

Comments on Jongmin Jerome Baek Paper “How to Solve Moral Conundrums with  
Computability Theory”  
By Liam Grantham

In his paper “How to Solve Moral Conundrums with Computability Theory”, Jongmin Jerome Baek proposes a framework to better deal with various moral problems by reducing value statements to a logical format. He begins the paper by briefly contextualizing the history of computability theory in moral philosophy. He introduces Gödel’s Incompleteness Theorem which he formulates roughly as showing that “consistency and completeness cannot coexist in a formal system”. He criticises this approach in past applications because it has been used to prove both the existence and nonexistence of God and both the possibility and impossibility of a fully conscious A.I., but states that such applications have been misinterpretations of the ideals. He states that a proper interpretation of such a theory claims that when we know something we change the way in which the knowledge is classified; the limits of that knowledge are broadened, introducing more avenues for understanding. In a practical sense, once we know something we can use it to gain more information. He then uses formal language to describe the way that knowledge acts in this way, coining it uncomputable. He states that free will is uncomputable because of this paradox of gaining knowledge. The belief that free will is good, then uncomputability is good. As a result, the moral value statements we make are incomplete. When we label something as morally good we are doing so because the uncomputable process we are using has been previously deemed successful in producing accurate judgements. However, the implications of Gödel’s Incompleteness Theorem mean that if we know that the process is good to begin with,

that same process cannot predict anything that is better than the previous good. With this understanding, he addresses moral questions like what is a good life and how should we concern ourselves with future generations.

The paper takes an admittedly highly formal (as in logical) approach to the classification of moral principles. However, it notes that this classification relies on a contradiction. It is not uncommon for the primary classifier of a notion to be a contradiction and does not in itself provide worry for the fallibility of the claim. However, once concern that he might want to address is whether or not what follows after the classification is acted on by that contradiction. He notes the difficulty that one would have using such a theorem to describe what is good, in fact it is impossible to fully know what is good. As a result the classification of a process as uncomputable, and therefore good, would be equally difficult, if fact impossible. If one cannot determine which processes are uncomputable then the value judgement needed is exposed by the same contradiction that it draws its power from. It is the paradox of the chicken and the egg. How can you determine a process that can be defined as uncomputable, without using a process that is equally uncomputable? But then how do you determine whether or not that process is uncomputable?

This objection could make the classification seem to be a non starter when the epistemology of moral values are questioned. However, once this objection is quelled, the implications that it has on moral problems is indeed interesting.

In addition, I believe that the questions answered toward the end should play a larger role in the presentation. The implications of such a classification better lend themselves to a presentation than the classification itself.