

Mentes, cerebros y programas

John R. Searle

Supongamos que estoy encerrado en una pieza y se me entrega un gran lote de escritos en chino. Supongamos, además, como es realmente el caso, que no conozco nada de chino, ni escrito ni hablado, y que ni siquiera confío en que pudiera distinguir escritura china de escritura japonesa, o de dibujos sin sentido. Supongamos, en seguida, que luego de este primer lote de escritura china se me entrega un segundo lote de escritos en chino junto a un conjunto de reglas para correlacionar el segundo lote con el primer lote. Las reglas están en inglés, y comprendo estas reglas tan bien como cualquier hablante nativo de inglés. Me permiten correlacionar un conjunto de símbolos formales con otro conjunto de símbolos formales, y lo único que significa “formal” aquí es que puedo identificar estos símbolos enteramente por sus formas. Ahora supongamos que se me entrega también un tercer lote de símbolos chinos junto a algunas instrucciones, nuevamente en inglés, que me permiten correlacionar elementos de este tercer lote con los otros dos lotes, y estas reglas me indican cómo debo entregar de vuelta ciertos símbolos chinos de ciertas formas en respuesta a ciertas formas que se me entregaron en el tercer lote.

Sin ser de mi conocimiento, las personas que me entregan estos símbolos llaman al primer lote un “guión”, al segundo una “historia”, y al tercero “preguntas”. Aún más, llaman a los símbolos que yo entrego de vuelta en respuesta al tercer lote “respuestas a las preguntas”, y al conjunto de reglas en inglés que me entregaron, “programa”. Para complicar un poquito este asunto, imagine que estas personas también me entregan historias en inglés, las que entiendo, y luego me hacen preguntas en inglés acerca de estas historias, y yo les respondo en inglés. Supongamos también que luego de un rato me vuelvo tan bueno siguiendo las instrucciones para manipular estos símbolos chinos, y los programadores se vuelven tan buenos escribiendo los programas, que, desde un punto de vista externo —esto es, desde el punto de vista de alguien fuera de la pieza en la que estoy encerrado—, mis respuestas a las preguntas son indistinguibles de las que darían hablantes nativos de chino. Nadie que viese mis respuestas podría decir que no hablo una palabra de chino. Supongamos, además, que mis respuestas a las preguntas en inglés son, como lo serían sin duda, indistinguibles de las de otros hablantes nativos de inglés, por la simple razón de que soy un hablante nativo de inglés. Desde un punto de vista externo, desde el punto de vista de alguien leyendo mis “respuestas”, las respuestas a las preguntas en chino y a las preguntas en inglés son igualmente buenas. Pero en el caso chino, a diferencia del caso inglés, produzco las respuestas manipulando símbolos formales sin interpretar. En lo que concierne al chino, simplemente me comporto como un computador; realizo operaciones computacionales sobre elementos formalmente especificados. Para los propósitos del chino, soy simplemente la realización de un programa computacional.

1. Imaginemos que una persona pide a ChatGPT las respuestas a un examen y la IA se las da correctamente. Podemos decir que ChatGPT sabe las respuestas? Por qué?