

# Resumen

Bajo la visión de la inteligencia artificial, un agente inteligente es una entidad autónoma la cual interactúa en un entorno a través de observaciones y acciones, tratando de lograr uno o más objetivos con la ayuda de varias señales llamadas recompensas. La creación de agentes inteligentes está proliferando durante las últimas décadas, y la evaluación de su inteligencia es un asunto fundamental para su entendimiento, construcción y mejora.

Recientemente la inteligencia social está obteniendo especial atención en la creación de agentes inteligentes debido a la visión actual de la inteligencia humana como altamente social. Normalmente la inteligencia social en sistemas naturales y artificiales se mide mediante la evaluación de rasgos asociados o tareas que se consideran que representan algunas facetas del comportamiento social. La agrupación de estos rasgos o tareas se utiliza entonces para configurar una noción operacional de inteligencia social. Sin embargo, esta noción operacional no representa fielmente a la inteligencia social y no sería posible una definición siguiendo este principio. En su lugar, en esta tesis investigamos la evaluación de la inteligencia social de un modo más formal y general, considerando la interacción del agente a evaluar con otros agentes.

En esta tesis analizamos las implicaciones de evaluar la inteligencia social utilizando un test que evalúe la inteligencia general. Con este objetivo incluimos otros agentes en un entorno inicialmente diseñado para un único agente con el fin de averiguar qué cuestiones aparecen cuando evaluamos a un agente en un contexto con otros agentes. A partir de este análisis obtenemos información útil para la evaluación de la inteligencia social.

A partir de las lecciones aprendidas identificamos los componentes que deberían considerarse al medir la inteligencia social y proporcionamos una definición *formal y parametrizada* de esta inteligencia social. Esta definición calcula la inteligencia social de un agente como su rendimiento esperado en un conjunto de entornos y con un conjunto de otros agentes organizados en equipos y distribuidos en alineaciones, reinterpretando apropiadamente las recompensas. Esto se concibe como una herramienta para definir bancos de prueba de inteligencia social donde podamos generar varios grados de comportamientos competitivos y cooperativos. Probamos esta definición analizando experimentalmente la influencia de los equipos y las alineaciones de agentes en varios sistemas multiagente con variantes de agentes Q-learning.

Sin embargo, no todos los bancos de prueba son apropiados para la evaluación de la inteligencia social. Para facilitar el análisis de un banco de pruebas de inteligencia social, proporcionamos algunos modelos de propiedades formales sobre la inteligencia social con el objetivo de caracterizar el banco de pruebas y así valorar su idoneidad. Finalmente, usamos las propiedades presentadas para caracterizar algunos juegos sociales y entornos multiagente, hacemos una comparación entre ellos y discutimos sus puntos fuertes y débiles para ser usados en la evaluación de la inteligencia social.