A Method to Assess AI Consciousness

Harry H. Porter III — Portland State University

Goal

Measure the degree of consciousness in an AI.
On a scale from 0-100, like an IQ test

Assumption

Consciousness is a collection of mental abilities.
Abilities are observed ⇒ Consciousness exists
Consciousness exists ⇒ Abilities are present

Method

44 questions about the AI's mental capabilities.

Answer each question
0 = ability is not present
1 = some ability
2 = substantial, but sub-human ability
3 = human level of performance
4 = exceeds human ability

Normalize and compute the final score
0 = no consciousness
100 = human level of consciousness
133 = maximum score possible

Problem

No agreed upon definition of consciousness.

Our Answer

Our definition is implicit in this assessment.

Pros

• Evaluation is easy and quick.
• The result is on a scale (unlike the Turing Test).

Example Questions

• Can the AI reason and make logical inferences?
• Does the AI have working memory? Long term memory?
• Does the AI have situational awareness (time, location)?
• Can the AI use natural language?
• Does the AI have goals, plans?
• Does the AI have likes, opinions?
• Does the AI have emotions, moods?
• Does the AI remember past personal experiences?
• Does the AI learn?
• Does the AI have curiosity?
• Does the AI have self-knowledge, an idea of who it is?
• Does the AI have a mental model of itself, of its mind?
• Can the AI perceive its own thoughts, feelings?
• Can the AI describe its thought processes?
• Can the AI control/alter its own thoughts?
• Does the AI have knowledge about others?
• Does the AI know about the other person's mental state?
• Can the AI assess the other person's intelligence?
• Does the AI seek info about the other's emotional state?
• Does the AI have a desire to imitate others?

Claim

***If these mental features are present,
the AI would appear to be conscious!***

Questions not included

• Does the AI hear an internal voice?
• Does the AI have visual, spatial abilities?
• Does the AI experience pain, pleasure, love, color?
• Does the AI have a body?
Are these mental attributes necessary for consciousness?

Cons

• Questions are difficult to evaluate precisely.
(Results may be imprecise, w/ error probabilities)
• May not match your definition of consciousness.

Paper at: `cs.pdx.edu/~harry/musings/ConsciousnessAssessment.pdf`