On the Physical Manifestation and Meaning of Parfit’s Relation R

Definitions
Further fact: a synthetic statement; true/false/indeterminate by virtue of empirical verification alone.

Kinetic Energy: ability to do work which a system possesses due to its being in motion.

Potential Energy: ability to do work which a system possesses due to its position.

Physicalism: Ontological doctrine that only the physical world exists.

Reductionism: All psychological states can be reduced to physical states.

Relation R: Psychological connectedness and/or continuity with any kind of cause. (Parfit).

Relation (P,Q): The unique set of relations any object P has to all other objects Q in a space.

Schrödinger Equation:  \[\hbar \frac{\partial}{\partial t} \Psi(r, t) = \left[\frac{-\hbar^2}{2\mu} \nabla^2 + V(r, t)\right] \Psi(r, t);\]

i.  \(i = \sqrt{-1}\)

ii.  \(\hbar = 1.054571800(13) \times 10^{-34}\) Joule seconds (reduced Plank constant)

iii.  \(\Psi(r, t)\) = The position-space wave-function of a position vector “r” at time “t”

iv.  \(\partial \) indicates the partial derivative of position vectors of the wave function with respect to time

v.  \(\mu\) indicates the “reduced mass” of a particle, which is a process of

vi.  \(\nabla^2 \Psi(r, t) = (\partial^2 \Psi(r, t) / \partial x) + (\partial^2 \Psi(r, t) / \partial y) + (\partial^2 \Psi(r, t) / \partial z)\) shows the second partial derivative of \(\Psi(r, t)\) in each of three spatial dimensions - in more familiar language this term denotes the kinetic energy of a particle.

vii.  \(V(r, t)\) indicates the potential energy of a particle.

viii. (vi) and (vii) together constitute the total energy of the system.

Note on the use of the Schrödinger Equation: the mathematical formulation listed above describes the evolution of a single particle in three dimensions over time as a probability function. The author lacks the mathematical sophistication to present a finalized form of the Schrödinger Equation for the system he proposes. It is important to this paper that the terms in brackets \(\left[\frac{-\hbar^2}{2\mu} \nabla^2 + V(r, t)\right]\) together represent the total energy of a single particle system. The terms in brackets are also called a Hamiltonian Operator. The solutions of the Schrödinger Equation for the term \(\Psi(r, t)\) are functions which describe the motion of waves. By using different Hamiltonian Operators which represent differently constituted physical systems, different solutions can be found to the Schrödinger Equation.
Introduction

In light of the recent death of Derek Parfit, his discussion in *Reasons and Persons* of “Liberation from the Self” takes on an indelible existential power. “When I believed that my existence was [a] further fact, I seemed imprisoned in myself. My life seemed like a glass tunnel, through which I was moving faster every year, and at the end of the tunnel there was darkness,” (Parfit, 451). As I sit composing this work in a room with large glass windows, through which can be seen the darkness of a cloudy Vermont night, the metaphor rings eerily true. Is the glass tunnel world plausible?

Parfit begins his discussion of personal identity with a section entitled “What We Believe Ourselves to Be” where he challenges our bodily and dualist theories of self with an intricate teletransportation thought experiment. We will discuss his thought experiment and the implications which he draws from it in due time, but first it is important for our purposes to understand the point of arrival before we step into the machine.

In this paper, I will develop and critique Parfit’s arguments concerning the nature of Relation-R. I will focus on the contingency of Relation-R within the framework of reductive physicalism to show how it is indeterminate whether Relation-R is an analytic or synthetic fact. I will end the paper with a fanciful speculation about the meaning of this indeterminacy, but, “as Einstein’s example showed, it can be useful to consider impossible thought-experiments,” (Parfit, 448).

What We Believe Ourselves to Be

Parfit prefaces his discussion of personal identity with a challenging thought experiment: imagine a teletransporter is built which will “destroy [one’s] brain and body while recording the exact states of all of [one’s] cells,” and then transmit this information via radio waves to a Replicator on Mars which reconstructs an exactly similar body from local materials on Mars (Parfit, 444). Supposedly, it is “in this body that [one] shall wake up,” (Parfit, 444). This is the case of “Simple Teletransportation,” and it is meant to challenge our intuitive bodily theorist notions of personal identity. If it is true, question beggingly, that “I shall wake up” in
the body on Mars, then it seems the materialist conception of the self cannot be true. Parfit then injects a strange twist: suppose the teletransporter does not destroy one’s body while recording the information necessary to construct an exact Replica on Mars. You step out of the machine on Earth, and an exact Replica of you steps out of the Replicator on Mars. In the words of Parfit’s contemporary, Pete Townshend, “who are you?”

The scenario wherein a person can coexist with their Replica is what Parfit calls a “branch-line” case and its logic is a primary motivator for revising Locke’s psychological theory of personal identity. Parfit claims Locke’s view that “experience-memory provides the criterion of personal identity” is not on its own, “a plausible view,” because it violates the principle of transitivity (Parfit, 446). Identity is a transitive relation, according to Parfit, and thus “the criterion of identity must be a transitive relation,” (Parfit, 447). Parfit expands Locke’s psychological view by appealing “to the concept of overlapping chains of experience-memories,” (Parfit, 446). He argues person X today could have direct memory connections to person Y in the past, or could have “continuity of memory” without such direct memory connections “if between X now and Y at that time there has been an overlapping chain of direct memories,” (Parfit, 446). This would make X and Y personally identical on the revised Lockean model, but, what about a branch-line case? If identity is a transitive relation, and X steps into the teletransporter on Earth in the branch-line scenario, then when Y steps out on Mars and Z (so called so as not to beg the question of the identity of X) steps out of the teletransporter on Earth, it follows Z must be identical to Y if Z and Y are both identical to X.

One approach would be to deny the existence of transitivity outright in the physical world. To argue Y and Z would have identical sets of memories assumes identical psychological experiences can manifest in diverse spatiotemporal locations in diverse contexts. Parfit writes, “reductionists admit there is a difference between numerical identity and exact similarity,” (Parfit, 447) but exact similarity is impossible in the physical universe, because every particle P is spatiotemporally related to the set of all other particles in the universe Q.
by some relation \((P,Q)\). If there are any discernible differences in the state of physical system \(X\) compared to \(Y\), then the two systems cannot be exactly similar. If \(X\) is constituted by particles which do not share all \((P,Q)\) relations with the particles that constitute \(Y\), then the systems are not exactly similar. Therefore, for two physical systems to have exactly similar \((P,Q)\) relations, they would need to be identical. With such a stringent notion of identity, it becomes obvious that neither \(Y\) or \(Z\) could be identical to \(X\), and thus, all notions of transitivity become senseless. A functionalist would argue psychological states could be instantiated in a plurality of physical systems, but an epistemologist would say the functionalist can never demonstrate this because the verification condition for being in a given psychological state is the quality of the state itself. A logician could chime in and say there is no further fact of the matter whether there could, or could not be identical psychological states instantiated in different physical systems. Given the complexity of this formulation, it isn’t surprising Parfit took another route, but it is important for our considerations because Parfit’s Relation \(R\), as we will see, describes psychological continuity independently of context, and this is his greatest error.

Transitivity and the Psychological Criterion

Instead, Parfit escapes the transitivity trap in the branch-line case via an even more radical revision of Locke. He uses his concept of overlapping chains of experience-memories and Lockean direct-memory connections to define two terms, psychological connectedness and psychological continuity. “Psychological connectedness is the holding of particular direct psychological connections,” and “Psychological continuity is the holding of overlapping chains of strong connectedness,” (Parfit, 446). Parfit shows strong connectedness (operationally defined as having enough direct psychological connections) is not a transitive relation. Parfit could be strongly connected to himself two or three days ago but “it does not follow that [Parfit is] strongly connected to [himself from] twenty years ago,” (Parfit, 446). Having established that connectedness is not a transitive relation, Parfit argues connectedness cannot be the criterion of personal identity over time because identity is
transitive. The concept of psychological continuity constituted by overlapping chains of direct experience-memories was engineered to be transitive. The Psychological Criterion for personal identity is thus defined as the “holding of facts like [X is psychologically continuous with Y; this continuity has the right kind of cause; there does not exist a different person who is also psychologically continuous with Y],” (Parfit, 447). It is interesting to note Parfit’s Psychological Criterion is defined by “facts,” rather than properties. It seems that the kind of relation Parfit wants to approach in his discussion of personal identity is in the same kind of ontological space as other logical relations. This is important because Parfit believes one’s continued existence does not “involve a further fact,” (447). One’s existence is predicated on physical and psychological continuity alone.

Concerning personal identity, Parfit writes “in some cases, there would be a real difference between some person’s being me and his being someone else who is merely exactly similar to me. Many people assume there must always be such a difference,” (Parfit, 447). But, if there were always a difference, that would depend on the existence of a further fact of the kind Parfit wants to refute. Appealing to the Simple Teletransportation thought-experiment, wherein one does not coexist with his Replica, Parfit claims “we could say here that my Replica will be me, or we could instead say that he will merely be someone else who is exactly like me,” (Parfit, 447). Given the discussion of exact similarity above, I differ with Parfit on this point in a very important way. Parfit believes that “for these to be competing hypotheses, [one’s] continued existence must involve a further fact,” (Parfit, 447). I don’t accept that there is no further fact, given Relation (P,Q), because the existence of this fact and a commitment to reductive physicalism motivates rejecting exact similarity by contextualizing all events.

Parfit uses the case of division to argue that the intrinsic nature of his relation to a future self must consist solely in psychological continuity. In division, Parfit’s brain is successfully split in half by a brilliant surgeon and implanted into two identical surviving bodies. The question of division is whether Parfit survives as one, both, or neither of the resulting people. If we are inclined to say Parfit has died, he retorts “it cannot
be the nature of my relation to each of the resulting people that, in this case, causes it to fail to be survival,” (Parfit, 449). Therefore, given that Parfit is psychologically continuous with both new bodies, he pragmatically declares, “the problem with double survival is that it does not fit the logic of identity,” (Parfit, 449). Parfit formulates “Relation R” to capture the intrinsic nature of the relation between two psychologically continuous persons: “R is psychological connectedness and/or psychological continuity with … any cause,” (Parfit, 449). To claim any causality could contribute to Relation R is bold, but warranted by Ockham’s Razor, because while the existence of causality is understood independently of psychological continuity, the existence of particular causes which can only affect psychological continuity is clumsy anthropocentrism or psychocentrism.

To address critics who would say we care more about personal identity than mere psychological continuity, Parfit analyzes personal identity (PI) and its relation to Relation-R.

He claims \( \text{PI} = \text{R} + \text{U} \)

Where U indicates uniqueness. He claims that to be “R-related to some future person, the presence of U makes no difference to the intrinsic nature of my relation to this person,” and therefore, U cannot matter all that much. This makes sense for Parfit, because he has no need for U in his picture of reality because he believes exact similarity without identity is possible. In the context of reductive physicalism, if R is the intrinsic psychological continuity between oneself at \( T_1 \) and \( T_2 \), then \( U = (P,Q) \), which, we remember was the set of relations of each particle in a system to all other particles in the universe. On this account, \( (P,Q) \) is every bit as important as R. Rather than take a logical positivist view and miss half the picture, take a logical polarist view and have twice the fun.

**Further Facts and the Nature of Relation-R**

Another strange result of the formulation of the psychological criterion (and thus Relation-R) in the language of facts is that because facts canonically come in two varieties, analytic and synthetic, we can ask the question “Is Relation-R an analytic or synthetic fact?” It would be analytic if Relation-R was true, that is to
say if a system were R-related, by virtue of meaning alone. It would be synthetic if Relation-R were true, that is to say if a system were R-related, by virtue of empirical observation. However, presupposed in the idea of psychological continuity is consciousness. So it seems at first glance that a system has a consciousness if it is R-related, and is R-related if it has a consciousness. The circularity points toward analyticity. However, the nature of consciousness is such that it is the only epistemically privileged empirical verification of Relation-R. This point may be contentious, because there are those reductive physicalists who say that if we knew the exact states of every particle in one’s body, the epistemic privilege they enjoy would disappear. True, I say, and if we could measure the ~10 octillion atoms in a human body and discover all their states, we would have a lot of information, but remember to be fully reductive physicalists we need to know (P,Q), which is clearly impossible. It is a further fact which defines the verification conditions for Relation-R, so R must be synthetic and epistemically privileged. In fact, because Relation-R only holds for a conscious system, the empirical observation necessarily generates the truth of the fact.

My hypothesis, however speculative, is that the logical weirdness that is happening is not actually all that weird. It’s a superposition of exactly the type that is observed in quantum mechanics. Superposition occurs in linear systems (the Schrödinger Equation is a linear equation) and means that for any given quantum state of a system, it is possible to decompose the quantum state into two other equally valid quantum states whose characteristics could add together to yield the first quantum state (Griffiths, 2017). We can imagine a set of possible R-Relations (where psychological continuity remains constant, but the cause is different due to differences in (P,Q)). If this set of R-Relations were a set of solutions to a Schrödinger Equation which also gave an adequate approximation of (P,Q), the superpositional characteristics of R-Relation would be explained and it could be demonstrated that consciousness collapses the wave-function.

Here’s how it would work:

This is the Schrödinger Equation for a single, nonrelativistic particle moving in three dimensions.

\[ i\hbar \frac{\partial \Psi (r,t)}{\partial t} = \left[ -\frac{\hbar^2}{2\mu} \nabla^2 + V(r,t) \right] \Psi (r,t) \] (Griffiths, 2017).
Solving this equation generates a wave-function which is a probability distribution of finding a particle at a
given time (Griffiths, 2017). Solutions to the equation for even the simplest systems are extremely difficult to
derive - senior physics majors at Middlebury College generally work on solving it for the hydrogen atom. The
term in the brackets is called the Hamiltonian Operator and it represents the potential and kinetic energy of a
system (in this case, a single nonrelativistic particle in three dimensions). The function is solved for $\Psi(r,t)$
(the presence of the Hamiltonian Operator means we can’t simply divide out this term). I propose something
like the following:

$$i\hbar \frac{\partial}{\partial t} \Psi(r,t) = \left[(P,Q) \nabla^2 \right] \Psi(r,t) \Rightarrow \Psi(r,t) = R_n$$

Where $(P,Q)$ is the potential and kinetic energy of every particle in the universe (dressing up Relation $(P,Q)$ in
physics language), $\nabla^2$ indicates a partial derivative of $\Psi(r,t)$ in three dimensions, and $R_n$ indicates a group of
R-Relations: $R_1, R_2, \ldots, R_n$ that would be the solution of this equation. We would then expect to find that any
observed R-Relation to fall on the wave-function, and indeed would be a collapse of the superposition of all
possible R-Relations to the single observed R-Relation. In practice, this means that every conscious
observation in the universe is predicted by a single wave-function. Of course, this is merely a speculative
consideration, but the point is this: Parfit comes to be liberated from his sense of self by reflecting on the fact
that his very manifestation in the world was not a deep, further fact about the nature of reality. But, if this
sort of picture is true, we go from “[living] in the open air” to being the very medium by which air manifests
as air. It is the existence of your consciousness, empirically verified through the collapse of the wave-function,
that is the deep further fact.

Reference