Equivalent Utility Transport

by Sven Nilsen, 2020

In this paper I discuss a technique for solving difficult ethical dilemmas by transforming the problems into equivalent but easier to solve problems. This axiom addresses some core problems of ethics.

In the famous Trolley Problem, the operator of the level has two choices:

A) Do nothing and allow the trolley to kill five people on the main track
B) Pull the lever, diverting the trolley onto the side track where it will kill one person

Simplified, the cost in terms of number of lives:

A) -5
B) -1

Or, one can view the problem as the number of lives saved:

A) 1
B) 5

What makes the Trolley problem difficult to solve is that the operator is actively pulling the lever, leaving out some missing information that seems important for decision making.

Normally, the debate is about whether it is right to pull the lever or not for this particular situation. However, a different approach is to study the transformation of the situation into an equivalent problem from a decision theoretical view point, where the problem can be easily solved. The subject of debate then becomes whether a such transformation is ethical.

The axiom of equivalent utility transport states that under utility maximization, adding or subtracting utility to every choice does not change the optimal decision. Hence, by subtracting one life saved to every choice, the following outcomes can be used to inform the decision maker:

A) 0
B) 4

Imagine a scenario where by doing nothing, no lives were saved. By pulling the lever, four lives were saved. Most people would agree that taking action to save lives is acceptable. View as costs:

A) -4
B) 0

Most people would agree that acting to reduce loss of lives is acceptable. Yet, both these examples directly contradicts the intuition that by pulling the lever, one is contributing to the death of a person.

Hence, the axiom of equivalent utility transport seems to hint at core problems of ethics: Inconsistency.

If the axiom is accepted, then seemingly difficult ethical dilemmas gets easy to solve. However, in practice these dilemmas seem still problematic in many ways.

I believe that the focus on pulling the level, or not, is missing the bigger picture. The fact that you can construct an equivalent utility transport that leads to opposite conclusions, or at best large amounts of uncertainty, points to a much deeper problem.

For example, is it correct to say that pulling the lever saves 4 lives?

In the core axiom of path semantics, if two symbols are equivalent, then their associated symbols are also equivalent. So, using the core axiom of path semantics in this context means that one can say pulling the lever saves 4 lives. However, does the core axiom of path semantics hold for ethics?

A counter example can be constructed by saying that 0 lives were lost when pulling the lever.

This is not true. Yet, one can say that 0 relative lives were lost when pulling the lever. The argument is that the transformed thought experiment measures relative utility and not absolute utility.

If person A is going to die or person B is going to die, but not both, then it seems a much harder ethical problem to solve than person A is going to die anyway and person B will survive no matter what happens. The first problem seems hard, the second seems trivial, because a choice has no impact at all. Yet, applying the axiom of equivalent utility transport leads to the conclusion that those two problems are the same. Does this seem surprising?

Proof:

∵ [(-1, 0), (0, -1)] ~= [(0, 0), (0, 0)]
∵ [(0, -1), (0, -1)] ~= [(0, 0), (0, 0)]

∴ [(-1, 0), (0, -1)] ~= [(0, -1), (0, -1)]

The utility is summed over a vector containing a component per person, of which one is permitted to add or subtract values as long the same amount is added or subtracted to all choices.

The same axiom of equivalent utility transport permits swapping values in the underlying vector.

It might look like two vastly different problems at first. This is because in the first problem, the decision maker is perceived to actively making a choice, causing the death of one or the other. In the second problem, the decision maker is perceived to make no choices since the outcome is qualitatively the same and therefore the decision maker is free of responsibility.

The optimal decision is not going to be impacted by adding or subtracting utility to each choice, but that does not say that it is possible to determine which decision is optimal. On the other hand, if contraction of vectors is permitted, a problem with equal utility for all choices is equivalent to no choices. This becomes a proof that the optimal decision is undecidable.

Another point of view is that only because a hard problem can be equivalent to an easy one, does not mean that making decisions becomes easy. Inconsistency might be desired to some degree for reasons that are not bound to particular situations. A possible way forward is to model side effects of transport.