



Enseñanza de Modelos de razonamiento práctico a agentes artificiales o sobre una ética para la inteligencia artificial

Miguel A. Fonseca

Universidad La Gran Colombia

miguel.fonseca@ugc.edu.co

Palabras clave: enseñanza, ética, inteligencia artificial, agencia moral artificial, Tomás de Aquino.

Resumen

El impacto de los desarrollos de la Inteligencia Artificial (IA) en el desarrollo de la humanidad, en un corto y mediano plazo, suponen el planteamiento de reflexiones fundamentales sobre la relación de lo humano con este tipo de sistemas y artefactos. Las preguntas que ineluctablemente surgen son, entre otras: ¿Cómo debemos relacionarnos con dichos sistemas?; ¿Cuáles son los alcances de la acción de la IA para con nosotros y nuestro entorno?; ¿Cómo deberían comportarse tales artefactos? En definitiva, dichas preguntas se sintetizan en el problema de qué debemos hacer con ellos y qué deberían hacer ellos con respecto a nosotros.

Las respuestas a estos interrogantes dependen de nuestra concepción ontológica de la IA en tanto objetos o sujetos (Müller, 2021). El primero se refiere principalmente al uso de la IA, y así, a elementos éticos derivados, tales como, manipulación, privacidad, sesgo, interacción y otorgamiento de autonomía. Desde la segunda concepción, se comprende a la IA como un sujeto y, por lo tanto, se derivan problemas éticos como las denominadas éticas de las máquinas y la agencia moral artificial (Moor, 2006).

En el presente trabajo se asumirá como campo de la reflexión la segunda perspectiva o concepción, y se entenderán a los sistemas y artefactos de la IA como sujetos doxásticos con capacidad relativa de razonamiento práctico. Dada esta premisa, se buscará comprobar, en el marco de una IA débil, que no se involucra directamente con problemas como la singularidad de la IA (Kurzweil, 2005), que existen modelos de razonamiento práctico susceptibles de ser enseñados a agentes doxásticos artificiales (Fonseca, 2020).



VI CONGRESO LATINOAMERICANO DE FILOSOFÍA DE LA EDUCACIÓN BOGOTÁ, JULIO 12 - 14 DE 2023

Hospitalidad y reencuentro: volvernos a ver para
pensar el sentido de la educación y de la filosofía

Como soporte básico del argumento, se expondrá la construcción de un modelo de razonamiento práctico artificial, basado en la ética de la virtud de Tomás de Aquino (Aquino, 1981), como un pretexto y recurso heurístico relevante.



Referencias

- Aquinas, T. (1981). *Summa Theologiae (A Treatise on Theology)*, Parts I [1265–8], I–II [1271–2], II–II [c.1271], III [1272–3]. London: Fathers of the English Dominican Province. Cristian Classics Edition.
- Fonseca, M. (2020). *Agencia Eidética: agencia material, artefactos y agentes eidéticos*, En: Moreno J, et al, (2020). *Tecnología, Agencia y Transhumanismo*. pp 39-46, Bogotá: USTA
- Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*, London: Viking.
- Moor, James H., 2006, “The Nature, Importance, and Difficulty of Machine Ethics”, *IEEE Intelligent Systems*, 21(4): 18–21. doi:10.1109/MIS.2006.80
- Müller, V. (2021). "Ethics of Artificial Intelligence and Robotics", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>