

ロボット倫理学における道徳的行為者性

福田 朗

久木田の「ロボット倫理学の可能性」(2009)によればロボット倫理学は軍事用ロボットや情報収集ロボットによるプライバシーの侵害の問題、AIや人工知能の倫理などを議論対象とする学問領域である。筆者は①ロボットに道徳的行為者性を帰属できるか、②ロボットに道徳的判断をさせてよいのか、ということについて明らかにしたいと考え、久木田、ウォラック、メッツインガーの論を取り上げ、考察をおこなった。

久木田は、ロボットの道徳的行為者性を否定する。また、道徳的判断を下すような自律性をもったロボットは、理論的には実現可能であるかもしれないが、実現可能性は低いという見解を示す。ウォラックは、道徳的判断を下すようなロボット(AMA:Artificial Moral Agent)が開発されることは不可避であると主張する。また、AMAを開発していく上で、既存の倫理学や技術哲学について再構築していく必要があると論じる。メッツインガーは、神経科学や脳科学の分野の知見から、我々の感じる「自己」の存在を否定し、「意識的自己モデル」という概念を提示する。また、それらのモデルを備えたロボットをエゴ・マシーンと呼び、その実現可能性について認める一方で、エゴ・マシーンを作り出すことにより、全世界の苦痛の総量を増やすことになるため、開発についての取り組みを一切執り行うべきではないと主張する。

三者の主張を比較した結果、①と②について以下のことが明らかになった。

まず、①の論点においては、カント的倫理論と哲学的決定論の間にある自由意志の存在の有無についての議論が背景にあること、カント的倫理論ではロボットに道徳的行為者性を帰属することは難しいが、ウォラックが支持するフロリディの倫理論においては行為者性の帰属は可能であるということ、また責任主体については自動運転車の実例を検討するならば、どちらの倫理論を用いても、現実的にはロボットに責任を帰属できないということが示された。

②の論点に関しては、人工意識の開発に否定的なメッツインガーの論拠が、ウォラックの指摘する先進技術に対する未来主義的な恐れであること、道徳的判断をさせてよいとするウォラックの主張にはたとえ問題点があるとしても尊重されるべき点があるということが明らかになった。一方で、ロボットの道徳的判断の基準となる倫理論の具体的な検討などについては課題が残った。ロボットの行為を正しく評価する枠組みがどのようなものかを考えることが今後の課題である。

(指導教員 横山 幹子)