Charlette Hwang

10/19/2021

Part 1: Context

In this class, we looked at the difference between causal and teleological explanations of actions. Causal explanations interpret the action as being caused by something. For example, my act of drinking water may be caused by my thirst. Teleological explanations, on the other hand, interpret the action based on the goal the agent wishes to achieve through the action. For example, I drink water in order to be hydrated. However, it seems that we can restate this teleological explanation in the form of a causal explanation. (i.e., my drinking water was caused by my desire to be hydrated) In fact, we could simplify any teleological explanation into the following form: "Agent A does action B in order to achieve world state C." If we accept this common form for teleological explanations, we can restructure this form into a causal explanation that goes as follows: "Agent A's desire to achieve world state C caused A to do action B." Based on this simplified form, it seems that all teleological explanations can be restructured or reduced to a causal explanation.

This relationship between teleological explanations and causal explanations is the basis of Davidson's argument. He defines an action in his own way to capture this relationship. His definition goes as follows:

1. An action is an event that is caused by a primary reason under a description which makes the primary reason relevant.

2. A primary reason is the combination of "a pro attitude towards actions with a certain property, and a belief of the agent that [the action], under the [description], has that property."[1]

While Davidson elaborates on the nuances of his definition for a primary reason, for the following paper, I will simplify the definition as such:

2. A primary reason is the combination of a desire and a belief that the action, under the description, would satisfy that desire.[2]

Under this definition, it seems that all actions can be explained causally. As such, this would mean that any teleological explanation would simply be a paraphrased version of a causal explanation.

Part 2: Sehon's Deviant Causal Chain Argument (DCCA)

In order for someone to argue for the possibility of a teleological explanation that cannot be reduced to a causal explanation, they must argue against Davidson's definition of an action. If the definition is faulty, one can argue that it fails to encapsulate every single action; as such, there may be an action that can only be explained teleologically. There are two main ways to argue against Davidson's definition of an action: argue that his definition is too inclusive (i.e., the definition classifies non-actions as actions) or that his definition is too restrictive (i.e., the definition classifies actions as non-actions). Sehon chooses to take the former method in the DCCA.

---

[1] The Journal of Philosophy , Nov. 7, 1963, Vol. 60, No. 23, American Philosophical Association, Eastern Division, Sixtieth Annual Meeting (Nov. 7, 1963), pp. 687, Davidson
[2] I will refer to this combination as a belief/desire pair throughout the paper.

Sehon gives the following example:

There is a philosopher at a conference who desires to disrupt the current speaker. They believe that they can do so if they knock over their glass of water. The recognition of this belief/desire pair agitates the philosopher and makes them shake violently, knocking over the glass of water.

Let us look at this under the Davidsonian definition. The philosopher has a belief/desire pair: the belief that knocking over their glass of water would disrupt the speaker and the desire to disrupt the speaker. This belief/desire pair causes the event of the philosopher knocking over the glass of water. As such, it seems that under the Davidsonian definition, this event should be classified as an action. Unfortunately, many would intuitively argue that this event is not an action; this event lacks a certain intentionality that we commonly associate with an action. The philosopher intended to knock the glass over, but they did not intend to knock it over by being agitated and shaking violently. As such, Sehon argues that because the Davidsonian definition classifies this "deviant causal chain" as an action, the definition itself is faulty and must be reexamined.

In fact, Sehon argues that there are an infinite number of deviant causal chains that achieve this purpose. He proposes the following formula:

1. Agent A has desire D and the belief B that action C would fulfill desire D.
2. Having belief/desire pair BD agitates agent A which causes an involuntary motion E.
3. Involuntary motion E happens to be action C.

In fact, I believe he can generalize this formula even further by taking out agitation:

1. Agent A has desire D and the belief B that action C would fulfill desire D.

2. Having belief/desire pair BD somehow causes an unintended event E.

3. Unintended event E happens to be action C.

Part 3: My best objection to Sehon

      Proponents of the Davidsonian definition have two main methods for arguing against the DCCA: 1) the Davidsonian definition in its current state or with minor modification accurately classifies the event as not an action, or 2) our intuitive judgement is wrong and the event is in fact an action. The second method unfortunately has little merit. Our goal is to evaluate whether the Davidsonian definition correctly classifies an event. To do so, we must have a preconceived notion as to whether or not the event is an action. If we take out intuition completely, there must be another classification method that classifies the event; this classification will show whether our original definition is accurate. However, to test the accuracy of this other classification method, there must also be another preconceived notion as to whether or not the event is an action. As you can see, this leads to an infinite regression. To avoid such a regression, our only hope is to rely on intuition to provide the final say in whether an event is an action and use this judgement as the basis for evaluating the accuracy of the definition. We can of course argue with each other on whether whose intuition is correct; however, for the most charitable approach to the DCCA, I'll assume that the intuition itself is accurate. With only the first option left, let us consider how we can modify Davidson's definition.

      The main issue with the original Davidsonian definition seems to be that it does not require any form of intentionality, a common attribute of what people consider an action. As such, an easy solution would be to modify the definition such that it requires intentionality. A modified version may go as follows:

1a. An action is an event that is caused by a primary reason under a description which makes the primary reason relevant and that is intentional under some description.

Our job seems to be to define what it means for an event to be "intentional under some description". In order to define this term, let us examine some properties of intention.

One important property of intention is that it is causally transitive. For example, if I intend to perform action A, knowing that A causes B, then I intend to cause B. From this statement, we can extract another important property of intention: it requires knowledge. Imagine I modified the example as such:

If I intend to perform action A, and without my prior knowledge of the relationship between A and B, A causes B, then I intend to cause B.

This statement simply seems false. If I intend to open the door, and the opening of the door, without my knowledge, triggers a mechanism that hurts someone, it seems unfair to say that I intended to hurt someone. As such, it seems that in order for me to intend to perform action A, I must have knowledge of what A is. Intuitively, this seems reasonable; if I am unaware of the existence of action A as a possible action to take, I cannot possibly intend to perform A.

Let us work with these two properties:

Property 1: If I intend A, knowing that A causes B, I intend B.

Property 2: In order for me to intend to perform action A, I must be aware of action A as a possible action to take.

It seems that the DCC violates Property 2 of intention. We can restructure the DCC as follows:

1. Agent A has belief/desire pair B.

2. B causes event C (the agitation) which causes event D (knocking over of the glass)

3. A did not recognize event C as a possible action they could take.

Under our definition of intention, A would not intentionally be causing C because (3) shows the lack of Property 2.

Before I continue, I can foresee a possible objection that argues that requiring awareness of an action as a possible action is much too restrictive. Indeed, it seems that under this definition, some sudden, instinctive decisions or actions we make would be excluded from being intentional. While I do believe that these actions are not intentional actions, I recognize the merit behind the argument that these actions are intentional. As such, for the rest of the paper, I will make my requirement for intention less restrictive. Let us modify property 2 as such:

Property 2: In order for me to intend to perform action A, I must have some knowledge about action A.

I understand that I'm being vague about the type of knowledge about action A that is required. I myself do not know the answer to this and hope that it will reveal itself later on. For now, at the very least we can use this property 2 to argue that if I have no knowledge at all about action A, that I cannot intend action A. Let us see how we can incorporate this into the current definition.

1a. An action A is an event that is caused by a primary reason under a description which makes the primary reason relevant and that is intentional under some description.

2. A primary reason is the combination of a pro attitude towards actions with a certain property P, and a belief of the agent that action A has that property P.

It is clear that we are unable to consider an infinite number of actions. As such, there is a finite number of actions that we consider. Each belief is about a certain action. If we consider a finite number of actions, we have a finite number of beliefs. As such, there is a finite number of actions that we believe have property P. If we require that the action A be included in the finite number of actions that we believe have property P, we can guarantee that we have some knowledge about the action A, mainly whether or not it has property P and that action A is a possibility. As such, we can guarantee that action A has property 2 of intentionality. We can add the following premises to the definition:

3. In order for action A to be intentional under at least one description, the action must be included in the set of actions that the agent believes has property P.

4. There are a finite number of actions that the agent believes to have property P.

Premise 4 is necessary because if Sehon argues that somehow, we can have infinite beliefs about infinite number of actions, he can simply say that any and all DCC would be included in this set.

Let us see how this definition handles the DCC presented by Sehon. The event of involuntarily shaking and knocking over the glass does indeed have the property of fulfilling the desire to disrupt the speaker. However, considering that the philosopher was likely unaware that this involuntary shaking was even an option for them, we can conclude that involuntary shaking was not a part of the set of actions the philosopher believed had the property of fulfilling the desire to disrupt the speaker. As such, while the belief/desire pair causes the event, the event is not intentional under any condition, disqualifying it as an action.

Part 4: My Best Reply on Behalf of Sehon

I will consider an objection based on an even more deviant cases and objections against the logical strength of the definition.

First, let us consider an even more deviant case by modifying the DCC provided by Sehon as follows:

> There is a philosopher who has a condition in which sudden emotional disturbances cause the philosopher to twitch violently. The philosopher knows that seeing a certain speaker would cause a sudden emotional disturbance which would cause the philosopher to twitch violently. They do know that if they were to twitch violently, they will disrupt that speaker. However, the philosopher is unsure whether the speaker will attend the conference. Currently, the philosopher has no desire to disrupt the speaker.

> The philosopher chooses to attend the conference. It happens that the speaker attended the conference, and the philosopher sees the speaker. They suddenly have a desire to disrupt the speaker. This desire causes a sudden emotional disturbance within the philosopher causing them to twitch violently and disrupt the speaker.

In this case, can we say that the violent twitching is a) intentional or b) an action? Let us analyze this under our new definition.

The philosopher has a desire to disrupt the speaker. They know that twitching violently will achieve this purpose. As such, the act of twitching violently would be included in the set of actions that the agent believes will achieve their purpose. The desire causes an emotional disturbance that causes the philosopher to twitch violently, disrupting the speaker. In other words, the desire causes the philosopher to disrupt the speaker.

Under our definition of intentionality, it seems that this indeed is an action. However, I could very well see Sehon arguing that because the philosopher was not sure whether the speaker would attend and that initially, the philosopher had no intention of disrupting the speaker, the action is not fully intentional.

At the same time, I could argue that the philosopher recognized the possibility of seeing the speaker at the conference and still chose to go to the conference, meaning that there was some intentionality. Unless Sehon can show that the action is completely unintentional, the definition, although imperfect, has merit against Sehon's original DCCAs

This imperfection of the definition is likely caused by its logical imperfection. I will explore how Sehon might argue that the definition is logically weak.

Consider the following analogy: I know that all Sprite cans are green. As such, when deciding whether something is a Sprite can or not, I consider whether it is green. If it is not green, I can definitively say that the object is not a Sprite can. However, if it is green, that does not necessarily mean that the object is a Sprite can. There exists a possibility that the object is a Sprite can, but it is not guaranteed. Similarly, I know that all intentional actions require that I have some prior knowledge about the action. However, having prior knowledge about the action does not necessarily mean that the action is intentional. As such, while premises 3 and 4 that I add accurately and necessarily excludes any events that are unintentional under any description, the fact that an event satisfies premises 3 and 4 does not guarantee that it is intentional. More simply, the new definition resolves the issue of the original definition being too inclusive by excluding Sehan's DCCs. However, it does not resolve any issues of the original definition being too exclusive. As such, I am able to fend off DCCAs but I am unable to prove that the definition

is completely accurate. Sehon could argue that this logical imperfection means that the definition does not completely eliminate the possibility of a teleological explanation that cannot be reduced to a causal explanation.

Part 5: Further thoughts

A less important but interesting thought I had was about the size of the set of all actions with property P. Theoretically, it seems that there is an infinite number, or at least, an unfathomably large number of actions with property P. For example, if there is a set of bodily motions with property P, we can technically slightly modify each bodily motion in some way. As such, even if there are only extremely minute differences, there can be an infinite, or very large, number of actions with property P.

The reason why I consider whether the set of all actions with property P is infinite or not is to combat Sehon's formula. If Sehon is correct in that there can be an infinite number of DCCs then the set of all actions with property P cannot be finite. With the current model, our set of actions that we believe have property P is a finite subset of the infinite set of all actions with property P. As such, the infinite number of DCCs that Sehon's formula generates can still all be included in the infinite set of actions with property P. All of these cases would still be outside of the set of actions we believe to have property P and as such, they would not qualify as an action.

I do believe that my new definition effectively defends the Davidsonian definition against DCCAs. However, Sehon's argument that the definition is logically imperfect seems too strong to ignore. In fact, I question whether a perfect definition is even possible.

Recall my argument regarding the necessity of intuition when evaluating the accuracy of a definition. If it is true that intuition can be the only stopping point for evaluating a definition,

what happens if our intuition gives mixed results? For example, what if exactly half of all people think event A is an action while the other half believe that it is not an action? If the democratic nature of assessing intuition is off-putting, imagine that you are completely torn in the sense that event A seems both an action and a non-action. Then, for this event A, is it possible for any definition to provide an accurate classification? As such, if such a case exists, we must consider the possibility that either, a) there is a third category that is missing from the definition (i.e., actions, non-actions, semi-actions) or b) the "action-ness" of something requires a dimensional evaluation rather than a categorical classification. I can see cases like the even more deviant case that Sehon might use as an objection where intuition may produce mixed results.

Regardless, while a perfect definition may be impossible, I believe that Davidson's definition with my modification, at the very least, can hold its ground against any DCCAs. While, as Sehon might argue, the possibility of a teleological explanation that cannot be reduced to a causal explanation may exist, Davidson's definition serves its purpose as a method of reducing most, if not all, teleological explanations into causal explanations.