Uploading and Personal Identity, part 1

Let’s apply the lessons learned so far to the following question:

Could you survive death by being “uploaded” into a computer or robot body?

Note that there are already people working on this. Some estimate that this technology will be available by as early as 2045 (also here; see also this interesting interview between Neil deGrasse Tyson and Ray Kurzweil). As odd as it seems, it is quite possible that this is a real question that you will face within your own lifetime.

Against Survival (The Pessimistic View): Since the uploading case is similar to a case of teletransportation, initially, the intuitions would seem to be these:

- **Body Theorists:** You WILL NOT survive an uploading because the uploaded person will not have the same body as you.

- **Psychological Continuity Theorists:** You WILL survive because the uploaded mind will be psychologically continuous with your own.

- **Soul Theorists:** It is unclear whether or not you would survive. Does your soul relocate to the uploaded mind?

It is the psychological view that seems the most promising for an optimistic view about surviving uploading. But, the same problem that arises for the psychological view in teletransportation cases arises for cases of uploading. Consider:

**Uploading Without Destruction:** Imagine that Dave uploads his mind. Every synaptic connection is perfectly replicated. The uploaded individual (call him ‘DigiDave’) claims to be Dave, has all of Dave’s memories, and so on. However, the original Dave is NOT destroyed. The physical Dave (call him ‘BioDave’) is still alive and well—and both BioDave and DigiDave claim to be the REAL Dave.

Contrast this with the following:

**Destructive Uploading:** Like the previous case, except that the original, physical Dave dies, or is destroyed; i.e., the only survivor is DigiDave.

Who is the real Dave? Most intuit that, in Uploading Without Destruction, the original Dave (BioDave) is the real one. But, then, we have the following argument:
Argument Against Uploading Survival

1. In Uploading Without Destruction, Dave clearly survives as BioDave, and not as DigiDave (i.e., Dave = BioDave, but Dave ≠ DigiDave).
2. If Dave doesn’t survive as DigiDave in Uploading Without Destruction, then he also doesn’t survive as DigiDave in Destructive Uploading.
3. Therefore, you cannot survive an uploading.

Objections: We can either accept the conclusion, or reject one of the two premises:

(a) Rejecting premise 1: Perhaps Dave survives as both BioDave AND DigiDave? No, that won’t work. For, if Dave IS BioDave (i.e., Dave and BioDave are not two distinct persons, but are, rather ONE person), then if it is also the case that Dave IS DigiDave, then it follows (by transitivity) that BioDave is DigiDave (i.e., they are not two persons, but one)! That is:

If <Dave = BioDave>, and <Dave = DigiDave>, then <BioDave = DigiDave>.

But that is absurd. BioDave and DigiDave are not one person, but two. They have distinct conscious experiences. If I pinch BioDave, DigiDave will not feel it. If we delete DigiDave, BioDave will survive (and vice versa). If BioDave commits a crime, DigiDave is undeserving of blame. And so on.

(b) Rejecting premise 2: Consider the following view of personal identity:

- Closest Continuer Theorists: You survive wherever your closest continuer exists; i.e., YOU are whoever has the greatest claim to being you, which may vary from case to case. For example:

  (a) In teletransportation where the original is NOT destroyed, the original is you (because that person has the same body AND psychology as you, they have a greater degree of continuity with you than the transported person, who MERELY has continuity of psychology, but not of body). Similarly, in Uploading WITHOUT Destruction, BioDave is Dave for the same reason.

  (b) But, in teletransportation when the original IS destroyed, the transported individual IS you (because that person is psychologically continuous with you, they have the greatest degree of continuity with your self). Similarly, in Destructive Uploading, DigiDave IS Dave, because he has the greatest claim on Dave’s identity.

If this theory is correct, then we have grounds for rejecting P2. But is the Closest Continuer Theory plausible?
Problem: If this view were correct, then identity would be an unacceptably extrinsic matter. For instance, if I just stepped out of a teletransporter, or just woke up in my new robot body, I couldn’t know whether I am the REAL Chad or not unless you confirm to me that the original Chad was destroyed. If he WAS destroyed, then I am the real Chad. If he was NOT destroyed, then I am a mere copy. Worse still, what if the destruction of the original is delayed for a month. Am I a mere copy for that month, and then I suddenly BECOME the real Chad once the original is destroyed? That seems utterly absurd.

[Note that closest-continuer theories are also unable to address the problems that arise in Derek Parfit's split-brain cases. See that lecture for more.]

In conclusion, it seems that we could NOT survive a digital uploading. Sorry!

(But stay tuned for a more optimistic conclusion in a future lecture!)