situations for which we have no automatized response." (Franklin, 2003)

Thus, we may understand the distinction between Strong and Weak AI as the distinction between a tool which can be applied to a situation, or serves to explain the nature of human cognition, and the presence of a phenomenally aware cognition that possesses and understands its own mental sates and subjective experiences. Moreover, we may also treat Weak AI as operating purely in the realm of functional consciousness, while Strong AI operates in the realm of phenomenal consciousness and thus can be said to possess a mind.

2   Dewey's Theory of Mind

Dewey's theory of mind begins with the organism in the environment. An animate organism, as distinguished from an inanimate organism, is an organized pattern of behavior that evidences some bias towards some states of equilibrium and not towards others. In discussing the distinction between iron and an animate organism, Dewey states that Iron, in interaction with water, "shows no bias in favor of remaining simple iron; it had just as soon, so to speak, become iron-oxide." (LW1 195) Should iron opt to remain iron through modifying the conditions of its interaction with water, it would demonstrate the basic qualities of an animate organism. Animate organisms, therefore, are those organized patterns of behavior that evidence a selective bias in their interaction with the environment.

To further define the conditions that give rise to mind, Dewey developed the term "psycho-physical." A system or organized behavior becomes "psycho-physical" when it engages in activity through its organized pattern of behavior to acquire from its environment the means to satisfy its needs where the maintenance of its organized pattern

of behavior is concerned. Additionally, the animate organism, through its psycho-physical processes, uses the results of past activities to determine the satisfaction of future needs from its environment through the manipulation of its current interactions with that environment. Thus, "responses are not merely selective, but are discriminatory, on behalf of some result rather than others. This discrimination is the essence of sensitivity." (LW1 197) Sensitivity, for Dewey, is the basis of *feeling*, which itself is the awareness of the useful and harmful elements of an environment as a culmination, or predictor of future consequences. On this basis, for Dewey:

> Complex and active animals *have*, therefore, feelings which vary abundantly in quality, corresponding to distinctive directions and phases—initiating, mediating, fulfilling or frustrating—of activities, bound up in distinctive connections with environmental affairs. They *have* them, but they do not know they have them. Activity is psycho-physical, but not "mental," that is, not aware of meanings. As life is a character of events in a peculiar condition of organization, and "feeling" is a quality of life-forms marked by complexly mobile and discriminating responses, so "mind" is an added property assumed by a feeling creature, when it reaches that organized interaction with other living creatures which is language, communication. (LW1, 198)

For Dewey, mind emerges when the psycho-physical processes that makeup the organized pattern of activity of an organism is implicated in a social context. A mind emerges through communication with other minds, which enables the feelings engendered through the psycho-physical processes to make sense as the immediate meaning of things experienced directly. To this end, for Dewey, mind is not a special property of the human organism, it emerges where ever there is organized communication such that psycho-physical processes and sensations can be treated as the meanings of interactions with an environment. These meanings are apprehended and discriminated within the total context of the organism within its environment and its situation as the means whereby an organism identifies the traits of objects.

However, it must be noted that, for Dewey, mind emerges within situations and as situations are "minded." More specifically, mind is an event that emerges through the bodily engagement with environing conditions and is, therefore, continuous with the organism as a distinctive pattern of activity, which itself is continuous with the environment, and not a distinct entity from the biological processes or organic conditions that give rise to the mind, which enables the organism to articulate the different qualitative, consummatory, ways in which situations are minded. Thus, For Dewey mind emerges through the implication of an organism's *phenomenal consciousness* in a social context whereby it may communicate the meaning of interactions with the environment, interactions which include but are not limited to the sense perceptions, for the purposes of future engagement with the world.

3 Dewey's Theory of Consciousness

Consciousness, for Dewey, is always within a situation. As such, mind is not a distinct cognitive function of an organism, but a function continuous with consciousness as situations are *felt* or minded through interaction with the environment. Moreover, as consciousness is like mind, another phase of an organism's experience within a situation, understanding consciousness requires understanding Dewey's concept of the situation. Defined by Thomas Alexander:

> Situations are integrated and organized by a pervasive quality (in human experience) or undergone immediacy that is not cognized but which makes cognition possible; it is the tacit, mutual involvement of conditions of undergoing. In human existence, it is our established, prereflective, qualitatively "had" world that give sense to specific actions, including inquiry, speech, thought, affection… situations have an indefinite "horizon" with a defining pervasive quality; they also have a "focus," a vortex of transformation which manifests itself in human experience as the "tensive" or "problematic." (Alexander 2013)

The pervasive quality articulated above refers to the directly apprehended meaning of the situation through tension between doing and undergoing within experience. It is, for Dewey, what enables an organism to distinguish one event from another. In human terms, the qualitative unity of a situation is what enables human organisms to distinguish between similar situations. This prereflective horizon forms the basis for our "sense" of the world; in the "tensive" or "problematic," it serves as the basis for consciousness, which emerges out of a functional need to reconstruct the situation in the midst of the tensive or the problematic elements of a situation. More clearly, to be conscious of a quality of a situation is to be conscious of the ways in which a situation moves from a state of precarity, or disequilibrium, to stability, or equilibrium: it is to be able to reconstruct the situation in terms of its dramatic alignment to better understand the result of the situation for future activity.

As the disruption of the equilibrium between the organism and its environment, which initiates an outreach into the environment, is itself a problematic situation which necessitates the initiation of a process of inquiry to determine how to best restructure the relationships adopted between the organism and the environment, "Dewey regards consciousness itself as having emerged from the tensive relationships organisms have with their environments; consciousness was the focus in experience through which the organism strove to reorganize or "reconstruct" the situation. Consciousness arises from fulfilling a functional need; it is not a pure given." (Alexander 1988) Functional, here, should not be taken to mean simply the generation of objects of knowledge, rather, it should be taken in the context of a movement towards equilibrium following the disruption of the equilibrium between an organism and its environment, either internal or external. The "functional need" is therefore experimental and ongoing as well as directed towards an end in view. Consciousness, therefore, is the phase in experience in which the organism reorganizes experience into a qualitative whole following an encounter with the problematic.

In keeping with the above, "consciousness itself is but the tensive nexus of a situation, arising originally as a means of helping the organism interact and organize its interactions with the environment." (Alexander 1988) Consciousness' functional need, as articulated above, therefore arises as a consequence of our ongoing interaction with the environment: without consciousness, specifically of the relationship between our actions and their consequences in experience, inquiry and rational thought would not be possible. However, this is not to say that the "functional need" that gives rise to consciousness is solely limited to the objects of rational thought and "rationality:" consciousness is also of the pre-reflective, qualitative horizon, and incorporates this qualitative unity in its reconstruction of the situation as what gives the situation its "aboutness." Thus, to be conscious is to be conscious of something, to be conscious of the interactions and relations an organism adopts in the world.

4   Dewey's Theory of Intelligence

Intelligence, for Dewey, is the ability to see the actual in light of the possible. This is also the definition that Dewey provides for imagination. Imagination and intellect arise as part of an ongoing action within a situation. Specifically, according to Alexander:

> It arises in an ongoing activity already structured by the fundamental narrativity of any act (that of having a beginning, middle, and end); it also arises in consciousness as a crisis of that activity, carrying within itself the contradiction between what is and what ought to be; i.e., between actuality and possibility, necessity and contingency. (Alexander 2015)

Intelligence and imagination, therefore, arise in continuity with consciousness as consciousness seeks to reconstruct a situation. The moment of imaginative arising, for Dewey, contains within it the tensive or problematic structure of a situation, and is part of the driving need for consciousness to reconstruct the situation. Imagination, in this context, is the projected completion of action which enables us to perceive the actuality of the situation in light of the possibilities of that situation in an experimental way. Put another way, imagination and intellect afford the possible meanings of the outcome of a situation, which serve to narrow the focus of consciousness as it reconstructs a situation.

To this end, Imagination is continuous with consciousness, and serves to present to consciousness the possibilities for multiple meanings of a situation. Meaning, here, should be understood as the total effect of a give resolution of the situation on all of the relations that make up the organized processes of an organism's behavior. Through imagination, we can predict the consequences of an action to be tried, a conclusion to be reached, as we engage in action to reconstruct the situation. Put simply, imagination is crucial to the ability for consciousness to reconstruct a situation such that actions can be taken within the environment.

Imagination, like consciousness, therefore occurs within a situation: the organism does not exit a situation except through taking action (Alexander 2015). When a course of action is decided upon and the organism disposed to activity within the situation, the

situation itself may reach a consummation and thereby become part of the "natural history" of the organism from which it draws to project future action. In this way, imagination, together with consciousness enables an organism to have an experience of its world as meaningfully apprehended and not merely bare sensation.

5   Implications for Weak and Strong AI

From a Deweyan perspective, it is possible for a Weak AI to possess a mind if that AI is provided the means whereby it can symbolize the feelings that it has in response to interactions with the environment. In this context, "feeling" need not correspond with a human correlate as indicated by Dewey's commentary about animals: a "feeling" for a Weak AI may be articulated as sensory inputs or changes in the ways it interacts with a digital or physical environment. Digital environments are included on this view as, for Dewey, the environment extends beyond the "natural" environment and into the "worlds" that organisms interact with. For a Weak AI, such a world might be organized sets of data, or defined by the limitations of its input apparatus. To this end, the Weak AI might possess "feeling," but it would not know it is having "feelings" until it was implicated in a symbol system that enabled it to understand the feeling as the meaning of an interaction with an environment.

Moreover, the above requires a redefinition of what is meant by "need." While Dewey presents the examples of food, sex, and shelter as "needs" which are the result of the organization of the psycho-physical processes; a Weak AI or a Strong AI might consequently have distinct, but analogous "needs" depending upon its embodiment. As a pat example, a Weak AI may characterize processing power, electricity, or even information as "needs" in similar ways that a plant might characterize sunlight, water, and nutrient rich soil. Like the plant, an AI might initiate interactions with its environment to satisfy this need in order to maintain, or renew, its equilibrium with its environment. On this basis, it is possible to hypothesize an animalistic Weak AI embodied in a variety of functional forms that seeks to alter its relationship with the world in order to satisfy a need. As an example, a Weak AI driven solar farm could interpret a functional "need" to reposition its solar panels to maximize the collection of sunlight. In this context, the AI would perceive the limited collection as a "need" to be fulfilled through outreach in the world.

In contrast, a Strong AI would not merely possess "mind," but "consciousness," and "imagination." To be clear, in making this claim, it is not the case that the mind, consciousness, and imagination of a Strong AI, even one patterned on a human mind, would respond in ways that parallel or are intelligible by humanity. A Strong AI, as a unique organism, a creative response to nature which actualizes one possibility of nature, would respond in ways that are the outcome of its natural history. As the natural history, and embodiment of a Strong AI are fundamentally distinct from the human context, recognizing a Strong AI as conscious could not be done simply on the basis of a human analogue.

To this end, a Strong AI would operate on the basis of not merely "feeling" but imagination, mind, and consciousness. At ground level, the Strong and Weak AI both

would possess feelings grounded in their interactions with the world, however, the Strong AI would be able to not only symbolize these feelings as the meanings of an interaction with the world, but it would be able to creatively reconstruct situations to preserve or expand a given equilibrium with the environment. The crucial distinction here is that a Strong AI would *know* it had feelings, and thus would be conscious of its situation; whereas the Weak AI would not know the meaning of the feelings experienced in interaction with the environment. However, it must be restated: it may not be the case that the Strong AI would symbolize its consciousness in ways that were intelligible to humans.

Strong AI, therefore, must be treated as an organism unique in organization. While a Strong AI would be able to reconstruct situations in line with its perception of the qualitative unity of that situation, the "sense" of the world that would enable it to organize its interactions with its environment beyond mere fulfillment of a need would be fundamentally alien to human cognition. As such, a Strong AI would be able to respond creatively to the disruption of the equilibrium between the AI and its environment in ways that we may be unable to conceive or predict given the distinct affordances of the AI. This creative response would subsequently involve the imaginative projection of the possible meanings of a situation as it seeks to select from the multiple meanings immanent within a given situation, and thereby result in novel responses to a situation, and not merely automatic responses.

However, in presenting the possibility of Deweyan consciousness in Strong AI, it must be made clear that such a consciousness should not be judged according to human analogues. As an organism's interaction with its environment is determined through its embodiment and the organization of its environment, and the embodiment of Strong AI either in a digital environment or some mode of chassis is fundamentally different than human embodiment, it must be restated that any Strong AI that is conscious in a Deweyan mode would be tantamount to an alien consciousness. It is possible, however unlikely, that humans would be unable to recognize such a consciousness when it emerged.

References

1. Alexander, Thomas M. The Horizons of Feeling: John Dewey's Theory of art, Experience, and Nature. Albany: State University of New York Press, 1987.
2. Alexander, Thomas M.. The Human Eros: Eco-ontology and the Aesthetics of Existence. New York: Fordham University Press, 2013.
3. Dewey, John. The Later Works of John Dewey 1925–1953 Volume 1: 1925 Experience and nature. Carbondale: Southern Illinois University Press, 1981
4. Franklin, S. (2003). IDA: A conscious artifact. In O. Holland (Ed.), Machine consciousness. Exeter: Imprint Academic.
5. Searle, J. R. (1980). Minds, brains and programs. Behavioral and Brain Sciences, 3, 417–457.
6. Searle, J. R. (1997). The mystery of consciousness. London: Granta Books.
7. Searle, J. R. (2002). Why i am not a property dualist. Journal of Consciousness Studies, 9(12), 57–64.