UNIVERSIDAD AUTÓNOMA METROPOLITANA

Unidad Iztapalapa

División de Ciencias Sociales y Humanidades

Departamento de Filosofía

EL DUALISMO MENTE-CUERPO DE ALAN TURING Y LA CRÍTICA DE JOHN SEARLE

Que para optar por el grado de:

Licenciada en filosofía

Presenta:

Itzel Donají Pérez Martínez

Director de la tesina:

Dr. Silvio Mota Pinto

México, CDMX., Febrero de 2020

ÍNDICE

INTRODUCCIÓN……4

CAPÍTULO I: DESCARTES, ¿QUÉ ES EL PENSAMIENTO?.....6

1.1 Introducción……..……6

1.2 Dualismo cartesiano ……..……..6

1.2.1 Sustancia pensante…….……7

1.2.2 Substancia material o corpórea..…….…..…11

1.2.3 Interacción cartesiana mente-cuerpo……..…….14

1.3 ¿Es eso el pensamiento? ……...…….……..17

1.3.1 Error categorial…….17

1.3.2 Otros aspectos sobre el ejemplo anterior……...……20

1.3.3 Primera persona……21

1.3.4 Tercera persona……..……24

1.4 Distinción mente-cuerpo…….…….....26

1.5 Conclusión ……..31

CAPÍTULO II: TURING DUALIDAD SOFTWARE-HARDWARE……32

2.1 Introducción …….……...32

2.2 Precedentes……..32

2.3 Turing……..37

2.3.1 Máquinas de Turing……..39

2.4 Otros argumentos funcionalistas……...…...50

2.5 Juego de la imitación o Test de Turing…….…51

2.6 Conclusión ……..55

CAPÍTULO III: SEARLE Y EL EXPERIMENTO DE LA HABITACIÓN CHINA…56

3.1 Introducción…….56

3.2 Pensamiento en las máquinas de Turing……...56

3.2.1 La habitación china…….……..58

3.2.2 Intencionalidad…….……65

3.2.3 Relación mente-cuerpo…….……..……..70

3.3 Conclusión…….……..75

CONCLUSIÓN……76
BIBLIOGRAFÍA……..……78

INTRODUCCIÓN

¿Crees que las computadoras pueden pensar?, y que tal los animales ¿pueden pensar? Tal vez el pensamiento es algo propio de la especie humana ¿tú qué opinas? Sin duda la cuestión respecto de ¿qué es el pensamiento?, ¿cómo funciona?, ¿cómo surge? es uno de los temas más polémicos en diversas áreas del conocimiento como lo es en la filosofía, la psicología, la inteligencia artificial, la biología, la neurología, etc. El pensamiento es uno de los misterios que el conocimiento humano aun no consigue develar. Por ello ha habido diversos movimientos a lo largo de la historia humana que se han dedicado al estudio del pensamiento, más concretamente al estudio del pensamiento humano, puesto que igualmente se ha debatido si es posible el pensamiento en otros seres no humanos como los animales, extraterrestres o incluso las máquinas.

A la fecha no parece haber ninguna teoría de la mente que consiga una aceptación considerable como para declararla como "la teoría de la mente", muy por el contrario, hay tantas teorías de la mente que proliferan por todas partes, cada una con sus fortalezas y debilidades que parece difícil desecharlas por entero. Dentro de todas esas teorías, sin embargo, hay dos corrientes que parecen disputarse el entendimiento del pensamiento. Éstas son, por un lado, la idea de que la mente es un ente sustancial distinto del cuerpo: la dualidad mente-cuerpo; por otro lado, está la idea de que el pensamiento es parte del mundo físico. Ambas corrientes han tenido diversos portavoces que las defiendan, pero entre ellos Alan Turing destaca por planear la idea de que somos capaces de crear seres no humanos con pensamiento, una idea que sin duda embargaría de curiosidad y duda a muchos. Tales seres que señala Turing son las computadoras, pues como informático advierte que estas son capaces de muchas habilidades propias del humano.

Desde el planteamiento de Turing a la actualidad esta idea ha ido desarrollándose exponencialmente de la misma manera en que las computadoras se han vuelto mucho más hábiles y sofisticadas. El tema de la mentalidad de las computadoras actualmente prolifera en toda nuestra cultura al punto en que incluso vemos como constantemente surge mucho contenido audiovisual que habla al respecto. Es un tema que sin duda ha tomado popularidad

y al cual se le dedican innumerables investigaciones y fondos económicos; los países incluso compiten por tener las computadoras más eficientes o inteligentes. Pero, ¿realmente podemos decir que eso que hace una computadora es más que solo un cierto tipo de inteligencia, a saber, realmente podemos decir como Turing que eso es un pensamiento como lo es el pensamiento humano? O eso que Turing describe como pensamiento de la máquina es algo distinto de lo que hacemos los humanos y entendemos como pensamiento.

Aquí se explorarán algunas teorías de la mente para acercarnos y entender el pensamiento de Turing, así como se buscará una respuesta a si eso que Turing advertía en las computadoras realmente podría ser considerado como pensamiento. Por ello, se partirá de una de las teorías del pensamiento más populares y debatidas, a saber, la dualidad cartesiana del pensamiento pues veremos que Descartes también toca este planteamiento respecto de las máquinas y también se verán algunas críticas tanto hacia Descartes como hacia Turing.

CAPÍTULO I:

DESCARTES: ¿QUÉ ES EL PENSAMIENTO?

1.1 Introducción

La cuestión sobre el pensamiento es una de las preguntas de gran interés filosófico que ha dado lugar a diversas teorías en busca de describir o proponer la manera en que se da y opera el pensamiento. Una de las más destacadas ha sido el dualismo cartesiano en la cual se ha propuesto que el pensamiento es un ente distinto del de los cuerpos, el cual opera con ciertos criterios que se consideran mentales y parecen mantener una relación con el cuerpo físico que los contiene. Pues, se contempla que esta mentalidad o pensamiento imprime ciertas facultades, propiedades y características propias, como la intelección, sentido o propósito, a las acciones de los cuerpos que la poseen. Empero, se dirá que no cualquier cuerpo puede poseer mentalidad.

Si bien ésta concepción ha sido ampliamente debatida, primeramente, se verá los orígenes de esta postura que surgió con Rene Descartes. De tal modo, se buscará entender la postura cartesiana, pero me enfocaré en explicar tanto la concepción cartesiana de pensamiento como la relación mente-cuerpo y además se enfatizará un singular planteamiento que menciona Descartes respecto a la idea de si existen o pueden existir otro tipo de seres con pensamiento. Puesto que la postura dualista cartesiana se inclina a decir que las facultades mentales son exclusivas de ciertos cuerpos, por consiguiente, se verán las razones que se tiene para hacer ésa consideración de que las facultades mentales solo son propias de ciertos cuerpos y por qué no podrían ser parte de otro tipo de cuerpos.

1.2 Dualismo cartesiano

Rene Descartes es uno de los filósofos más destacados y polémicos en filosofía de la mente, al postular una teoría dualista del pensamiento, él propone considerar la existencia de una sustancia pensante. La cual, se caracteriza por ser un fenómeno de naturaleza esencialmente no física (Churchland, 1999: p. 17), y suele tener diversas connotaciones como “mente”, “espíritu” o “alma” entre otras. En su dualismo, la contraparte de esta sustancia pensante son los fenómenos de naturaleza esencialmente física o corpórea. Aunado a esto, Descartes plantea que estos fenómenos de naturalezas opuestas son de hecho capaces de coexistir e

inclusive de interactuar entre sí en condiciones aparentemente exclusivas y privilegiadas. Esta postura se conoce como el problema mente-cuerpo de la filosofía, y ha incitado a un sinnúmero de pensadores a cuestionarse sobre el tema. Así como, a representado una gran influencia para el trabajo de muchos.

En distintas de sus obras, Descartes muestra gran interés por abordar temas epistemológicos, ontológicos y metafísicos. Entre ellos un punto focal en sus trabajos es su concepción del pensamiento humano, mediante el cual, él realiza el tratamiento y exposición de otros temas. Pues, independientemente del tema que esté tratando en sus trabajos, Descartes siempre muestra una relación entre su tema de interés y el pensamiento. Ya sea que buscará probar su propia existencia o la existencia de Dios, explicar o justificar la conducta humana, el funcionamiento del cuerpo, indagar sobre la verdad y falsedad o sobre nuestras capacidades de soñar, conocer, imaginar, sentir, etcétera. El pensamiento resulta ser un punto influyente dentro de sus investigaciones.

1.2.1 La substancia pensante

La primera cuestión a tener en cuenta respecto de su planteamiento es que para Descartes el pensamiento es un tipo de ente o substancia muy peculiar en tanto que su esencia se caracteriza por ser de una naturaleza no corpórea o inmaterial, pues advierte que a diferencia de lo material dicha substancia es indivisible, inmortal e imperceptible mediante las sensaciones corporales. Por otro lado, para entender la concepción cartesiana de pensamiento, a mi parecer, también es importante comprender el peso que implican ciertas ideas que el expresa en sus textos. Pues, cuando Descartes habla de pensamiento suele decir cosas como “es contradictorio creer que no existe aquello que piensa mientras piensa” (Descartes y Leibniz, 1989: p.31), “Soy una cosa que piensa, es decir, que duda, afirma, niega, conoce pocas cosas, ignora muchas otras, ama, odia, quiere, no quiere y también imagina y siente.” (Descartes, 2002: p.165). Este tipo de comentarios o más bien afirmaciones sobre el pensamiento pueden ser confusas y poco reveladoras, especialmente cuando apenas estamos conociendo o tratando los textos de Descartes.

Aquí, una de las cuestiones que se pretende abordar es que esta manera de presentar su idea del pensamiento es, de hecho, la manera en que lo concibe. Pues, para empezar, esta presentación de sus ideas tiene que ver en parte con la profunda insatisfacción que él expresa

en sus textos, como en la primera parte del *Discurso del Método*, respecto al método de las enseñanzas que recibió. A pesar de que estos comentarios parecen poco relevantes ciertamente llegan a tener gran peso dentro de su concepción, ya que él buscará apartarse de las enseñanzas que recibió. Prueba de ello se advierte en su texto *Investigación de la verdad por luz natura*l, cuando dice:

> No he tenido la intención, ni la tendré nunca, de vituperar el método ordinario de enseñar que se sigue en las escuelas… pues todo lo que enseñaron era tan dudoso que, si hubiera sido más conforme a la razón, quizá me habría contentado con la pizca de razón que hubiera obtenido, y eso me habría hecho más remiso para indagar cuidadosamente la verdad (Descartes, 2011: p. 87).

En ese mismo texto, Descartes aclara que ese "método ordinario de enseñar" es la manera en que se suele problematizar un tema al plantear preguntas como ¿qué es el pensamiento?, ¿qué es la existencia?, etcétera. Su repudio hacia ese método radica en que, a su parecer, estos planteamientos generan una especie de enajenación con temas que somos capaces de conocer o, de hecho, conocemos.

> Hay cosas que, cuando queremos definirlas, las hacemos más obscuras, porque, siendo muy simples y claras, no podemos saberlas ni percibirlas mejor que por sí mismas… Ahora bien, entre las cosas que son así de claras y se conocen por sí, pueden incluirse la duda, el pensamiento y la existencia… Añado incluso que es imposible que alguien aprenda estas cosas por otro medio que por sí mismo, ni que se convenza de ellas de otro modo que por su experiencia y conciencia o testimonio interno que cada cual experimenta en sí mismo cuando las considera. (Descartes, 2011: p. 93)

De tal forma, Descartes no pretende en ningún momento cuestionar o preguntar sobre el pensamiento, pues para él, el pensamiento es algo que no requiere ser explicado para ser conocido en tanto que es algo que se conoce mediante su experimentación o vivencia interna; así como mediante la capacidad que tiene el pensamiento de ser consciente de su propio acaecimiento o mediante su propia introspección. Por ello adopta un método de enseñanza en el que para hablar del pensamiento basta con afirmar cosas de él, como si fuera algo familiar o como si ya supiéramos de qué está hablando. Ahora bien, según Descartes la experiencia del pensamiento tiene diversos modos de acaecer, los cuales, también son cosas

que de hecho ya conocemos o nos resultan claras por experiencia propia y distintas por introspección o conciencia; estos modos se dividen en dos géneros del pensamiento.

Por un lado, está el género de las voliciones en el cual se manifiesta la capacidad que tiene el pensamiento para realizar operaciones de acción, es decir, de manifestar una cierta decisión, inclinación, afección o voluntad respecto de algo. Según Descartes, este tipo de experiencia y capacidad proviene directamente del propio pensamiento o bien la ocurrencia de estas voliciones es producto exclusivo y propio del alma, por ello, dirá que estas parecen depender sólo del pensamiento mismo. Este origen de las voliciones se puede advertir mediante la conciencia o introspección de dichas voliciones. Además, dicho género del pensamiento tiene la particularidad de ser considerado por Descartes el de mayor perfección o más bien de ser absolutamente perfecto, pues como él dice "de entre todas las cosas que hay en mí, no hay ninguna que sea tan perfecta y grande que no reconozca yo que podría muy bien ser aún más perfecta y grande" (Descartes, 2002: p.186).

Algunos modos de querer o tipos de voliciones son el desear, rechazar, temer, afirmar, dudar, negar, odiar, amar, etcétera; y podemos aplicarlos tanto a objetos materiales como a objetos no materiales, esto es, que nuestro pensamiento puede manifestar su voluntad, querer o no querer, etcétera tanto de objetos físicos como inmateriales como a una mesa o alguna idea. Estas dos aplicaciones de la voluntad Descartes las describe como acciones del alma que terminan en el alma misma o bien se aplican a objetos inmateriales, y acciones del alma que terminan en el cuerpo o se aplican a objetos físicos. Algo interesante y que se tratará más adelante es que, Descartes dirá que, las acciones que terminan en el cuerpo, de hecho, generan eventos físicos. Además, a diferencia de lo que veremos en el otro género de facultades del pensamiento, aquí al ser un género perfecto el cual consiste en la emisión de decisiones, no hay posibilidad de error. Pues la voluntad del pensamiento en sí misma no puede ser equívoca, es decir, el pensamiento no puede estar equivocado al manifestar su propia voluntad respecto de algo.

Por otro lado, está el género de las pasiones en el cual se manifiesta la capacidad del pensamiento para percibir, esto es su capacidad para conocer y entender algo, así como para adquirir y formular ideas o juicio respecto de ellos. A diferencia del genero anterior, este tipo de operaciones se encuentran en el pensamiento o provienen de él, pero tales percepciones

no son completamente propias del pensamiento o no son generadas únicamente mediante él. Más bien, aquí, el alma es capaz de recibir información de cosas que le son ajenas y a partir de ella puede formular en él ideas o representaciones de tales cosas, por ello Descartes dice que las pasiones "son como imágenes de las cosas y sólo a éstos les conviene propiamente el nombre de idea: como cuando me represento un hombre, una quimera, el cielo, un ángel o el mismo Dios" (Descartes, 2002: p.167).

Así, los modos de percibir son el imaginar, entender, conocer y entre ellos también se encuentra el sentir o la conciencia. Los modos de percibir también son de dos tipos, los que tienen por causa el alma y los que tienen por causa el cuerpo. Los que tienen por causa el alma se refieren a la capacidad o facultad del pensamiento para percibir, advertir, enjuiciar e incluso modificar sus propias voliciones, intelecciones, imaginaciones y demás modos de pensar. Cabe decir que la conciencia o sensación en Descartes es precisamente una pasión del alma que tiene por causa el alma misma, es decir: es una percepción respecto de las ocurrencias o eventos que se generan en el pensamiento o que tienen lugar en él (Descartes, 1989: p.32). En cambio, las pasiones que tienen por causa el cuerpo son las percepciones, ideas o impresiones que se generan directamente por influencia de objetos ajenos al pensamiento, las cuales llegan al pensamiento solo mediante el cuerpo al cual está estrechamente unido el alma.

El género de las pasiones, además, resulta no ser perfecto como el de las voliciones ya que su alcance de conocimiento es limitado, por ello en éste siempre se puede aspirar a un conocimiento más grande y "nunca llegará a tal punto de perfección, que no pueda acrecentarse más" (Descartes, 2002: p. 177). Esta imperfección, según Descartes, genera que el pensamiento pueda tener errores, pues la sola percepción o acaecimiento de las ideas que se generan en el pensamiento no pueden ser como tal un error del mismo, en tanto que "las ideas son en mí como cuadros o imágenes que pueden, es cierto, descender fácilmente de la perfección de las cosas de donde han sido sacados" (Descartes, 2002: p. 172). El error del pensamiento surge de la interacción entre ambos géneros del pensamiento, esto es, cuando la voluntad se adelanta en la formulación de una percepción y afirma o niega algo respecto de dicho conocimiento.

Entonces, el pensamiento tiene la particularidad de ser un tipo de experimentación o vivencia no corpórea a partir de la cual se tiene la capacidad de decidir libremente sobre sus acciones y/o voluntades, así como permite generar un entendimiento o conocimiento respecto de todas sus experiencias. Ahora bien, como se dijo, parte de estas experiencias del pensamiento son provocadas mediante el cuerpo, y en Descartes dicho cuerpo aparentemente no puede ser cualquier cuerpo.

1.2.2 La substancia material o corpórea

A parte de la substancia pensante Descartes reconoce la existencia de una segunda substancia que es la materia, a la cual también se refiere como cosas u objetos físicos, corpóreos o materiales. Esta substancia la describe como algo extenso, divisible, perceptible mediante los sentidos, mortal e inanimada. Respecto a esta última característica cabe resaltar que en Descartes la diferencia entre cuerpos animados e inanimados no radica solo en que los primeros sean en sí mismos capaces de presentar algún tipo de movimiento mientras que los segundos carecerían por entero del mismo, pues para él hay cuerpos inanimados que son plenamente capaces de mostrar tantos movimientos y operaciones como lo harían los cuerpos animados. “Y así se ha creído, sin razón, que nuestro calor natural y todos los movimientos de nuestros cuerpos dependen del alma: mientras que se debía pensar, al contrario, que el alma sólo se ausenta cuando uno muere por que ese calor cesa y los órganos que sirven para mover el cuerpo se corrompen” (Descartes, 1997: p.61). La diferencia radicará en que los cuerpos animados aparentemente tienen pensamiento. Por tanto, las habilidades que éstos muestran supuestamente operan y funcionan de un modo distinto al proceder de las habilidades de los cuerpos inanimados.

Dentro del ámbito de las cosas físicas Descartes plantea que hay cosas más perfectas que otras o que contienen más realidad que otras. Por ejemplo, el hecho de que hay cuerpos que no presentan movimiento, como las piedras, muestra que tales cuerpos poseen una menor perfección o realidad respecto de aquellos que sí tienen la disposición para moverse, como ocurre con los animales “Pues, en efecto, las que me representan sustancias son sin duda algo más y contienen, por decirlo así, más realidad objetiva, es decir, participan por representación a más grados de ser o perfección que las que sólo me representan modos o accidentes” (Descartes, 2002: p. 170). Asimismo, un cuerpo animado presentara mayor perfección que

un cuerpo inanimado. Esta jerarquización cartesiana de la perfección, en su extremo de lo menos perfecto, comienza catalogando objetos físicos que poseen menor complejidad y capacidad de disposición en sus cuerpos, es decir: que posean una constitución o composición física poco compleja, así como una precaria diversidad de acciones (destrezas), como el que generen calor, el que sean capaces de respirar, que tengan órganos, etcétera. En tanto que en su extremo de mayor perfección tenderá hacia aquellas cosas que él identifica que poseen atributos de mayor abstracción respecto de los objetos físicos, es decir verá como más perfectas aquellas cosas a las que se les pueden atribuir habilidades (funcionamientos) y propiedades (una constitución o esencia) que no sean propiamente físicas sino, más bien, mentales.

En la quinta parte del *Discurso del método* Descartes reconoce tres tipos de cuerpos que presentan un alto grado de perfección, a saber, el cuerpo humano, los cuerpos de los animales y las máquinas o autómatas. Los cuales destacan por la diversidad y complejidad de disposición y propiedades que muestran en sus cuerpos, ya que sus constituciones físicas están compuestas de diversos materiales, órganos y mecanismos que a su vez se encuentran organizados de tal manera que les permiten llevar a cabo una o varias acciones. Si bien tales constituciones físicas no son plenamente iguales entre sí y no tienen exactamente las mismas disposiciones de sus órganos o mecanismos y destrezas. Aun así, entre ellas llegan a poder producir acciones o cosas muy similares como desplazarse de un lugar a otro, emitir ciertos sonidos, comportarse de cierta manera, manipular o hacer uso de otros objetos, e incluso en algunos casos pueden emitir palabras.

Descartes buscará determinar cuál de ellos es el de mayor perfección diciendo, por un lado, que las semejanzas de los cuerpos respecto a su capacidad de acción se pueden distinguir desde el punto de vista fisiológico, es decir, por los tipos de mecanismos y materiales con los que operan los cuerpos según su constitución y mediante los cuales se llevan a cabo sus acciones. Por ejemplo, pensemos en la acción de lanzar un objeto, dentro del ámbito meramente fisiológico o mecánico éste hecho implica el imprimir cierta fuerza a un objeto, haciendo que cambie abruptamente del lugar donde estaba hacia otro más alejado. Esta acción será la misma sin importar que cuerpo la realice, empero la distinción radica en que en el caso de los humanos y los animales esto se lleva a cabo mediante huesos, músculos,

nervios, y demás tejidos biológicos que en ambos casos por su naturaleza operan de la misma manera; mientras que, en los autómatas esta acción se realiza mediante engranes, resortes, placas de metal, cables, y demás mecanismos o circuitos, los cuales por su naturaleza no operan de la misma forma que los organismos biológicos. Además, Descartes va a considerar como más perfectas aquellas cosas que estén hechas por Dios, por lo cual una máquina siempre será menos perfecta por ser obra del humano. De esta manera, para Descartes las máquinas ya serían los menos perfectos dentro de estos tres tipos de cuerpos. No obstante, habrá otra razón filosóficamente más interesante para desacreditarlos.

> El cuerpo de un animal, consideren este cuerpo como una máquina que, por ser hecha de manos de Dios, está incomparablemente mejor ordenada y posee movimientos más admirables que ninguna otra de las que puedan inventar los hombres. Y aquí me extendí particularmente, haciendo ver que, si hubiese máquinas tales que tuvieran los órganos y figura exterior de un mono o de otro cualquiera animal, desprovisto de razón, no habría medio alguno que nos permitiera conocer que no son en todo de igual naturaleza que esos animales (Descartes, 2002: p.113-114).

Por otro lado, aunque se puedan distinguir las acciones de estos cuerpos por el grado de dominio que muestran sobre ellas, éstas diferencias de acción para Descartes en realidad no permiten determinar una diferencia en la perfección de los cuerpos, pues el grado de eficacia y destreza que se muestre solo se debe a las características particulares de sus disposiciones fisiológicas que, por su naturaleza y organización, les permiten mostrar mayor dominio sobre ciertas actividades de la que se lograría con otros materiales y otra organización de los mismos. Por ejemplo, volar es una actividad que las aves dominan con mucha más destreza y eficacia que los humanos y máquinas; en cambio calcular es una acción que las máquinas tienden a dominar con mucho más eficacia y destreza que los animales y humanos; mientras que el uso de herramientas es una actividad mucho más desarrollada en humanos que en animales y máquinas.

> Es también muy notable cosa que, aun cuando hay varios animales que demuestran más industria que nosotros en algunas de sus acciones, sin embargo, vemos que esos mismos no demuestran ninguna en muchas otras; de suerte que esto que hacen mejor que nosotros no prueba que tengan ingenio… sino más bien prueba que no tienen ninguno y que es la

> naturaleza la que en ellos obra, por la disposición de sus órganos (Descartes, 2002: p. 115-116).

Así, a parte de una diferencia meramente fisiológica y de domino sobre la realización de tal o cual acción él advertirá que, a diferencia de los animales y las máquinas, los humanos presentan un modo diferente de realizar acciones. El cuál será la diferencia indudable de que hay mayor perfección en tales acciones pues, dirá que, ese modo es propio del pensamiento o alma.

1.2.3 Interacción cartesiana entre mente-cuerpo

Las acciones que presentan los humanos, según Descartes, tienen la particularidad de ser realizadas por conocimiento esto es que el humano tiene la capacidad de entender, juzgar, querer y de no querer respecto de sus acciones. Por lo cual, en sus acciones físicas también se pueden advertir operaciones características del alma, y, Descartes dirá que, esto se advierte más claramente al momento de hacer uso de palabras, pero más concretamente al mostrar un domino en el uso del lenguaje. Pues, aunque el dominio de una actividad, como se dijo, no es garantía de mayor perfección; aquí Descartes parece decir que ésta capacidad va a tener algo así como una doble perspectiva de domino, es decir, hay dos modos en que se puede mostrar un mejor manejo de tal o cual actividad o, al menos, es en esta acción en particular donde podemos advertir con mayor claridad la posibilidad de que haya dos maneras de mostrar dominio respecto de una misma acción física. Cabe aclarar que, aquí, empleo el término "acción física" para destacar la idea de que hablo de fenómenos o hechos del ámbito corpóreo, es decir, me refiero a cualquier actividad que tenga una realización física.

Si bien como acción física, para Descartes y su época, el humano es quien muestra un mayor dominio sobre el uso de palabras o símbolos respecto de los animales y máquinas, pues éste es capaz de proferir y utilizar una abrumadora cantidad de palabras o símbolos. No obstante, vista como una ejecución de eventos físicos dicho dominio sigue sin ser prueba de mayor perfección, ya que podría ser una ejecución producto de la disposición de los órganos del cuerpo humano, así como lo es con los animales y las máquinas. Empero, para Descartes, el humano a parte de mostrar un dominio meramente fisiológico sobre sus palabras, también muestra un dominio mental sobre ellas. Pues mediante su uso éste es capaz de transmitir ideas (conocimientos, imaginaciones, percepciones, entendimientos, juicios, etc.) y voliciones

(deseos, quereres, disgustos, alegrías, tristezas, miedos, etc.) que le permiten comunicar sus pensamientos. Es aquí cuando la acción (física) de usar palabras se torna una función, pues se vuelve una acción mental o bien una actividad que tiene una realización en el pensamiento, y, además, pasa a ser considerado propiamente como un uso y dominio de un leguaje en lugar de un mero uso de palabras.

Así para Descartes es mediante el uso de palabras (el lenguaje) y el obrar por conocimiento (pensamiento) que se puede distinguir, indudablemente, al cuerpo humano como el objeto físico de mayor perfección.

> Por esos medios puede conocerse también la diferencia que hay entre los hombres y los brutos, pues es cosa muy de notar que no hay hombre, por estúpido y embobado que éste, sin exceptuar los locos, que no sea capaz de arreglar un conjunto de varias palabras y componer un discurso que dé a entender sus pensamientos; y, por el contrario, no hay animal, por perfecto y felizmente dotado que sea, que pueda hacer otro tanto (Descartes, 2002: p. 114-115).

Ahora bien, las otras actividades que realiza el cuerpo humano seguirán siendo acciones físicas y solo algunas de ellas tendrán también una realización en el pensamiento, como resolver un problema matemático, armar un rompecabezas o cualquier actividad que pudiera implicar algún tipo de sentido o pasos no visibles en la acción física. Sin embargo, para Descartes estas acciones, a diferencia del uso de palabras, no pueden ser corroboradas de manera indudable como acciones mentales que se realizan por conocimiento; aun así, aceptará que tales actividades pueden pasar a ser entendidas propiamente como funciones (Descartes, 2011b: p.710). Es decir, en el ejemplo de lanzar un objeto, se dijo que la acción física, como tal, del cuerpo es imprimir fuerza a un objeto. Pero, lo que ocurra después con ese objeto, el que cambie abruptamente de un lugar a otro, sería un evento distinto o propio de ese otro cuerpo. Empero, como acción mental estos dos eventos tienden a verse como parte de un mismo sistema, donde la acción del primero es causa directa del actuar del segundo cuerpo en tanto que el primer cuerpo, por su capacidad de conocimiento o entendimiento, es capaz de contemplar y calcular la reacción del otro cuerpo respecto del estímulo que le va o le está propiciando. En este sentido se dice que el primer cuerpo es capaz de manifestar mediante su acción física y la reacción del otro cuerpo un cierto pensamiento o bien un cierto propósito en sus acciones.

Lo siguiente es ver cómo ocurre esto según Descartes, pues si bien el cuerpo humano es un complejo sistema biológico, él dirá que hay ciertos órganos mediante los cuales opera la mente. El primero es la glándula pineal en la cual reside el alma y donde ocurren todas las operaciones de la misma, de ahí procederán todas las disposiciones del alma, a saber, las funciones manifestadas mediante acciones físicas. Después, a grandes rasgos, están los órganos de los sentidos (ojos, oídos, piel, etc.), los músculos, el cerebro, el sistema nervioso y el sistema vascular (las venas) en el cual además "lo que llamo aquí espíritus animales son solamente cuerpos, y no tiene más propiedad que la de ser cuerpos muy pequeños" (Descartes,1997: p.72).

Los espíritus animales como se dijo son diminutas partículas que se generan en el corazón y circulan por el sistema vascular hacia el cerebro y la glándula pineal, también son capaces de circular por el sistema nervioso y así actuar sobre los músculos. Además, actúan como transductores de información entre la glándula pineal y el resto del cuerpo humano. De igual manera, los órganos de los sentidos actúan como transductores de información entre el mundo externo y el cerebro. Así cuando los órganos de los sentidos reciben información, ésta se transmite mediante el sistema nervioso hacia el cerebro. Ahí, los espíritus animales logran captar la información y la hacen llegar a la gandula pineal, de donde posteriormente volverán a salir hacia el cerebro, los nervios y músculos llevando a cabo las disposiciones del alma (Descarte,1997: p. 71-72; Descartes, 2011b: 715-718).

Un último aspecto a tratar sobre la interacción mente cuerpo es la memoria pues Descartes dirá que, ésta, propiamente es parte del cerebro por lo cual es algo corpóreo. La información que es almacenada en ella se da porque los espíritus animales trazan la información en el cerebro a medida que lo recorren para trasmitir información entre la glándula pineal y el cuerpo, de esta manera el cerebro almacena tanto información del mundo exterior como la información proveniente de la glándula pineal. Además, es mediante esa información almacenada en el cerebro que la mente puede realizar la operación de imaginar, puesto que hace uso de dicha información para percibir cosas que no se están presentando como hechos ante los órganos de los sentidos (Descartes, 2011b: p. 719).

1.3 ¿Es eso el pensamiento?

Algunos filósofos como Ryle Gilbert consideran que no es tan evidente e indudable la distinción entre los eventos físicos que son producto de acciones físicas y aquellos que son producto de acciones mentales, pues consideran que la idea cartesiana respecto a que las mentes supuestamente imprimen funciones a eventos físicos está siendo malentendida por el propio Descartes. De tal manera Ryle dirá que la distinción cartesiana mente-cuerpo en realidad es un error categorial, esto es que Descartes está atribuyendo a la mente una categoría lógica que no le pertenece. Para entender a qué se refiere Ryle con "error categorial" él nos ofrece varios ejemplos en su libro *El concepto de lo mental*. No obstante, aquí he optado por ofrecer un ejemplo de error categorial propio que considero expone adecuadamente la idea sobre error categorial de Ryle a la vez que introduce varias de las observaciones que él advierte respecto a la postura cartesiana y, asimismo, da una perspectiva hipotética pero común de cómo se pueden generar estos errores, lo cual considero ayuda a entender mejor este punto al proporcionar un panorama un poco más detallado y familiar.

1.3.1 Error categorial

Piense en lo que implica realizar un paltering, esto es, lo que comúnmente se conoce como *mentir diciendo la verdad*. Imaginemos, entonces, el caso en que un profesor al entrar al salón de clases profiere un acto locucionario y pregunta a sus alumnos si hicieron la tarea. Supongamos que la tarea de hecho consistía en resolver varios ejercicios. Todos los alumnos al unísono contestan al profesor, mediante otro acto locucionario, que sí hicieron la tarea. Lo que el profesor, de hecho, no sabe, y no corresponde propiamente al acto ilocucionarion o significativo de su pregunta, es que algunos de sus alumnos audazmente logran realizar o copiar todas las respuestas de los ejercicios antes de que él entrará al salón y otros sólo lograron finalizar unos cuantos ejercicios. Sin embargo, la respuesta que los alumnos dieron a su profesor en cierto modo responde verídicamente a su pregunta, a su acto locucionario, en tanto que los términos 'hicieron la tarea', como tal o en sí mismos, pueden no implicar el no hacer la tarea poco antes de ser impartida la clase e incluso en algún sentido podría no implicar el haberla terminado. Esto último podría ocurrir cuando el alumno admite que, al dejar varios ejercicios, el profesor en realidad dejó muchas tareas, y que con el término 'la

tarea' el profesor se refiere solo a uno de los ejercicios que les dejó o bien podrían estar admitiendo que terminaron una tarea, aunque esa tarea no es de ese profesor.

En el ejemplo, por mor del contexto o costumbre, en el acto locucionario del profesor se comete una falta de proferencia y/o especificación significativa respecto de su acto ilocucionario. Dando, así, la posibilidad a sus alumnos de tomarse la libertad para asignar, a conveniencia, un contenido significativo al acto locucionario emitido por su profesor que no corresponder con el acto ilocusionario del mismo. De tal forma, encontramos que los alumnos astutamente consiguen adoptar un acto ilocucionario distinto al de su profesor que, sin embargo, les permite responder con un acto locucionario tan abstracto, similar o aparentemente equivalente al de su profesor, ya que para el profesor tal proferencia consigue responder con aparente coherencia a su acto ilocucionario. Así, tenemos dos tipos de actos locucionarios (alumnos/profesor), y ambos logran ser válidos dentro de dos actos ilocucionarios o significativos, a saber, el de los alumnos y el del profesor, debido a que omiten o esconden su significado específico.

Pensemos ahora ¿Qué sería un "falso paltering"? Supongamos que, de hecho, los alumnos no advierten la diferencia entre su acto ilocucionario y el acto ilocucionario del profesor, es decir, para ellos su acto ilocucionario corresponde al de su profesor, por lo cual, para ellos no debería haber diferencias significativas entre el tipo de acto locucionario que profieren y el de su profesor. Además, al no ver como incorrecto, indeseable o impertinente su respuesta, los alumnos no están cometiendo paltering. Simplemente cometen o caen en un error. Pues, tal vez a estos chicos nunca se les explicó que las tareas son un ejercicio para ayudar a repasar y asentar conocimientos, y que, idealmente, una buena sesión de repaso involucra el que uno mismo realice los ejercicios, así como contempla un plazo de descanso que permita asentar los conocimientos o destrezas practicadas antes de volver a ser repasadas; o nunca se les dijo que cuando el profesor te deja de tarea resolver distintos ejercicios sobre los temas abordados en una clase, esos ejercicios en conjunto, se toman como una sola tarea correspondiente a la sesión impartida; o bien, no advierten que si un profesor de cierta clase te pregunta por la tarea, en realidad te está preguntando por la tarea que él te dejo y no por tus otras tareas.

Este "falso paltering" o confusión de los términos y de la lógica de los mismos es lo que Ryle identifica como error categorial, aunque, el paltering por si solo también es un tipo de error

categorial o bien una conveniencia categorial a nivel individual o una adecuación categorial arbitraria; esto ocurre en tanto que la modificación de los términos (palabras) y su lógica (connotación) se realiza deliberadamente por un individuo. Ahora bien, la persona que genera tanto el error categorial como la adecuación categorial arbitraria está otorgando una categoría lógica (connotación o clasificación), a su respuesta o creencia (palabras o términos o locución), que no corresponde al sistema lógico (contexto connotativo o ilocución) que se está tratando.

Dicho de otra manera, esto es que un cierto individuo a partir de un conjunto de clasificaciones o de un determinado contexto connotativo (conjunto de categorías o ilocución) de conceptos, es capaz, mediante la generalidad o abstracción que permiten los términos connotativos en sí mismos, de reasignar o modificar la clasificación (categoría) de los conceptos o términos involucrados; generando otro contexto connotativo (sistema lógico) en el que dichos términos también resultan válidos. Pues, cabe destacar que aquí no se modifica propiamente la estructura lógica o la manera en que operan tanto los sistemas lógicos (contexto connotativo) como los términos que utiliza (el lenguaje), simplemente uno de los sistemas (el del alumno) genera nuevos valores equivalentes o compatibles con el modelo operativo del lenguaje utilizado por el otro sistema (el del profesor). Empero, lo que Ryle intenta decir es que, esto no significa que los conceptos que han adquirido nuevas connotaciones puedan aplicarse indistintamente dentro del sistema lógico previo a tales modificaciones. Sino que se debe reconocer que son sistemas lógicos distintos, en tanto que poseen connotaciones diferentes, por lo cual no hay una correspondencia lógica entre las nuevas connotaciones (aunque pudieran ser los mismos términos) y el sistema lógico original.

> Los errores categoriales teóricamente interesantes son los cometidos por personas capases de usar conceptos, por lo menos en situaciones que les son familiares, pero que, sin embargo, al pensar en abstracto pueden asignar dichos conceptos a tipos lógicos distintos de aquellos a los que pertenecen (Ryle, 2005: p. 31)

1.3.2 Otros aspectos sobre el ejemplo anterior

Algo interesante del error categorial y la adecuación categorial arbitraria es que, si se le analiza en retrospectiva se puede advertir que su sola ocurrencia denota necesariamente la interacción de al menos dos sistemas lógicos con semánticas formales propias y un lenguaje formal común a tales sistemas o dos sistemas ilocucionarios distintos (alumno/profesor) y un lenguaje común; esto último es la proferencia de los actos locucionarios, el uso de las palabras o términos en la interacción. De tal manera, estos casos son entendidos como interacciones semánticas o connotativas entre dos sistemas lógicos o dos mentalidades. Donde, empero, uno de los dos sistemas es tomado como el sistema base o veritativo de la interacción, en estos casos fue la ilocucion del profesor. Mientras que el otro sistema parece generarse con base en ese primer sistema o al menos surge a partir del acto locucionario del primer sistema y aparentemente con el fin de darle explicación al mismo.

Ahora bien, en el caso de la adecuación categorial arbitraria (en el paltering) se presume que el alumno conoce o entiende plenamente la semántica formal que el profesor está aplicando al leguaje formal que ambos utilizan. No obstante, el alumno arbitrariamente rechaza como veritativa esa aplicación semántica de los términos y genera su propia versión semántica para los mismos sin perturbar la operación del lenguaje formal. Por otro lado, en el error categorial se está diciendo que el alumno no tiene un conocimiento o entendimiento suficiente respecto de la semántica formal del profesor, pero es capaz de generar su propia semántica para los términos utilizados. Como vemos lo que da paso a tales interacciones entre ambos sistemas es la manera indirecta o abstracta de comunicarse mediante un lenguaje formal, en el cual ambos sistemas tienen la posibilidad de manipular las connotaciones del mismo según su propio funcionamiento o sus propias reglas, axiomas y criterios de operación.

En retrospectiva, también, se podría decir que tanto la ocurrencia de los errores categoriales y el paltering así como su dificultad para detectarlos en la vida cotidiana podría ser indicio de esa comunicación indirecta o abstracta entre las mentes, ya que es muestra de la imposibilidad de los individuos para acceder directamente a los estados mentales del otro. Pues, siguiendo el ejemplo del profesor y sus alumnos, normalmente en estas situaciones el profesor muy probablemente llega a dudar de la respuesta de sus alumnos, pero, automáticamente desecha su duda o simplemente lo deja pasar. ¿Por qué ocurre esto? En

primera instancia, se podría decir que la situación resulta tan común, normal o cotidiana para el profesor que simplemente por costumbre o hábito lo deja pasar. Pero, en realidad, a simple vista no hay ninguna señal tangible, para el profesor, que le permita detectar una diferencia significativa como para advertir la mentira o el error de sus alumnos. Ciertamente se puede dar la sospecha al respecto, pero no hay nada concluyente que permita o dé razones para seguir sosteniendo la duda. Esta incapacidad del profesor para sostener la sospecha se debe, en parte, al desconocimiento de lo que implica la respuesta de sus alumnos dado que ese contexto de implicaciones es determinado por el acto ilocucionaro o la mentalidad de los alumnos que la emiten. Pero, también se debe a que él está asumiendo y delimitando el contexto de esas respuestas con su propio acto ilocucionario o sus propias categorías lógicas. De esta manera, este tipo de errores parecen tener lugar en esa incapacidad de corroborar la mentalidad de la otra persona o bien en la posibilidad de generar un sistema lógico paralelo a partir de otro ya establecido, aunque probablemente Ryle preferiría decir que se debe a un cierto desconocimiento, ignorancia, falta de práctica o familiaridad.

1.3.3 Primera persona

Respecto a la idea de que la distinción cartesiana mente-cuerpo es un error categorial, Ryle y otros autores como Bernard Williams advierten como una señal clara de esto el que la dualidad señalada por Descartes ocurre a causa de la manera en que él expone y concibe a la mente humana. Pues, como se mencionó al principio del presente capítulo, Descartes concibe al pensamiento como distinto del cuerpo y como algo indudable que conocemos por conciencia, introspección o si se prefiere por la experimentación propia y directa de nuestras acciones mentales. Tanto Williams como Ryle dirán que éste modo de abordar el pensamiento promueve que adoptemos una postura en la cual las diferentes percepciones, voliciones, ideas e incluso experiencias sensibles son vistas como parte de un acaecimiento inmaterial que conforma el flujo del pensamiento, dado que éstas llegan a ser aprehensiones y operaciones mentales. Esta postura, además, contempla la idea de que todas las operaciones propias de una mente o todo el flujo del pensamiento inevitablemente se nos presenta o se presenta ante la conciencia de dicha mente. Por tanto, se está diciendo que las mentes tendrían que poder advertir todas las implicaciones o circunstancias mentales que rodean la ocurrencia de una cierta operación mental, sean cuales fueren (Williams, 2012: p. 103-104). Esto es que

cada individuo dotado de pensamiento justifica, explica, sabe y entiende por qué y de qué manera se da cada una de las acciones mentales dentro de su flujo de pensamiento.

Ahora bien, esta concepción cartesiana del pensamiento como conciencia también se conoce como la perspectiva de primera persona, pues se está contemplando al pensamiento como un tipo de experiencia propia y exclusiva de un individuo. Por ello, Descartes siempre describe al pensamiento en términos de acciones que uno mismo realiza o narrándolas desde el punto de vista de quien lo experimenta, y dónde el propio individuo asume que es consciente de la acción que realiza o bien sabe lo que está realizando. Para Ryle, es a partir de ésa perspectiva de que somos conscientes de nuestro flujo de pensamiento donde, también, se genera la idea de que la mente imprime una función o sentido a los eventos físicos que esta produce. Pues, entiendo que para Ryle la idea de una conciencia y conocimiento de los estados mentales sería la condición preliminar de los humanos para que se de lo que Descartes llama "obrar por conocimiento", en la cual plantea que el humano es capaz transmitir sus pensamientos mediante acciones físicas.

Descartes estaría planteando, entonces, que las acciones o eventos físicos que produce una mentalidad son acciones que se han dado previamente en el pensamiento o bien son acciones de las que es consciente el individuo que las realiza. De esta manera, las acciones físicas humanas son realizadas conforme a los estados mentales que acaecen en el flujo de pensamiento de tal o cual individuo o bien se hablaría de que tal acción se da al obrar por conocimiento. Esta conformidad de las acciones físicas respecto de las acciones mentales es lo que les otorgaría a las primeras la capacidad de adquirir una cierta función, sentido o significado que es propia a la mentalidad del individuo que las realiza y de las cuales carecen las acciones físicas que son producto de meros objetos corpóreos, pues se estaría contemplando que las mentes y los cuerpos operan bajo sus propias condiciones, criterios o reglas de acción. En el caso de los cuerpos, según Descartes, estos operan bajo disposiciones meramente mecánicas o bien por las leyes naturales; las cuales determinan cómo reacciona un cuerpo ante ciertos estímulos según la manera y los materiales con que está constituido el cuerpo en cuestión. Mientras que las mentes serian algo así como cuerpos inmateriales (almas) que operarían bajo sus propias condiciones, reglas, leyes y criterios mentales preestablecidos.

Por ejemplo, si una máquina realizara una cierta acción física, como lanzar un objeto, Descartes diría que esto es solo producto de la mera disposición de sus mecanismos, cuyo funcionamiento depende de su inmersión y/o seguimiento de las leyes naturales; pues, para que la máquina logre realizar tal acción se requiere que sus mecanismos operen bajo ciertos criterios de fuerza, fricción, resistencia, peso, etcétera. Por otro lado, para que una mente lleve a cabo esa misma acción fisca previamente tendría que haber presentado acciones mentales, inmateriales y privadas, que determinaran cómo y porque se realizará dicha acción física. De tal manera, su realización estaría sujeta a las especificaciones y seguimiento de los estados mentales que se generen en el pensamiento del individuo que la lleva a cabo, es decir, su acción estaría determinada por las diversas voliciones que motiven la realización del evento, así como al entendimiento, planificación y cálculos mentales implicados en su realización. Teniendo, así, de antemano presente los requerimientos o pasos a seguir para su realización física; lo cual, contempla que su operación mental involucraría el conocimiento de las acciones corpóreas requeridas en su realización, así como las causas o efectos de las mismas.

En este sentido, si la mente tiene aprehensiones del mundo exterior, a su vez que es consciente de todas sus operaciones (acciones y percepciones) mentales y puede influir causalmente sobre objetos corpóreos, entonces, se puede plantear que todo evento físico producto de una mentalidad ha sido previamente planificado en y por la propia mente. Además, Ryle está diciendo que desde esta perspectiva de primera persona se considera que los fenómenos causados por una mente son producto de la aplicación de criterios o reglas preestablecidas por la misma.

> Argumentando que toda actuación inteligente presupone la observancia de reglas o la aplicación de criterios. Se sigue de ello que toda operación considerada inteligente debe ser precedida por un reconocimiento intelectual de esas reglas o criterios, esto es, que el sujeto debe primero reconocer internamente ciertas proposiciones acerca de lo que debe hacer… para comportarse —sólo después— de conformidad con tales dictados (Ryle, 2005: p. 43).

Por otro lado, Descartes está admitiendo que las mentes tienen la capacidad para generar eventos dentro del ámbito corpóreo lo cual promueve la idea de que, en cierto modo, tales mentes son algo equivalente a los cuerpos mecánicos. Pues, ambos son capaces de generar eventos físicos, es decir, tanto las acciones mecánicas como mentales parecen tener un efecto

causal sobre objetos o fenómenos físicos. Esta causalidad compartida de mentes y cuerpos da paso a la consideración de que ambos operan al mismo nivel lógico, a saber, dentro de la causalidad de los sucesos físicos propios del ámbito corpóreo. Consiguientemente, si bien se dijo que la mente no es en sí misma parte del ámbito físico, se está diciendo que es capaz de influir sobre los objetos físicos o generar fenómenos dentro del espacio, al igual que cualquier otro cuerpo. Me permito hacer énfasis sobre esta última frase "al igual que cualquier otro cuerpo" puesto que me parece una forma muy burda pero atinada para destacar lo que Ryle quiere expresar cuando dice que la perspectiva dualista cartesiana entiende a la mente como algo que pertenece a la misma categoría lógica que los cuerpos.

Lo que se está queriendo mostrar en este punto es que Descartes concibe a la mente como siendo el equivalente inmaterial de los cuerpos físicos, y dicha equivalencia parece radicar precisamente en su capacidad de tener causalidad sobre sucesos físicos. Cabe mencionar que dicha equivalencia también se puede observar en el hecho de que la mente igualmente tiene criterios o reglas para la elaboración de fenómenos físicos del tipo leyes naturales, es decir, las leyes y criterios mentales operan de manera similar a las leyes naturales ya que determinan el desarrollo causal de los fenómenos físicos. Aunque, en el caso específico de la mente sus criterios parecen ser causa directa en la generación de los fenómenos y, por ende, de la interacción propia de los cuerpos, mientras que en los cuerpos la causa directa más bien seria la interacción casual entre los cuerpos, la cual se ve determinada por las leyes naturales.

1.3.4 Tercera persona

Ahora bien, se habló de una primera persona (una mente) que aparte de ser consiente de todas sus acciones y ser capaz de generar eventos dentro del ámbito corpóreo, también es independiente del cuerpo e inmaterial. Al respecto, Ryle destaca que con esto se está advirtiendo que la ocurrencia de la mente, así como las operaciones que ésta involucra, es privada; pues esas características apuntan a una existencia completamente distinta y ajena a lo corpóreo. Esto supone la falta de la ocurrencia mental dentro del ámbito público y propio al de los objetos físicos, a saber, el espacio; lugar en el que interactúan los objetos físicos. Tal condición de las mentes cartesianas sugiere que algo como un cierto evento mental ocurre dentro del ámbito privado de la mente que lo experimenta y solo ésta misma puede ser consiente de dicho evento. En consecuencia, dice Ryle, "no existe conexión causal directa

entre lo que sucede a una mente y lo que le pasa a otra… La mente de una persona solo puede afectar a la mente de otra únicamente a través del mundo físico" (Ryle, 2005: p. 27).

Se está diciendo, entonces, que las mentes no tienen un lugar común en el que interactúen de manera indirecta entre sí, por lo cual, hay una imposibilidad de comunicación mente-mente. Empero está privacidad de las mentes también involucra a todo su flujo de pensamiento, es decir, todas sus operaciones mentales y, por ende, a los criterios o reglas que estas aplican y contemplan previamente a la realización de una acción física. Pues a diferencia de los fenómenos físicos aquí no se puede estudiar o advertir objetivamente los eventos de una mentalidad como para calcular o determinar los factores involucrados que motivaron la realización de una determinada acción (Williams, 2012: p. 126-127). Los únicos factores o reglas que se podrían advertir, estudiar o tomar en cuenta de forma objetiva son los apreciables por distintos individuos, a saber, los propios al ámbito corpóreo o bien los propios al fenómeno resultante de los cálculos mentales.

El problema respecto a esta incomunicación de las mentes es que imposibilita cerciorarse plenamente si es correcto o verídico cualquier significado, contenido mental o función que se pueda atribuir, y que supuestamente una mente puede imprimir al uso de palabras y acciones que otra realiza. Lo que es más, esta consecuencia impide la posibilidad de corroborar plenamente la existencia de otras mentes, dado que no hay un acceso directo a sus operaciones, y solo se podría inferir su existencia mediante acciones físicas supuestamente generadas por tales mentalidades. En este punto, Ryle advierte una clara dificultad en los criterios que Descartes propone para distinguir fenómenos producto de una mentalidad y aquellos que son producto de acciones mecánicas.

Pues criterios como "obrar por conocimiento" no parecen ser viables en la distinción de fenómenos mentales y fenómenos mecánicos en cuanto que no pueden ser corroborados plenamente ni siquiera al utilizar el lenguaje. Tómese en cuenta, por ejemplo, que bien puede darse el caso en que distintos cuerpos llevan a cabo acciones que generan un mismo efecto bajo circunstancias similares. Ya sea una máquina, un animal o un humano podrían lanzar una botella de tal manera que caiga en un bote de basura o bien podrían proferir una misma serie de palabras que respondan coherentemente con otro evento lingüístico. En cualquier caso, a excepción del tipo de cuerpo que lanzó la botella o profirió las palabras, no se da una

diferencia observable en el comportamiento y las acciones de estos tres cuerpos; los tres eventos presentan una secuencia de desarrollo idéntica. Pero, para Descartes, solo el actuar humano se calificaría como inteligente o propio de una conciencia.

Tanto Ryle como Williams señalan entonces que la postura cartesiana solo tiene sentido desde la primera persona, pero desde el punto de vista de quien atestigua tales sucesos o las acciones que esos cuerpos realizan, es decir, desde la tercera persona o el observador no se puede corroborar objetivamente la supuesta distinción entre el actuar humano, el del animal y el de la máquina. Ahí no existe tal diferencia, a lo mucho la supuesta función mental impresa en tales sucesos solo puede ser inferida a partir de los sucesos atestiguados, empero no habría razón suficiente para inferir algo del actuar humano que no se pueda inferir de los otros cuerpos. Pues se planteó que los tres sucesos se dieron de la misma forma por lo cual, los tres podrían implicar la misma función mental, y no habría razón observable o corroborable lo suficiente para creer que alguno de esos sucesos es distinto de los otros dos. Además, siguiendo los propios planteamientos de Descartes, aunque los sucesos fueran físicamente distintos y los cuerpos mostraran dominios distintos de la misma actividad esto no sería muestra de una mentalidad o mayor perfección del cuerpo, simplemente seria debido a su disposición mecánica.

1.4 Distinción mente-cuerpo

Es a partir de las observaciones anteriores como se logra advertir en la perspectiva cartesiana dificultades respecto a su distinción mente-cuerpo, que, primeramente, apuntan a sostener y a atribuir una categoría de acción errónea respecto de las mentes. Dicha perspectiva se aprecia en que durante el proceso de desarrollo de una acción, en los cuerpos, se puede ver ese suceso como consistiendo de un primer momento disposicional que se advierte como la constitución mecánica de los cuerpos, y un segundo momento en el que se da la realización o el ejercicio de la acción a la cual estaban dispuestos tales mecanismos. Estos momentos, de igual manera, se podrían reconocer en las mentes, donde el primero sería la planificación de la acción y el segundo la realización de la acción. Seria en ese primer momento de las mentes donde Descartes estaría señalando una distinción categorial, la cual, a diferencia del proceso de acción mecánico, en las mentes dicho momento consistiría de causas no observables o no pertenecientes al ámbito físico. Así como en que las acciones de dicha naturaleza vienen

provistas o impregnadas de intelección, mientras que su realización o ejercicio son efectos observables a los cuales se les ha transmitido dicha facultad.

Sin embargo, Ryle dirá que, aunque haya una concientización o planificación mental no observable respecto de nuestras acciones observables ésta no hace que tales acciones posean algún tipo de superioridad, intelección, o perfección por encima de las mecánicas. Por el contrario, considera que tanto las operaciones mentales como las operaciones físicas poseen el mismo grado de intelección y tales operaciones son por igual facultades o habilidad adquiridas mediante la práctica. Pues, para él, aunque haya o no reglas y una planificación mental previa a un suceso, tal suceso será tan torpe o inteligente como el individuo que la realiza. Por ejemplo, él menciona que leer o realizar cálculos de manera introspectiva es una habilidad adquirida mediante la práctica de esas facultades de la mente; mas no implica que, por el hecho de llevar a cabo tal o cual actividad de esta manera, dichas actividades introspectivas sean mejores que aquellas que se han realizado de manera pública.

> Porque un niño que está calculando en voz alta o por escrito puede estar razonando correctamente y organizando su demostración metódicamente. Por el hecho de que sus cómputos sean públicos en lugar de privados, no dejan de ser una operación intelectual cuidadosamente realizada. Su comportamiento es, en consecuencia, el ejercicio de una facultad mental (Ryle, 2005: p. 48).

De esta manera Ryle está buscando mostrar que la actividad mental u oculta no es algo que en sí misma posea la misma categoría de los cuerpos y, a su vez, ella misma sea de un tipo o categoría diferente a los mismos. En otras palabras, que mente y cuerpo sean dos tipos de cosas o entes distintos pero pertenecientes a la misma categoría ejecutiva o con las mismas capacidades de acción. Más bien él dirá que cuando Descartes habla de mente y cuerpo en realidad está hablando de facultades, aptitudes, habilidades, capacidades, hábitos o destrezas distintas. Entonces, si bien Descartes propone la idea de dos entidades distintas, Ryle parece advertir que esta dualidad esencial es una idea cartesiana que se generó erróneamente y muy probablemente con posterioridad a la determinación de las facultades que dichos entes poseen. Además, como se ha mencionado anteriormente, la caracterización del ente mental se da con base en la idea de un paralelismo de la operación y funcionamiento del cuerpo o, como menciona Ryle, es una "hipótesis paramecánica descrita con un vocabulario negativo" (Ryle, 2005: p.33). Donde se pretende dar una caracterización de la mente al delimitar su

operación con criterios de la misma índole o categoría operativa que los cuerpos, pero siendo de un tipo sustancial distinto a los mismos.

En este sentido se ha buscado dar cuenta de las observaciones que Ryle hace acerca de porque la dualidad cartesiana es un error categorial; por ello, se ha señalado cómo la idea del ente mental parece tener su origen a partir del ente físico. Ahora, se verá porque para Ryle las facultades, propiamente dichas, adscritas a lo mental y lo físico no pertenecen a la misma categoría de acción como arguye Descartes. Así, parte del error que se advierte en esta perspectiva radica en la noción cartesiana de "disposición", pues Descartes parece entenderlo como un concepto que refiere al señalamiento de un determinado evento, estado, experimentación o cambio de hecho. En este caso, se habla sobre una disposición de acción o de realización de hechos que poseen los cuerpos y las mentes. Por ello, cuando habla de la capacidad de los cuerpos para llevar a cabo una actividad él describe y hace referencia a la composición y organización de los cuerpos o a la constitución del mecanismo de ejecución, así como a los propios eventos que éstos generan y dan paso a la acción en cuestión. De igual manera, cuando refiere a la acción ejecutiva de la mente, él habla sobre eventos o estados mentales los cuales se describen como una sucesión de pasos o leyes a seguir que, a su vez, son una especie de fenómenos no físicos que determinan y generan la acción fisca.

Para Ryle las disposiciones no son propiamente eventos, estados, experiencias, cambios, ejecuciones, ejercicios o acciones particulares de un ente. Así como tampoco es una mera composición y ordenamiento de materiales o leyes, sino es más bien un proceso de adquisición o contención y expresión de una facultad, habilidad o bien dice es "un factor de tipo lógico tal que no puede ser visto o no visto" (Ryle, 2005: p. 47). Es decir, es más bien una susceptibilidad, tendencia, habilidad, hábito o propensión a tener, sufrir, padecer o encontrarse en ciertos eventos, estados, experiencias o cambios particulares cuando se cumple una cierta condición. Esto último se refiere a que las disposiciones requieren del cumplimiento de circunstancias particulares para promover o permitir el desenvolvimiento de sus habilidades y estados.

Ahora bien, Ryle señala que hay diversos tipos de disposiciones, pero de ellas destaca que hay disposiciones simples y disposiciones compuestas o inteligentes. Las disposiciones simples en los objetos serian, probablemente, a lo que Descartes se quiere referir cuando dice

que los cuerpos actúan solo por disposición de sus órganos, pues son disposiciones que no experimentan ni permiten actualizaciones ya que se caracterizan por ser acciones que se ejecutan de la misma manera una y otra vez. Cabe decir que estas disposiciones simples igualmente se dan en el humano y Ryle los identifican como hábitos, además menciona que estos se desarrollan o aprenden mediante la práctica, repetición o condicionamiento de una cierta actividad. Al respecto Ryle nos da dos ejemplos, en el caso de los objetos nos habla de la fragilidad del vidrio y nos dice que "frágil significa que si alguna vez es, o ha sido, golpeado o forzado se hará o se ha hecho añicos" (Ryle, 2005: p. 57), en cuanto a hábitos humanos nos habla sobre el ser fumador y dice "ser un fumador no implica que en este o aquel instante esté fumando, sino que soy propenso a fumar cuando no estoy comiendo, durmiendo, leyendo, atendiendo un funeral o cuando ha transcurrido algún tiempo después del último cigarrillo" (Ryle, 2005: p. 57).

Las disposiciones compuestas o inteligentes serian aquellas que Descartes identifica como producto de una planificación mental. Sin embargo, para Ryle estas disposiciones no necesariamente son mentales u ocultas, sino que pueden darse de manera expresa o corpórea como se mostró con el ejemplo del niño que calcula en voz alta. Estas disposiciones se caracterizan por tener actualizaciones, es decir, son modificables o mejorables; dicha actualización comprende el seguimiento de cánones, reglas o la aplicación de criterios. A diferencia de las disposiciones simples, estas conllevan un aprendizaje más complejo que requiere de la habilidad para entender y aplicar dichos cánones, así como una capacidad para innovar sobre tales actividades. Algunos ejemplos que nos proporciona Ryle son "la inteligencia para boxear se exhibe al dar y parar golpes y no al aceptar o rechazar proposiciones referentes a ellos. Del mismo modo, la habilidad para razonar se muestra en la construcción de argumentos válidos y en la detección de falacias, y no en la admisión de las fórmulas de los lógicos" (Ryle, 2005: p. 63).

Aquí, la diferencia entre las disposiciones simples y compuestas reside en el grado de complejidad de su proceso y no en que las primeras sean operaciones mecánicas o expresas y las segundas sean operaciones mentales u ocultas. Además, se busca mostrar que, a diferencia de lo que piensa Descartes, la ejecución de una acción no es una operación distinta a las operaciones intelectuales de las cuales supuestamente proviene, sino que ambas son una

misma acción. Esto es algo que podemos apreciar haciendo el cambio de perspectiva antes mencionado, es decir, al concebir la mente no como esa corriente de pensamiento inmaterial dentro de nuestra cabeza sino advirtiendo que desde la tercera persona (el observador) las disposiciones son algo apreciable y verificable mediante la conducta, así como a partir de criterios que satisfacen las proposiciones hipotéticas y semihipotéticas que las describen o bien de proposiciones condicionales y de su aplicación particular.

Se está proponiendo, entonces, que tenemos habilidades o facultades disposicionales que pueden ser realizadas de manera explícita o interna y, sea cual fuere la expresión que estas tengan, tales pueden ser y son cuestionadas y evaluadas mediante diversos criterios específicos; asociados a las habilidades que se estén desempeñando, las cuales determinan si dicha habilidad está siendo realizada de manera adecuada, pertinente, correcta y eficiente. Esto último se refiere a que en cualquier ámbito a desempeñar se tiene ciertas metodologías, procedimientos, normas o principios que caracterizan y determinan si una actividad está siendo realizada correctamente o si lo que se está haciendo corresponde con lo que se espera o se debería de hacer. Entonces no es que las facultades mentales y físicas correspondan a entes distintos, sino que hay facultades con diversos grados de complejidad que podemos y aprendemos a ejecutar tanto de manera explícita como íntima (interna).

No obstante, dentro de esa misma perspectiva, dichas facultades no son ellas mismas o en sí mismas "ni informes de estados de cosas observadas u observables, ni tampoco informes de estados de cosas no observados o inobservables. No narran incidente alguno" (Ryle, 2005: p. 69). Esto quiere decir que Ryle al advertir como disposiciones, lo que serían las facultades mentales de Descartes, está dando cuenta que estas no deben ser entendidas como eventos, sino como capacidades y se describen o dicen mediante proposiciones condicionales o hipotéticas, las cuales describen algo que podría, puede o tiene posibilidades de ocurrir. Estas proposiciones, dice Ryle, se asemejan a las autorizaciones de inferencia puesto que, si son verdaderas las condiciones que estos establecen, como el que un vidrio se romperá tras ser golpeado por una piedra, estarán siendo satisfechas por los sucesos y condiciones que estas mismas narran o predicen.

1.5 Conclusión

Encontramos, entonces, en la postura de Descartes una teoría del pensamiento que es ampliamente considerada e incluso aceptada, en la cual a partir de la perspectiva de primera persona se genera un dualismo que da paso a determinar que los cuerpos humanos son los únicos poseedores de mentalidad. No obstante, al ser considerada desde la tercera persona nos encontramos con diversas dificultades incompatibles con dicha postura, en la cual sus criterios mentales; como el obrar por conocimiento, el uso de palabras e incluso el hecho de llevar a cabo actividades ocultas; no garantizan las facultades mentales del cuerpo que realiza tales acciones.

A partir de la postura de Ryle, podemos encontrar en él a un crítico del dualismo cartesiano que parece conciliar la posibilidad de que la mentalidad pudiera ser considerada algo viable para otros cuerpos o que al menos pone en duda el carácter oculto de las actividades intelectuales, así como la incapacidad de corroborar la existencia de una supuesta sustancia mental. Pues aquí ya se habla de las facultades mentales como algo cuya verdad o falsedad puede ser corroborado mediante las proposiciones condicionales en que se expresan y su aplicación particular.

CAPITULO II
TURING: DUALIDAD SOFTWARE-HARDWARE

2.1 Introducción

La comprensión del pensamiento humano es una de las cuestiones que despierta mayor incertidumbre e interés no solo dentro de la filosofía, sino en cualquier área de investigación. En filosofía de la mente una de las posturas de gran influencia al respecto ha sido el dualismo mente-cuerpo, la cual también se ha arraigado en las creencias comunes de las sociedades humanas especialmente por su ascendencia en la teología y por ende en el ámbito religioso, pues comúnmente es aquí donde se habla de un alma inmortal, independiente, trascendente a lo material con propiedades y capacidades distintas del cuerpo que ocupa. Alan Mathison Turing, un brillante matemático y pionero de la computación, nacido el 23 junio de 1912 en Londres (Coperland: 2017: p. 4), influido en parte por ideas dualistas abordó la cuestión filosófica del pensamiento desde la perspectiva del funcionamiento de las computadoras. Ámbito en el cual destacó al punto de ser considerado padre de la ciencia de la computación.

Turing se interesó especialmente por la programación de computadoras, fue ahí donde advirtió que dichos programas podrían ser elaborados para exhibir capacidades singularmente humanas. Es más, notó que, de hecho, las computadoras ya podían exhibir algunas de esas capacidades e incluso dominarlas con mayor eficacia que el humano, tal fue el caso de los cálculos matemáticos. Esto lo llevó a cuestionar si las máquinas pueden pensar y su respuesta no fue ignorada por la filosofía. Por el contrario, diversos autores como Ned Block, Hilary Putnam, Jerry Fodor entre otros se profirieren ante tales ideas. Muchas de las cuales serán adoptadas por el funcionalismo quien ofrece una tesis filosófica sobre la naturaleza de los estados mentales, la cual se apoyará en la descripción del funcionamiento de las máquinas de Turing.

2.2 Precedentes

Alan Turing desarrolla una concepción de pensamiento que se ve influenciada por el filósofo Rene Descartes quien sostuvo una postura dualista del pensamiento en la cual, probablemente y en parte debido a los avances científicos de su época, había señalado ciertas capacidades o

habilidades humanas como aspectos o características que nos permitirían distinguir a los humanos como seres pensantes respecto de los animales, máquinas o autómatas como seres carentes de pensamiento. Tales aspectos son el que los seres dotados de pensamiento tienen capacidad de usar palabras y signos en una amplia gama de modos para responder conforme a lo que se está diciendo en su presencia, pero más precisamente son capaces de tener un lenguaje o bien de dar una respuesta significativa. Así mismo señala que los humanos tienen la característica de obrar por conocimiento, lo cual aparentemente es lo que les permite tener un alcance de acción considerablemente amplio o universal. Pues, la razón les permite, por un lado, impregnar de significado sus acciones y, por otro, permite algo como una capacidad numéricamente indeterminada o considerablemente amplia para la realización de diversas acciones que, además, resulta fisiológicamente económica en el sentido de que les permite llevar a cabo tales tareas sin requerir de un cambio o actualización de su composición y constitución física (Descartes, 2002: p. 114).

Además, en Descartes encontramos una caracterización de la interacción mente-cuerpo en la que se detalla más la descripción de la transición de las acciones mentales a través del cuerpo que la propia conformación de dichas operaciones mentales. Pues respecto a los eventos mentales encontramos que Descartes los identifica y los define como voluntades (o acciones) y pasiones (o entendimientos), si bien él describe el tipo de operación que son y cómo operan, él no cuestiona ni explica con detalle cómo se generan estas operaciones. Dado que para él tales son propiedades mentales, es decir, son características de la mente que nos resultan completamente intrínsecas y que ya conocemos.

Por otro lado, Descartes señala que la capacidad de acción de las máquinas y animales es considerablemente más limitada que la de los humanos. Asimismo, advierte que estos seres no poseen lenguaje ni la capacidad de obrar por conocimiento, lo cual lo lleva a asumir que tales seres no poseen mentalidad. Estas carencias parecen sentenciar su capacidad de acción a limites muy determinantes, a saber, los límites de lo corpóreo. Pues él concibe a estos seres como cosas imperfectas y restringidas, donde sus limitaciones se caracterizan por la forma, figura o constitución de su propio cuerpo, es decir, por la disposición espacial de sus mecanismos y (el tipo de) material del que están constituidos. En este sentido, él parece asumir que en tales seres lo que permite la realización de acciones es una suerte de interacción

coercitiva o causal entre los objetos físicos involucrados en la realización de tal o cual acción. De tal forma, para Descartes tanto la cantidad como el tipo de capacidades y acciones que presenta una máquina o un animal están sujetas estrictamente al hecho de tener un cierto tipo de cuerpo con tales y cuales características.

Por ejemplo, pensemos en un bolígrafo sencillo; para que realice adecuadamente la labor de dejar un rastro de tinta sobre ciertas superficies se requiere de una superficie adecuada como el papel, un cartucho de tinta, tinta, la caña o el cuerpo de la pluma, un tapón y la punta, la cual, a su vez se compone de un balín, porta balín y un soporte. Cada una de estas partes tiene un diseño determinado que al ser ensamblados bajo cierto orden se da el que la tinta entre en contacto con una parte del balín, impregnándolo de una porción de tinta que, posteriormente, cuando el balín entra en contacto con la superficie de papel y es arrastrado sobre ésta, rota de tal forma que la parte impregnada de tinta toca el papel impregnándola de esa porción de tinta mientras otra parte del balín se impregna de otra porción de tinta.

De esta manera para Descartes, las máquinas, ya sea un bolígrafo o máquinas con mecanismos mucho más complejos como un reloj, un autómata o un animal operan solo en virtud de las disposiciones de sus componentes u órganos, y aunque pudieran llevar a cabo cierta actividad con mucho más destreza que nosotros no podrán llevar a cabo con igual destreza la totalidad de las habilidades humanas, mostrando así lo limitado de sus acciones. Ya que para cada nueva acción requerirían una disposición física distinta de mecanismos o un aumento de los mecanismos requeridos para sus acciones. Consiguientemente para Descartes y su época resulta inconcebible una máquina que pueda operar con un sistema puramente mecánico y a su vez sea capaz de actuar o responder a situaciones tan diversas como las que podría enfrentar un humano, especialmente a aquellas donde se hace uso de los aspectos propios de la razón (Descartes, 2002: 109, 115, 116).

Esta noción mecanicista cartesiana de las máquinas es compartida por otras dos posturas filosóficas a tomar en cuenta, que surgen en respuesta al dualismo, pues igualmente son antecedentes de cómo van a operar las máquinas de Turing, así como nos indicarán las ideas bajo las cuales Turing construye su noción de pensamiento. Una de ellas es el *conductismo,* postura sostenida por Ryle, en la cual se propone que toda actividad, acción o respuesta (física o mental) de un individuo es explicable a propósito de su conducta observable. Esto se debe

a que el conductismo advierte que todos los organismos sean humanos o no-humanos realizan acciones empíricamente observables. Identifica, además, que los cuerpos están sujetos a condiciones orgánicas que los constituyen, las cuales determinan su capacidad de acción. Asimismo, se ven inmersos a la interacción con su ambiente pues constantemente son afectados por diversos estímulos ambientales que inmediatamente parecen incitar acciones, reacciones o conductas en los organismos. Siendo así, considera que las acciones pueden y deben ser explicadas tomando en cuenta tales disposiciones (habilidades propias o pertinentes a una cierta constitución), así como las diversas condiciones (interacciones ambientales) para la conducta.

Así, para el conductismo hay un medio ambiente (un contexto de circunstancias, criterios o leyes) que influyen o inducen las acciones de los organismos mediante estímulos o eventos físicos ambientales, ya que los cuerpos inevitablemente responden con un comportamiento al que son afines y que suele ir acorde a sus disposiciones o capacidades motrices (físicas) (Graham, 2017). Por ejemplo, para un conductista el que un bolígrafo deje un rastro de tinta sobre alguna superficie es un comportamiento que se debe a que hay un cierto estímulo externo, como una cierta fuerza, que al entrar en contacto con el bolígrafo provoca que su constitución física reaccione naturalmente o automáticamente dando lugar a ejercer dicha afinidad conductual propia de sus capacidades o aptitudes físicas. Por ello aquí se considera que cualquier actividad que realice un organismo debe poder ser explicada en temimos conductuales, prescindiendo de los términos mentales, de tal manera que se pueda dar una descripción de ellas mediante proposiciones que planeen o relaten las disposiciones implicadas y condiciones requeridas para que tal o cual acción pueda tener lugar.

La otra postura se opone a la idea de que una constitución física, por si sola o en sí misma, no pueda ser lo suficientemente compleja como para responder con igual eficacia que lo haría la mentalidad a la que alude el dualista. De hecho, sostiene la existencia de una constitución física capaz de realizar cualquier actividad propia de la mentalidad. Esta postura es la *teoría de la identidad mente-cerebro*, al igual que los conductistas toma en cuenta las interacciones estimulo-respuesta y la conducta observable, sin embargo, ellos dan mayor importancia al cuerpo mismo. Pues consideran que las interacciones se dan en cuanto existe una cierta constitución física (cuerpo u objeto) cuyas partes y materiales que lo conforman son los

responsables directos de presentar ciertas tendencias de acción que, igualmente, se activan o llevan a cabo cuando un cierto estimulo exterior entra en contacto con el objeto en cuestión (Smart, 2000).

Al igual que el conductismo para la teoría de la identidad mente-cerebro todas las acciones emitidas en términos mentales son también explicables mediante las interacciones estimulo-respuesta, empero, aquí dichos estados mentales (creencias, deseos, temores, etc.) son idénticos a los sucesos físicos o procesos cerebrales que se describen mediante los términos de conductas observables (Smart, 1959). Así, cuando alguien informa sobre un cierto estado mental como "creo que está anocheciendo" esa persona de hecho está reportando un suceso físico o proceso cerebral, en tanto que dicho informe es una experiencia propia de la tendencia de su constitución física a percibir los cambios y estímulos de luz que emite su entorno.

Tanto en el conductismo como en la teoría da la identidad, toda actividad o conducta que sea capaz de llevar a cabo un humano es un producto propio de las tendencias conductuales o físicas del cuerpo y de los estímulos que entran en contacto con tales condiciones físicas. Cabe enfatizar que en la postura cartesiana esto también se da, pero solo en los animales y máquinas. Empero, el conductismo y la teoría de la identidad tienen un carácter antidualista, por un lado, el conductismo simplemente evita la idea de una substancia mental al decir que toda actividad es explicable mediante las interacciones estimulo-respuesta de los individuos, mientras la teoría de la identidad adopta una oposición más directa al decir que no hay tal sustancia ya que en realidad los estados mentales son estados cerebrales. De esta manera, la teoría de la identidad no estaría de acuerdo con la idea cartesiana de que es imposible la existencia de una constitución física lo suficientemente compleja como para poder realizar las acciones mentales propias de un humano, pues para ellos la propia constitución física humana es y posee en sí misma la complejidad suficiente que permite producir tales acciones. Por lo cual para ellos no habría dificultad en creer que solo mediante propiedades físicas se puede lograr un sistema que pueda llevar a cabo y satisfactoriamente tantas actividades como las adscritas a la mentalidad.

No obstante, sí consentirían la idea cartesiana de que el pensamiento es algo exclusivo del humano o más bien que hay una exclusividad humana para realizar ciertas acciones ya que, como se dijo, tales capacidades van de acuerdo a las propiedades de la constitución fisca

humana, es decir, son producto de ese tipo específico de constitución física. Pues se debe entender que la teoría de la identidad está postulando a los estados mentales humanos como cosas indistintas, de manera alguna, o como siendo lo mismo que los estados cerebrales (físicos) propios del cuerpo o cerebro humano. Entonces, para el teórico de la identidad los humanos son capaces de mostrar un dominio de habilidades de las que ningún otro ser es capaz y esto se debe solo a causa de su propia constitución física mas no porque tengan algo así como una substancia mental.

2.3 Turing

A sus diecinueve años Alan Turing escribió un breve ensayo titulado *Nature of Spirit* (Turing, 1932) en el cual expresó su inquietud respecto a querer explicar la interacción entre objetos físicos (cuerpos) capaces de llevar a cabo actos, y un espíritu (alma, voluntad, mente) que al disponer de dichos objetos promueve determinantemente esos actos, pero a su vez muestra independencia de ellos. En dicho ensayo Turing advierte que el espíritu no está ligado a un solo cuerpo vivo, sino que puede ocupar o interactuar con otro cuando aquel que ocupaba muere. Esto claramente refleja su interés por la explicación de la mentalidad humana, así como muestra afinidad hacia la postura dualista cartesiana, pues al igual que Descartes está postulando dos entidades independientes. Sin embargo, a diferencia de esa postura dualista tradicional, Turing no solo caracterizará la relación e interacción mente-cuerpo como causal, sino que da un paso más allá que Descartes al desarrollar una explicación mucho más sistemática, compleja y detallada respecto a cómo opera concretamente el pensamiento al momento de generar y llevar a cabo una acción o evento mental.

Hay que decir que en *Nature of Spirit* Turing formula el primer esbozo de su teoría de la mente, por lo cual es natural que estos primeros planteamientos resulten burdos. No obstante, partiré de aquí para dar pauta de ciertas nociones que caracterizarán su concepción de pensamiento. En este ensayo, Turing señala a los átomos y electrones como capaces de llevar a cabo actos o de interactuar en el espacio y menciona que éstos componen o constituyen los cuerpos (objeto físico o materia). Por lo cual, cabe aclarar que aquí Turing ya está haciendo referencia a mecanismos electrónicos computacionales donde se trabaja con la emisión, flujo y control de los electrones que dan paso al procesamiento de información mediante diversos mecanismos eléctricos. Además, también pareciera referirse a que las acciones de estas partes

de la materia (a saber: los átomos) serian equivalentes a algo así como las unidades base, primarias o mínimas de acción de los cuerpos. Asimismo, él hace una mención burda de que las acciones de los átomos pueden ser reguladas o determinadas de dos maneras: una es mediante la voluntad humana (espíritu) que se ejerce en el cerebro y, la cual, se potencializa o amplifica por el resto del cuerpo humano; la otra es por mera casualidad ("be regulated by pure chance") o por coacción de los cuerpos (Turing, 1932).

Según Turing, la capacidad del espíritu para regular o determinar las acciones de la materia le permite otorgar a la materia un sentido (curso o guía) de la cual carece por si sola la materia. Más adelante, se verá a que se refiere Turing con esta idea de que las acciones de la materia pueden tener "sentido". Otro aspecto a tomar en cuenta es cuando menciona que los cuerpos vivos son los que tienen un cierto mecanismo que es capaz de capturar al espíritu haciendo que este cuide del cuerpo y haga uso de él o disponga de su constitución física. Este es el sentido en que, para Turing, el cuerpo le estaría dando algo a hacer al espíritu. De esta manera, Turing estará promoviendo una relación de mayor reciprocidad entre mente y cuerpo que la propuesta por Descartes, pues ambas partes de alguna manera proveen de algo a la otra parte. Por otro lado, dice que el espíritu (mente) es algo que puede ser o existir y hacer cosas con total independencia del cuerpo, así como puede ocupar distintos cuerpos una vez que el mecanismo que lo retenía en un determinado cuerpo muere.

Tales planteamientos iniciales de Turing sobre la mentalidad humana y su relación con el cuerpo serán un tema que permearía sus investigaciones y trabajos posteriores, ya que más tarde identificaría y desarrollaría dichas nociones dentro de su proceso de invención, construcción y especialmente para él en la programación de máquinas (computadoras). Pues advirtió que el funcionamiento de estas máquinas se ajustaba a esas nociones características de la mentalidad humana que había identificado en *Nature of Spirit*. Veremos, entonces, que las máquinas que propone Turing implicarán la elaboración de un sistema que implementa la división dualista mente-cuerpo (software-hardware), es decir, tendrán una parte mental (software) e independiente de objetos físicos que a su vez será capaz de operar en distintas máquinas físicas (hardware) y dispondrá de estos dando sentido a sus acciones.

2.3.1 Máquinas de Turing

Cuando Turing dice que el espíritu es capaz de regular las acciones de los cuerpos y darles un sentido quiere decir que, para él, el espíritu es algo capaz de hacer un uso específico de las distintas disposiciones propias a cada tipo de constitución física de tal manera que sea capaz de generar una cierta acción en concreto. Empero y aún más importante es que esta capacidad del espíritu para regular las acciones que los cuerpos llevan a cabo o emiten es dada o determinada por el espíritu conforme a una cierta finalidad, función o sentido de su realización. Por ello Turing advierte que los actos generados por los cuerpos con espíritu no son meras casualidades.

Ahora bien, la parte mental de las máquinas de Turing la llamará *tabla de instrucciones* o *programa* y en ella residirá la capacidad o aptitud para realizar tareas específicas, enfocadas o dirigidas a un cierto contexto o ámbito. No obstante, para Turing el que la tabla de instrucciones genere acciones enfocadas no implica que ella misma se encuentre limitada a tales acciones y tampoco impide o limita la capacidad de la máquina para llevar a cabo otras acciones. Por el contrario, estas tablas de instrucciones les permitirán ser máquinas universales pues, según Turing, su sistema operativo está ideado para imitar cualquier aptitud intelectual, comportamiento o acción humana, y por ende para realizar cualquier tipo de operación propia de un humano (Turing, 1950a: p. 8, 9, 13; D.C. INCE (ed.), 1992: p. 111-§5). Esta universalidad de las máquinas de Turing consiste en que un mismo sistema operativo o programa es capaz de ser instruido para llevar a cabo diversas tareas específicas, por lo cual no sería necesario crear una máquina nueva cada vez que se quiera realizar una tarea nueva; simplemente habría que programarla para poder realizar dicha tarea.

Cabe señalar que dentro de las múltiples operaciones intelectuales que se contempla puede realizar la máquina de Turing destacan su capacidad de realizar operaciones matemáticas de distintos grados de complejidad, la posibilidad de dominar o manejar símbolos lingüísticos e incluso traducir cualquier lenguaje o hacer criptografía, así como participar satisfactoriamente en distintos juegos como el ajedrez, póker, etcétera (D.C. INCE (ed.), 1992: p. 87-§2, 103-§17 y 18, 117). Por ello, Turing encuentra en su máquina una interacción mente-cuerpo muy similar a la humana que exhibe un amplio dominio de habilidades que normalmente son consideradas solo propias de la mentalidad humana. De hecho, para Turing

en realidad no parecería haber diferencia alguna respecto a lo que hacen sus máquinas y lo que hace un humano, pues veremos que para él tanto las acciones de los humanos como las de sus máquinas pueden ser descritas de la misma manera.

En este sentido, la postura de Turing no se distingue tanto de la postura cartesiana en cuanto a los parámetros de acción o actividades propias de una mentalidad. No obstante, en principio la distinción más obvia entre el dualismo cartesiano y el dualismo de las máquinas de Turing se observa en el tipo de objetos físicos capaces de contener mentalidad. Pues Turing muestra mayor flexibilidad respecto al tipo de objetos o cuerpos que son capaces de contener mentalidad. Desde *Nature of Spirit* se podía advertir ésta flexibilidad cuando dice que el espíritu es captado por "cuerpos vivos", ahí ya daba paso a suponer que hubiera una diversidad de objetos físicos, además del cuerpo humano, que de hecho pudieran presentar mentalidad. Esto marcará una clara diferencia con la postura dualista cartesiana pues llevará a Turing a cuestionar la exclusividad humana del pensamiento (mente o razón) y a aseverar que sus máquinas de hecho piensan, lo cual sería algo inconcebible para Descartes.

Esta diferencia no solo se opone al dualismo cartesiano, sino también a la teoría de la identidad ya que irá contra la idea de que la mentalidad es algo exclusivo de una constitución física determinada. De tal manera las tablas de instrucciones de Turing podrán ser implementadas en distintos cuerpos o hardware sin importar cuan diversa pueda ser su constitución material (Turing, 1950a: p. 10); por ello, sus acciones también presentan la peculiaridad de poder ser realizadas por distintas vías físicas o de diversas maneras. Esta peculiaridad, también llamada múltiple realización de las máquinas de Turing da paso a una relación mente-cuerpo distinta de la planteada por Descartes, ya que en cierto grado veremos que las características físicas influirán sobre la eficacia y eficiencia mental implicada en la realización de tareas.

Así, el hardware de la máquina que propone Turing adapta la idea antes mencionada de que los cuerpos vivos tiene un mecanismo que captura al espíritu, pues independientemente de los materiales que se usen en su elaboración él dirá que éstos tendrán que cumplir con tres capacidades operativas base, al parecer responsables de su múltiple realización, ya que la labor que desempeñan puede ser satisfecha de diversas formas y con materiales distintos. (Turing, 1950a: p.10; D.C. INCE (ed.), 1992: p. 23). En *Computing Machinery and*

Intelligence menciona que el tipo de máquina o computadora capaz de contener mentalidad son las computadoras digitales y dichas computadoras tiene un hardware que consta de al menos tres partes: almacenamiento, unidad ejecutiva y el control (Turing, 1950a: p. 7-8). Cada una de estas partes serán las características básicas y capacidades operativas del hardware que debe tener toda realización física de su máquina para poder contener mentalidad, podemos decir que para Turing cualquier objeto físico que cuente con una constitución física con estas características y capacidades operativas sería lo que él llamaba "un cuerpo vivo capaz de contener al espíritu" en *Nature of Spirit*.

El almacenamiento dice Turing es un depósito de información finito o bien es la disposición del hardware a almacenar una carga finita de información y este almacenamiento puede estar constituido de distintos materiales físicos. Por ejemplo, puede ser una larga pila de papeles, cinta perforada, cinta magnética, disco duro, tarjetas de memoria, discos ópticos, etc. Empero, a pesar de que todas presentan la disposición para almacenar información, cada tipo de material con el que esté constituido presenta otras disposiciones que pueden limitar o potenciar el funcionamiento de la máquina de distintas maneras. Ciertos materiales pueden ser más inestables y fácilmente se puede perder información, otros requieren de ciertos cuidados como mantenerlos a ciertas temperaturas, cada uno tiene distintos tiempos para acceder, manipular o mover la información de un punto a otro, y por supuesto la capacidad de almacenamiento que presentan no es la misma. Algunos requieren físicamente de mucho más espacio para tener un mayor almacenamiento, etcétera (D.C. INCE (ed.), 1992: p. 22-24). Turing, además, menciona que el almacenamiento es la parte física que representa mayores limitaciones a las capacidades o aptitudes mentales que pudiera llevar a cabo la máquina, pues será aquí donde se contendrá a la mente (el software) y, por consiguiente, a la diversidad de tareas que esta puede ofrecer. Por ello, de la cantidad de información que pueda contenerse en la memoria y de la eficacia con que se almacene también dependerá la cantidad de actividades mentales que pueda realizar físicamente la máquina. (Turing, 1950a: p. 8, 12)

Por otro lado, la unidad ejecutiva, procesadora o la parte aritmética central (CA) es la parte que lleva a cabo las diversas acciones o cálculos mentales de la máquina. Para entender el funcionamiento de la unidad ejecutiva, Turing retoma la idea de interacciones estimulo-respuesta de los conductistas, pues dirá que esta parte de la máquina en principio o a grandes

rasgos se puede entender como la ejecución de una acción que tienen por inicio una entrada o estímulo y el término de esa acción se identifica por tener una salida o dar una respuesta a dicha entrada. Sin embargo, su funcionamiento está ideado de una manera más compleja que eso, pues en ella se contemplan muchos más sucesos que el de una única operación individual y lineal, es decir, el de una única entrada y una única salida. Su proceso contempla la interacción de una diversidad de posibles entradas, así como una diversidad de posibles respuestas. Se está hablando, entonces, de un sistema que es capaz de ejecutar un conjunto de operaciones individuales que se realizan, dan o ejecutan a la manera de los estimulos-respuestas. El tipo y conjunto de operaciones individuales que pueda ejecutar varía de una máquina a otra según la tarea que deban realizar: estas pueden ser la acción física de mover información de un punto a otro dentro de la memoria o entre los propios mecanismos de ejecución, para manipular la información de cierta manera. Por ejemplo, al implementar la información de la memoria, de las entradas e incluso de salidas internas al sistema (operación individual) dentro de mecanismos que permiten realizar diversas operaciones aritméticas como multiplicaciones, sumas, restas u operaciones lógicas tipo "A & B", "A ∨ B", etc. (Turing, 1950a: p. 8; D.C. INCE (ed.), 1992: p. 26-27).

Se está hablando, entonces, de un conjunto de acciones y cada acción diremos es una operación individual, un microcircuito o un elemento de circuito, la cantidad de elementos de circuito que conformen una unidad ejecutiva dependerá de la complejidad y capacidad de acción de la misma. Así, si bien esta unidad recibe entradas del medio con el cual interactúa (entradas externas) también puede tener entradas internas como sería el trasladar alguna información de la memoria al procesamiento de un microcircuito. De igual manera ocurriría esto con las salidas, es decir, hay salidas externas a la unidad e internas a ella. Ahora bien, la manera en que la unidad procesadora o el conjunto de elementos de circuito llevan a cabo sus acciones se pueden entender en principio como el hecho de acarrear información de un lugar a otro; empero, dependiendo de la tarea que se quiera realizar y la entrada que se reciba la información será transitada o llevada mediante ciertas vías o circuitos. Pues cada elemento de circuito está diseñado para fungir con un determinado procedimiento resolutivo propio de una operación lógica, aritmética o del tipo pertinente para la tarea a realizar.

Por ejemplo, si se requiere realizar una operación lógica del tipo disyuntiva (X o Y) lo que ocurrirá es que la tabla de verdad correspondiente a la forma en que opera esta función será representada físicamente, esto podría ocurrir mediante circuitos eléctricos o algún otro material físico que permita ese modo de representar dicha operación. En el caso de un circuito eléctrico, esta función se recrearía usando cables, botones, resistencias, leds, transistores, etc. De tal manera esta operación presenta cuatro posibles casos donde tres de ellos son verdaderos cuando por lo menos una de las entradas es verdadera, y solo una resultará falsa cuando ambas entradas sean falsas, a saber: V + V = V, V + F = V, F + V = V, F + F = F. El circuito podría entonces ser representado de la siguiente manera: tenemos un botón (entrada) A, un botón (entrada) B y una luz led. Cuando el botón A es accionado, pero B no, permite la entrada suficiente de electricidad en el circuito como para que el led se ilumine (V + F = V); cuando B es accionado y A no, también corre suficiente electricidad para iluminar el led (F + V = V); cuando A y B son accionados a la vez también corre suficiente electricidad para iluminar el led (V + V = V); por último, cuando A y B no son accionados no corre electricidad por el circuito como para iluminar el led (F + F = F). Así este principio puede y será representado de distintas maneras y con distintos materiales. Lo mismo se aplicaría para otro tipo de funciones.

Al igual que con el almacenamiento, la eficacia y eficiencia con la cual la máquina ejecuta sus operaciones individuales dependerá de los materiales con los que pudiera estar conformada la unidad ejecutiva. Algunos de esos materiales podrían ser el silicio, electricidad, tubos de vacío, rieles electicos, circuitos digitales o bien también podríamos pensar en otro tipo de circuitos a gran escala como la infraestructura de los trenes, carreteras o incluso podemos compararlo con los mecanismos y materiales de diversas formas de comunicación, es decir, el teléfono, cartas o entre personas, etc. La unidad ejecutiva es, entonces, toda la infraestructura que se utilizará para realizar las transferencias de información y las operaciones o procesamientos físicos de la misma en la transición de entradas y salidas, ya sean externas o internas.

Por último, el control será una especie de intermediario entre el espíritu (mente) que está contenido en la memoria y la unidad ejecutiva, pues se encarga de vigilar que esta última obedezca correctamente a la mente. Cabe señalar que Turing caracterizara a la mente como

un conjunto de instrucciones u órdenes. Así, el control debe de hacerle llegar adecuadamente dichas órdenes a la unidad ejecutiva para que esta las ejecute en el orden pertinente y así se pueda realizar apropiadamente la tarea. Turing dirá que el control está constituido de tal manera que necesariamente eso ocurrirá. (Turing, 1950a: p. 8) El control será, entonces, el sistema encargado del manejo correcto de la información, por ello éste se servirá de instrumentos y mecanismos como los traductores (codificadores y decodificadores) que le permitirán identificar apropiadamente la información de las entradas y salidas tanto externas como internas, esto incluye las órdenes de la mente, para posteriormente canalizarlas adecuadamente para su procesamiento. Estos traductores del control operan de la misma manera que los circuitos vistos en la unidad ejecutiva, solo que su operación está diseñada para transformar la información de un lenguaje a otro o de una cierta presentación a otra, como puede ser pasar de ser un sonido a una palabra escrita. De la misma manera en que ocurriría con la traducción del código morse, aquí también se elabora un circuito (conjunto de acciones) capaz de actuar conforme a principios que permitan hacer esa traducción.

Ahora bien, parte de las traducciones que realiza el control son de las entradas y salidas externas, es decir, mediante sensores externos u órganos externos transforma los estímulos del ambiente en pulsos eléctricos o cual fuera el método de procesamiento de información que maneje la máquina y viceversa. Igualmente cuenta con traductores internos que le permiten cambiar la presentación de la información ya sea para que los circuitos de la unidad ejecutiva realicen su labor o para que sea adecuadamente identificada en las ordenes de la mente, pues estas últimas están escritas en lo que Turing llama un leguaje de máquina, y así se obtenga el orden o la manera en que dicha información debe ser procesada. El control también contará con sistemas de cronometraje: estos son sistemas que permiten regular la rapidez con la que se emite o procesa una información. Así como tendrá métodos para priorizar ciertas operaciones sobre otras como es el que las instrucciones de la mente tengan un orden de ejecución, es decir: que estas se encuentren ordenadas a manera de una lista que debe ser leída y ejecutada en ésa secuencia. Otro elemento a mencionar del control es la forma en que está almacenada la información pues a propósito la información está segmentada y almacenada en paquetes pequeños para facilitar su manejo, pero además a dichos paquetes se les asigna una clasificación que permitirá su identificación de forma

sistemática. Tal clasificación a su vez es dada conforme a la codificación de información implementada en la máquina (Turing, 1950a: p. 8; D.C. INCE (ed.), 1992: p. 22,34, 46-51).

Un ejemplo sencillo que ofrece Turing al respecto es el de una instrucción codificada en el lenguaje de la mente o lenguaje de máquina dónde dicha instrucción estaría expresada de la siguiente manera: 6809430217. Ahí nos dice que ésta instrucción es un paquete de información codificado que consta de 10 dígitos, donde el 6809 y 4302 son las clasificaciones de otros dos paquetes de información y el 17 es la operación que se deber realizar con dichos paquetes. Para nosotros, o bien en lenguaje natural, lo que ésta instrucción diría es: suma la cifra almacenada en la posición 6809 a la situada en 4302 y devuelve el resultado a la última posición del almacenamiento (Turing, 1950a: p. 9). Así pues, el control es tanto una serie de circuitos de lectura, de codificación, de decodificación, de ordenamiento que dan paso al uso determinado de los diversos mecanismos de la máquina, así como a la adecuada manipulación de la información según las especificaciones de la mente.

Para Turing cualquier objeto físico (hardware) que presente estas capacidades operativas podría contener mentalidad (software) sin importar de que materiales este constituido o mediante que materiales se lleven a cabo dichas capacidades, pues aquí lo importante es tener un conjunto de elementos físicos que al trabajar permitan la contención, detección, lectura y la apropiada ejecución de las instrucciones de la mente. Pues, si bien la memoria es el mecanismo principal que contiene la mente, sin los mecanismos del control y la unidad ejecutiva la mente no podría desempeñarse adecuadamente pues no podría hacer uso de ninguna disposición física y simplemente estaría contenida en algún medio físico. Se podría decir que la memoria es la primera instancia que da paso a que la mente se involucre en las acciones del cuerpo, mientras el control son los mecanismos que hacen posible que la mente disponga plenamente del cuerpo (hardware), a saber, que disponga de la información almacenada y los mecanismos de ejecución y, por último, la unidad ejecutiva son los mecanismos que ejercen o realizan físicamente las órdenes de la mente.

De esta manera Turing formula un argumento en contra de la exclusividad humana del pensamiento, donde recalca que los materiales con los que se constituye el cuerpo no son un factor fundamentalmente determinante en lo que respecta a la capacidad de realizar acciones mentales. Sin embargo, sí influyen en la eficiencia y eficacia de tales acciones ya que algunos

materiales son mejores que otros en la precisión, rapidez, contención y manejo de la información. Para Turing, la relevancia de los cuerpos recae en la forma en que sus componentes materiales están dispuestos u organizados para poder dar cauce a las tareas de la mente. Para él, un cuerpo capaz de contener mentalidad es un cuerpo que esta de alguna manera diseñado para ello y puede haber más de un diseño que funcione para dicha labor; de aquí la idea de que Turing sostiene una relación de mayor reciprocidad con la mente que la propuesta por Descartes y los teóricos de la identidad. Veremos que esta noción de Turing respecto de priorizar el funcionamiento o la capacidad operativa también se extenderá a su noción de mentalidad.

Esta concepción de Turing será compartida por el funcionalismo cuya postura filosófica apoya, especialmente, la idea de caracterizar a la mente como un sistema que describe y contiene el proceso de elaboración de una tarea con base en los principios causales operativos o funcionales que debe satisfacer la realización de dicha tarea. Por decirlo así, su caracterización de lo mental se enfocará en la capacidad para dar solución a un problema. Por ello, para Turing y los funcionalistas la capacidad que muestran las máquinas de Turing en llevar a cabo satisfactoriamente acciones, no radica simplemente en su capacidad de reaccionar (responder) a estímulos (entradas). Sino que las operaciones del software (mente) están diseñadas para efectuar una serie causal de estados mentales que desembocarán en una determinada salida dada una cierta entrada. Así al ser ejecutadas por el hardware generan que las operaciones físicas, también, estén dirigidas a la realización de una cierta actividad o rol causal. Por ejemplo: el rol causal de un corazón seria bombear sangre, el rol causal de un bolígrafo es dejar un rastro de tinta sobre ciertas superficies, mientras que el rol causal de una máquina de refrescos es emitir refrescos. Entonces, de acuerdo al rol causal que se quiera generar, el software contendrá la descripción de cómo llevar a cabo esa labor.

Se está sosteniendo que lo que caracteriza al pensamiento radica en esa capacidad para describir la realización de funciones de tal forma que puedan ser desempeñadas. En el caso de las máquinas de Turing esta capacidad se da en tanto que el software está compuesto por un sistema de instrucciones que describe el proceso de la elaboración de una cierta tarea. Piénsese por ejemplo en la manera en que un recetario te dice cómo elaborar un platillo o en el instructivo de un juego que te indica las reglas del juego, los movimientos que puedes

realizar, como se deben usar las fichas o cartas que contiene, etcétera. Por ello, Turing dirá que sus máquinas están ideadas para realizar todo tipo de operación intelectual o cognitiva que normalmente se considerarían propias e incluso exclusivas del pensamiento humano (Turing, 1950a: p. 8). Pues se advierte que este diseño o estructuración del software a manera de instructivo permite describir e imitar cualquier tipo de actividad, solo hace falta identificar las condiciones bajo las cuales se da tal o cual acción y el proceso que acarrea su elaboración. Así podrán contener cualquier tarea o acción que pueda ser descrita en esos términos ya sea escribir, tocar una canción, jugar ajedrez, realizar operaciones matemáticas, mostrar una sonrisa (o una representación de ello), un llanto, emitir palabras de desánimo, alegría o cualquier cosa que pudiera esperarse ocurriera bajo ciertas circunstancias y cuyo proceso de realización causal pueda ser descrito y representado mediante instrucciones.

Solo haría falta elaborar un programa (software) en el cual se describa el proceso que deseamos llevar a cabo e ingresarlo en el sistema físico (hardware) de la maquina mediante el sistema de almacenamiento de información, y cada que deseemos implementar una nueva tarea, simplemente hay que reemplazar esta descripción por una nueva o bien añadirla dentro de la memoria sin necesidad de cambiar la posición de los mecanismos físicos de la máquina. Esta peculiaridad de la máquina se conoce como la universalidad de las máquinas de Turing. Tal caracterización de lo mental muestra, además, cómo el funcionalismo de Turing conserva parte del dualismo cartesiano, a su vez que muestra desacuerdo con el conductismo y la teoría de la identidad, respecto a la idea de que la naturaleza mental no es producto de meras reacciones de interacciones fiscas. Pues aquí tales reacciones deben estar reguladas o determinadas con base a una capacidad funcional, en otras palabras, para que una tarea pueda ocurrir o darse satisfactoriamente se deben llevar a cabo ciertas acciones en concreto bajo determinadas circunstancias y finalidades. Así, la interacción durante la realización de una tarea entre software y hardware sería similar a tener una guía de ensamblaje, dónde la guía es software y nosotros somos el mecanismo físico que lee y realiza el proceso de ensamblaje.

Turing describe el funcionamiento de la mente de su máquina como una “tabla de instrucciones” en la cual están contenidas instrucciones o determinaciones de las acciones del cuerpo que indican la manera o el proceder que deben tener los mecanismos del mismo para llevar a cabo una tarea. La tabla de instrucciones opera bajo un lenguaje de máquina,

esto es, que toda la información que maneja está codificada en un lenguaje que facilita y agiliza el manejo de la información. Normalmente se contempla el uso de un lenguaje lógico-matemático como el sistema binario, no obstante, podemos imaginar cualquier sistema de codificación de la información como podría ser el braille, el código morse o el sistema de codificación numérica de 10 dígitos que anteriormente se ejemplifico. Ahora bien, el modo en que opera la tabla de instrucciones es mediante el desarrollo de estados mentales. Tanto los funcionalistas como Turing entienden los estados mentales de la máquina como el producto de la interacción, relación o procesamiento causal de lo que sería la información de las entradas y salidas, el programa o las instrucciones de la mente y otros estados mentales, a saber, la información o los resultados que se han obtenido de otras operaciones internas y que, por ende, está almacenada en la memoria (Block, 2007: p.17).

Para entender cómo se da esa operación mental, la relación causal de información, ésta se representa esquemáticamente dentro de una tabla o cuadrícula de operaciones, donde cada tabla se distinguirá por tener un numero finito de estados (información finita que puede manipular la tabla), entradas que puede recibir y salidas que puede dar. Ned Block dice que una tabla de máquina tiene dos funciones que la especifican (Block, 1980: p.173). Sin embargo, me parece más prudente decir que son dos maneras de determinar lo que se realizara con una cierta entrada (o dos tipos de relaciones causales), según sea la información que manipula la tabla con base en la función (tarea) para la que está programada y debe desempeñar. Esos tipos de relaciones causales son: 1) una desde entradas y estados a salidas, estas relaciones son a lo que nos referimos cuando decimos que la máquina puede emitir una cierta salida externa a la entrada que recibe; y 2) una desde entradas y estados a estados, esto es, que se genera un estado propio o interno al sistema, ya sea que tenga que ir a otro estado (paquete de información, operación o generar otro resultado) o modificar algún otro contenido tanto de la memoria como de la tabla de máquina. Además, cada tipo de relación causal estará provisto de un equipo o conjunto de condiciones (reglas) para cada combinación de estados y entradas que puede presentar la máquina, esto según sea la relación causal para la que está programada la máquina. El conjunto de condiciones son las instrucciones o reglas contenidas en la tabla de máquina.

Un ejemplo muy sencillo de una máquina de Turing que da Block sería una máquina de refrescos con dos estados o contenidos de la memoria S1: desear diez pesos, y S2: desear cinco pesos. También tiene dos entradas posibles E1: entrada de cinco pesos y E2: entrada de diez pesos. Así cuando la máquina está en S1 y tiene una entrada E1 el conjunto de órdenes que se provee es: "no emitir una salida e ir a S2"; aquí el tipo de relación causal que se llevó a cabo es de tipo 2. Si se está en S1 y se tiene E2 se da "emitir un refresco y permanece en S1" ésta es una relación causal de tipo 1. Si se está en S2 y se tiene E1 se da "emitir un refresco e ir a S1"; si se está en S2 y se tiene E2 se da "emitir un refresco y cinco pesos, después ir a S1". Las dos últimas combinaciones que se llevaron a cabo presentaron relaciones causales tanto de tipo 1 como 2 (Block, 2007: p. 30).

La forma esquemática de la tabla de instrucciones se da de la siguiente manera:

	S1: desear diez pesos	S2: desear cinco pesos
E1: Cinco pesos	No emitir una salida e ir a S2	Emitir un refresco e ir a S1
E2: Diez pesos	Emitir un refresco y permanece en S1	Emitir un refresco y cinco pesos, después ir a S1

Esta manera esquemática muestra cómo opera la mente de la máquina en lenguaje natural, puesto que en lenguaje de máquina esto estaría codificado. De tal manera aquí el modo de operar, resolver problemas o la capacidad de realizar tareas mentales se muestra por el esquema de tabla mientras que anteriormente en el hardware la realización física de esta operación mental se veía a manera de una lista de órdenes que se leía y ejecutaba. El esquema de tabla, además, muestra como Turing idealmente concibe a sus máquinas como "máquinas de estado discreto", esto es que sus máquinas tienen un grupo de operaciones finitas que permiten individualizar o distinguir cuando se está llevando a cabo un cierto estado y cuando ha pasado a otro, pues su realización está caracterizada y delimitada por la relación de un conjunto de información. Esto no ocurre con las máquinas de estado continuo, donde las condiciones que deberían determinar a los estados resultan tan amplias, variables o relacionadas entre sí que resulta difícil distinguir cuando se ha terminado una operación y cuando ha iniciado otra (Turing, 1950a: p.10-11).

2.4 Otros argumentos funcionalistas

Con los teóricos de la identidad y los conductistas se decía que la causalidad implicada en una acción era de índole meramente mecánica, empírica, objetiva o bien de primer orden. Por ello, al momento de describir la realización de lo que para ellos serían las acciones mentales, ellos no usan términos referentes a una substancia mental incuantificable (subjetivo). Sino que utilizan términos empíricos que describen esas acciones en términos de sucesos físicos.

Los funcionalistas argumentarán contra esa subjetividad mental a la que apelan los teóricos de la identidad y los conductistas, al decir que la máquina de Turing puede describir sucesos mentales omitiendo el uso de términos mentales, sin negar la presencia o influencia de los estados mentales, y conservando una objetividad de tipo lógica de primer orden. Esto al utilizar un lenguaje más abstracto como el lógico-matemático o una codificación, pues dicha simbolización aludirá de manera formal a la cuantificación objetiva de la realización de los estados mentales. Block da un ejemplo en el cual muestra cómo se puede hacer el cambio de un enunciado sobre estados mentales en lenguaje natural a un leguaje formal, donde describe un cierto rol causal de la función de dolor. Aquí, la función de dolor la representa con el termino formal "*S*", empero dentro de todos los ejemplares de dolor (*S*), él describe el ejemplar de dolor que es provocado al sentarse en una tachuela y al cual se le representa con "*x*" (Block, 2007: p. 17-18). Entonces, el enunciado sobre ese dolor en lenguaje natural seria:

1) Tener dolor = Sentarse en una tachuela, y esto causa ansiedad y decir "auch"

Este primer enunciado está dicho utilizando términos mentales como "dolor" y "ansiedad". La ocurrencia del rol causal de dolor *x* involucraría otros dos estados mentales (o propiedades funcionales) individuales, a saber, "sentarse en las tachuelas" (*P*) el cual provocaría el segundo estado "tener ansiedad" (*Q*) y a su vez emitir "auch". Así la relación causal de estados que produce este ejemplar (*x*) de dolor (*S*) puede describirse como:

2) Tener un *S* = Tener una *x* tal que $\exists P \exists Q$ [Sentarse en una tachuela (tener un ejemplar *Sx*) causa *P*, y *P* causa tanto *Q* como emitir "auch", & *x* está en *P*]

Aquí vemos que los términos mentales han sido remplazados con símbolos. No obstante, este lenguaje abstracto permite hablar de un estado mental cualquiera en tanto que los símbolos

que los representan aluden a la realización de dichas propiedades funcionales, es decir, a su realización física. Así como, describe la causalidad mental al dar cuenta del proceso relacional que ocurre entre dichas propiedades funcionales y dar paso a la realización del rol causal *S*. De tal manera *S* es una función de segundo orden (mental) en cuanto que está constituida de otras propiedades funcionales (*P* y *Q*) cuyas realizaciones son cuantificables y están causalmente relacionadas una con otra. De esta manera, para los funcionalistas los términos mentales como "creo", "deseo", etcétera implican un proceso causal de estados que es propio al rol causal que se realiza, así como será propio de la programación de la máquina en cuestión o, si se quiere, de la mentalidad que utiliza dichos términos.

2.5 Juego de la imitación o test de Turing

Siendo ese el funcionamiento de las máquinas que propone Turing, él advierte que el potencial de estas máquinas puede ser tan amplio y capaz como el de cualquier humano, por lo cual bajo condiciones de prueba plausibles considera que una máquina bien programada y con las características físicas pertinentes podría desempeñarse tan bien como un humano en cualquier tarea. Por máquina bien programada se está entendiendo el que tenga un libro de reglas o tabla de máquina con el conjunto de instrucciones requerido para determinar la realización de la tarea en cuestión. Por supuesto entre más compleja sea la tarea a realizar o si un mismo software contiene más de un grupo de instrucciones para realizar más de una tarea, el hardware tendrá que poder cumplir con las condiciones para soportar la cantidad de información que se debe manejar.

Con condiciones de prueba plausibles me estoy refiriendo a la exigencia con que se prueba a la máquina, así como las expectativas que se esperan de sus resultados, debe ser la misma que se haría y tendría con un humano. Por tanto, si de un humano se pueden esperar ciertas equivocaciones, lentitudes y debilidades de desempeño, de igual manera la máquina puede presentar tales dificultades. Por ejemplo, si bien las máquinas pueden presentar un dominio casi perfecto en ciertas áreas como cálculos matemáticos, en otras áreas puede ser que muestren mayor torpeza o deficiencia como al componer un poema. Estas debilidades de acción son válidas no solo para la máquina sino también para los humanos, pues a algunos humanos se les dan mejores las matemáticas que a otros, o bien, unos pueden ser desastrosos escribiendo poemas y otros ser unos genios en el área (Turing, 1950a: p. 15-20).

Así Turing propone una metodología para poner a prueba sus máquinas. Esta prueba o test de Turing lo llama "juego de la imitación": básicamente consiste en un experimento que consta de dos habitaciones entre las cuales se puede mantener comunicación, pero solo mediante la escritura a máquina. Es decir, solo se pueden recibir y enviar textos entre una habitación y la otra, ya sea en un papel mecanografiado, mediante un fax o en un monitor. Inicialmente en una de las habitaciones hay una mujer (A) y un hombre (B), en la otra habitación hay otra persona que es el interrogador (C). El interrogador solo conoce a los individuos de la otra habitación como A y B, su tarea es descubrir quién es el hombre y quien la mujer, por lo cual él puede realizarles cualquier tipo de pregunta que se le ocurra para determinar quién es el hombre y quién es la mujer. Por ejemplo, puede preguntar que ropa traen puesta, sus gustos, hábitos o cualquier cosa.

La tarea tanto de la mujer como del hombre en la otra habitación es persuadir al interrogador de ser la mujer de esa habitación, es decir, A simplemente tiene que responder con sinceridad a las preguntas de C. Imagínese las respuestas típicas como tengo pelo largo, llevo tacones y vestido, uso maquillaje e incluso puede mencionar las marcas de su ropa y accesorios para convencer al interrogador, decir que le gustan las películas románticas, etcétera. A puede optar por añadir cualquier comentario a sus respuestas que le pudieran servir de justificación y persuasión para que C no se equivoque al determinar que ella es la mujer. En tanto que B debe dificultar la tarea de C, por lo cual B buscará mentir para convencer al interrogador que él es la mujer en esa habitación. De tal forma al igual que ocurrió con A, B debe mostrar conocimientos o parámetros de comportamiento y habilidades característicos a los de una mujer, en consecuencia, B debe poder ajustarse y llevar a cabo dichos parámetros para poder convencer a C de ser la mujer de la habitación. Así, él podría dar todas las respuestas típicamente esperadas de una mujer y aseverar que él es la mujer o incluso si es ingenioso puede también dar respuestas que, al igual que A, justifique su falta de conocimientos femeninos con alguna otra excusa.

Turing señala que esta metodología del juego de imitación traza una clara separación entre capacidades físicas y capacidades intelectuales (Turing, 1950: 5), pues claramente limita toda interacción física directa entre los participantes de las respectivas habitaciones. Por lo cual, la interacción depende solo de su capacidad en el dominio tanto del lenguaje y la escritura

como de los temas que el interrogador les plantee para determinar su género. Además, el interrogador no puede siquiera cerciorarse de la manera o las condiciones en que las respuestas a sus preguntas son elaboradas, su única herramienta son las palabras escritas a máquina que recibe.

Esta dinámica cambia un poco cuando Turing propone que una máquina tome el lugar de B en el juego, ahora A y B deben persuadir al interrogador de que son humanos y el interrogador debe determinar quién es humano y quien máquina. En el caso anterior se asumía que el interrogador tiene criterios para determinar el uso típico de las palabras y temas que una mujer y hombre utilizarían; en este caso, el interrogador debe tener criterios para determinar el uso normal o típico de un humano y una máquina. Por otro lado, ahora B (la máquina) debe contener una tabla de instrucciones que le permita imitar satisfactoriamente el dominio del lenguaje utilizado por los humanos, así como le debe permitir ser capaz de imitar las respuestas típicas que podría dar un humano a ciertas preguntas (entradas) e incluso que la dote de poder justificarse o excusarse ante cualquier falta del domino de ciertas habilidades humanas o ante su capacidad superior a las humanas. Por ejemplo, si C le pide a B realizar una cierta operación matemática, B muy probablemente sería capaz de resolver el problema en fracciones de segundo. Ante esto su tabla de máquina debe poder determinar que emitir una salida tan presurosa va en contra de las habilidades de un humano promedio por lo cual bien podría emitir una justificación como "soy un genio en matemáticas" o si está mejor programada la tabla de máquina muy probablemente emitiría la orden interna "esperar 10 minutos antes de emitir la respuesta", todo esto con tal de convencer a C de que es el humano al interior de la habitación.

De tal manera para Turing resulta completamente plausible que una máquina sea capaz de imitar cualquier conducta y tarea intelectual humana, pues la manera en que está elaborado y en que opera la tabla de máquina (software o programa) le da la posibilidad de imitar la elaboración de cualquier tarea que se le proponga o programe. Pues recordemos que para programar una tabla de máquina solo sería cuestión de que en ella se contenga la descripción de cómo se lleva a cabo una cierta tarea. Por supuesto, entre más detallada sea la descripción o entre más posibles variables (entradas, salidas y otros estados mentales) involucradas se

tomen en cuenta para la programación de la tabla respecto a la elaboración de la tarea, el software será capaz de realizar la tarea con mucho mayor precisión.

Ante las características con que opera su máquina, para Turing, resulta claro que es posible construir una máquina capaz de mostrar un desempeño satisfactorio en el juego de imitación; para él, solo sería cuestión de dotarla con una programación adecuada que tome en cuenta las diversas variables que podrían surgir en el juego. De la misma manera en que ocurriría con cualquier persona que se prepara para un examen, a la máquina también se le debe haber provisto de la información requerida para su prueba, así como dotarla con las reglas pertinentes e instruirla en la manera en que debe emplear dicha información. Todo esto ocurre al momento de programar la máquina. Cabe dejar dicho a manera de reflexión que el hecho de que Turing perciba las capacidades mentales como describibles en términos de instrucciones nos indica como él concibe los estados mentales a manera de procesos silogísticos, donde uno constituye y recrea mediante proposiciones (premisas y conclusión) una descripción de relaciones causales características de acontecimientos involucrados en el acaecimiento de acciones teóricas y prácticas.

Así es como se está contemplando que las máquinas de Turing son capaces de tener pensamiento pues aquí la operación que realizan estas máquinas no solo describe la manera en que funciona el pensamiento, es decir, no es solo una herramienta que ayuda a entender cómo funciona el pensamiento humano, sino que de hecho eso es propiamente pensamiento. En este sentido la distinción que guardan con la postura dualista cartesiana consiste, por un lado, en que Turing fue capaz de concebir y plantear un sistema que cumpliera con las exigencias dualistas sobre las capacidades que caracterizan el desempeño propio de la mente, pero que además tiene la peculiaridad de poder ser implementada en diversas constituciones físicas. Lo cual es un punto de divergencia con Descartes. Por otro lado, Turing consigue plantear una noción de pensamiento presuntamente más objetiva en el sentido en que se concibe al pensamiento como un sistema de premisas (instrucciones) elaboradas con mira a dar lugar a una cierta causa, fin o acción bien definida y, además, hay que decir son sistemas que se pueden reconstruir o a los que se puede acceder y son asequibles para todo ser con mentalidad, dado que está constituido con base en la capacidad para realizar una función. Descartes plantea una idea de pensamiento más subjetiva al dar mayor peso a la idea de que

la realización de acciones se ve motivada y realizada con base en criterios privados y propios a cada mentalidad como son las voliciones, deseos, entendimientos o juicios. Descartes no apunta tanto al hecho de cómo realizar una acción sino a lo que motiva la acción.

2.6 Conclusión

La postura dualista cartesiana encuentra en las ideas de Turing y los funcionalistas un sucesor capaz de plantear y sacar adelante su postura a reserva de modificar su concepción respecto al tipo de seres que pueden contener mentalidad, lo cual se consiguió al llevar el dualismo al ámbito del desarrollo tecnológico, dónde se consigue establecer una analogía entre la dualidad mente-cuerpo con la dualidad software-hardware. Ahí se advierte que el hardware puede tener una constitución física de materiales y disposiciones variables pero que debe cumplir con un soporte de mecanismos operativos básicos para correr el software (mente), como tener un sistema de almacenamiento de información, una unidad ejecutiva capaz de modificar, manipular y procesar todo tipo de información, y un sistema de control u orden que traduce y regula la información entre la tabla de máquina (las órdenes del software) y el resto de los mecanismos del hardware, a saber, la unidad ejecutiva y la memoria con la información de operaciones internas o distinta de la información contenida en la tabla de máquina.

Asimismo, esta analogía lleva a modificar la concepción de la mente al proponer un funcionamiento, de la misma, dónde se sistematizan sus operaciones al concebirlas como un conjunto de reglas o instrucciones ordenadas y diseñadas para llevar a cabo una tarea específica. Pues al entrar en interacción con otro tipo de información, a saber, las entradas, salidas y otros datos de la memoria estas influyen determinantemente sobre el tipo de acciones que se producirán. De igual manera, se le está revocando el carácter privado o individual de la mente al plantear que dicho sistema está elaborado de tal manera que cualquier mentalidad sería capaz de corroborarlo, puesto que los eventos mentales de la máquina son descripciones meramente funcionales u objetivas de cómo llevar a cabo una tarea. Y no una descripción subjetiva, privada a un individuo. De tal manera se consigue postular un tipo de seres físicamente muy distintos de los humanos pero que pueden tener un domino satisfactorio de capacidades consideradas propias del pensamiento humano como es el dominio del lenguaje y que se ajusta a la concepción dualista mente-cuerpo.

CAPITULO III

SEARLE Y EL EXPERIMENTO DE LA HABITACIÓN CHINA

3.1 Introducción

El planteamiento de Turing es un tema que llamó la atención de John Searle, filósofo de la mente, quien cuestiona si realmente el funcionamiento del pensamiento humano puede ser entendido de la manera en que Turing y los funcionalistas describen la operación de las computadoras como pensamiento. Searle considerara que una descripción como esa es bastante imprecisa o más bien que el funcionalismo no está contemplando, e incluso está desechando, una característica crucial del pensamiento humano que a su parecer es lo que hace al pensamiento humano ser pensamiento. Al no estar de acuerdo en que las máquinas descritas por los funcionalistas pudieran tener la capacidad de pensar, Searle propondrá una postura en oposición a dicha idea del pensamiento, la cual, a diferencia del funcionalismo dará prioridad al carácter semántico del pensamiento, esto es: la idea de que el pensamiento humano tiene intencionalidad o es sobre algo.

3.2 Pensamiento en las máquinas de Turing

Lo primero a tener en cuenta respecto a la postura de Searle es una aclaración que él mismo hace sobre cuál es la postura a la que se está oponiendo, pues advierte que hay dos versiones sobre la inteligencia artificial (IA). Por un lado, está la versión débil de la IA que ve a los modelos computacionales como herramientas útiles en el estudio de la mente, es decir, aquí las máquinas no se perciben propiamente como si tuvieran la capacidad de pensar, más bien, son algo que puede ser usado para un estudio más preciso de dicha capacidad, de la misma manera en que éstas resultan útiles para el estudio de otras áreas como la economía o la predicción del clima. Por otro lado, está la versión fuerte de la inteligencia artificial, la cual defiende la postura de que las máquinas no son simples herramientas para el estudio de la mente, sino que, ellas mismas son mentes en tanto que simulan al pensamiento. Esta postura sostiene que las máquinas que operan como computadoras digitales, por el hecho de operar según un cierto programa, son capaces de tener estados cognitivos y realizan tareas o funciones como las que presentan los humanos. La idea es que las computadoras son capaces de actuar de tal manera que un experto no pueda distinguir entre su actuación y la de un

humano. Para la IA fuerte, ésta habilidad sería razón suficiente para considerar a dichas computadoras como poseedoras de las mismas habilidades cognitivas que tiene el humano.

> La meta es diseñar un programa que simule la cognición humana de tal manera que pase el test de Turing[1] (Searle, 1990: 26-27).

Respecto a la versión débil de la inteligencia artificial, Searle no parece tener inconveniente en aceptar que las máquinas pueden ser útiles en el estudio de la mente, empero estará en contra de la consideración propia de la IA fuerte de que las acciones realizadas por las máquinas que proponen son pensamiento. De esta manera, Searle también se estará oponiendo a Turing quien defiende la IA fuerte, pues para Searle no es concebible que el pensamiento se entienda de tal forma. El problema con la versión fuerte es que aquello que hacen las máquinas no parece implicar ni ser necesariamente pensamiento, en tanto, dicha caracterización únicamente contempla e implica un cierto manejo de la forma de los símbolos.

Cabe aclarar que por *programa* nos estamos refiriendo a un grupo específico de reglas o instrucciones que son las que caracterizan el funcionamiento de una máquina o determinan la vía de acción que ésta debe llevar a cabo. Por decirlo así, tanto el grupo de reglas como los símbolos en que están expresadas son lo que hace que la máquina sea lo que es: son parte esencial de la máquina. Siendo así, cuando hablamos del pensamiento de la máquina nos estamos refiriendo únicamente a la descripción del funcionamiento que presenta su programa según los propósitos para los cuales haya sido diseñado. Además, las instrucciones son construidas por programadores a manera de instructivos concernientes o pertinentes a la elaboración de una cierta función mediante el uso de los símbolos. Así, la IA fuerte considera que mediante un cierto grupo de instrucciones se puede operar o instruir correctamente la manipulación de símbolos, de tal manera que mientras dicha manipulación sea capaz de generar las salidas correctas a las entradas percibidas: esto sería garantía suficiente de que la máquina es capaz de elaborar tareas o funciones de la misma manera en que normalmente un individuo con mentalidad podría llevarlas a cabo.

[1] Traducción propia

La IA fuerte está entendiendo a la máquina como algo equivalente a la mente, donde, todo el sistema funcional que presenta el programa vendría a ser la manera en que operan las mentes y, dado que éste es un funcionamiento que únicamente consta de ciertas instrucciones para poder llevar a cabo cierta tarea, se estaría señalando al programa o a este funcionamiento como algo que solo se determina dentro de un ámbito lógico abstracto. Por ello, si bien seria cierto que podríamos acceder a las instrucciones del programa y corroborarlas al ponerlas a prueba, su elaboración no deja de darse de manera puramente formal. Además, también señalan que el programa es independiente de cualquier realización física de la máquina.

3.2.1 La habitación China

No obstante, para Searle la propuesta de Turing y la IA fuerte ha dejado de lado un aspecto muy importante del pensamiento humano que ya se había advertido desde Descartes. Pues tanto a consideración de Descartes como de Searle las acciones de la máquina o la ejecución de las instrucciones que permiten ese manejo de figuras, requieren o deberían implicar algo más que solo una cierta manipulación de símbolos, a saber, dice Descartes "nunca podrían hacer uso de palabras ni otros signos, componiéndolos, como hacemos nosotros, para declarar nuestros pensamientos,… no se concibe que ordene en varios modos las palabras para contestar al sentido de todo lo que en su presencia se diga" (Descartes, 2002: p. 114). Searle busca dar un paso más al identificar esas faltas como una carencia de intencionalidad en las acciones y estados mentales de la máquina.

A manera de objeción, Searle postula el caso de la habitación China, donde explica e intenta mostrar porqué las máquinas de Turing exhiben una caracterización del pensamiento que difiere lo suficiente respecto a lo que entendemos como pensamiento humano. Pues dice que dicho funcionamiento de las máquinas a pesar de lograr llevar a cabo funciones aparentemente exclusivas del pensamiento humano, parecen quedar cortas o no logran emular de manera exacta las características esenciales de la mentalidad humana. Por ello, Searle cree que no deberíamos considerar el funcionamiento que presentan tales máquinas como pensamiento propiamente, no al menos en el mismo sentido en que consideramos al pensamiento humano como pensamiento.

En sus textos *Minds, brains, and programs* y *Is the Brain's a Computer program?* Searle plantea el argumento de la habitación China como un experimento mental en contra de la

postura sostenida por la versión fuerte de la IA. Este experimento consiste en plantear una situación donde un humano está encerrado dentro de una habitación; fuera de ella, hay un grupo de personas quienes proveen a él con dos pilas de papel distintas, las cuales llevan impresos símbolos chinos, y un libro de reglas que le permitirá relacionar los símbolos de cada pila en una secuencia determinada. Además, en este experimento se hace la suposición de que la persona no conoce absolutamente nada sobre el idioma chino, por lo tanto, no entiende ni sabe cuáles son los significados de los símbolos en las pilas de papel que se le han proporcionado. Para el individuo, tales símbolos más bien serian garabatos o figuras, de ahí también la importancia de que el libro de reglas permita relacionar los símbolos tan solo por su figura. El libro de reglas deberá estar en el idioma nativo del individuo al interior de la habitación para que éste sea capaz de seguir y ejecutar las reglas del mismo. De esta manera, el individuo podrá relacionar los garabatos adecuadamente tan solo por su forma, y según lo indicado por las reglas.

Posteriormente, se le entregan al individuo al interior de la habitación otro conjunto de símbolos chinos. A partir de este conjunto, el individuo es capaz de llevar a cabo las respectivas indicaciones que le especifica el libro de reglas. Cabe recalcar que el individuo solo es capaz de emplear los símbolos chinos en tanto que es capaz de identificarlos por su forma y por la mención de dicha forma en el grupo de instrucciones que le indican que hacer o como mover u ordenar tales figuras para crear o estructurar otras figuras que deberá ser devueltas al exterior. La tarea del humano solo reside en su capacidad de leer y ejecutar las instrucciones del programa.

Tenemos, entonces, una situación en la que se está ejemplificando la manera en que operan las máquinas de Turing pues el libro de reglas de la habitación vendría siendo el equivalente a el programa computacional o la tabla de instrucciones de las máquinas de Turing, las pilas de papel con símbolos chinos serían la base de datos o la memoria, las personas fuera de la habitación vendrían a ser los programadores, el conjunto de símbolos entregados por los programadores son las preguntas que se le hacen a la computadora o bien las entradas de la máquina mientras que el conjunto de símbolos devueltos por el individuo son las salidas de la máquina. Por último, Searle señala al humano como el equivalente a la computadora. “The people who wrote it are programmers, and I am the computer” (Searle, 1990: p.26). Empero

a mi parecer es más preciso decir que el humano vendría a ser el equivalente a dos partes del hardware de las máquinas de Turing, a saber, es tanto el control como la unidad ejecutiva en las máquinas de Turing[2], donde su dominio del lenguaje en que está escrito el libro de reglas y su capacidad para identificar la forma de los símbolos chinos le permitiría ser el equivalente al control en la máquina de Turing, mientas que su capacidad para realizar dichas ordenes sería el equivalente a la unidad ejecutiva de la máquina de Turing.

Al igual que con las máquinas de Turing, el libro está diseñado para instruir la manipulación de información de tal manera que se generen ciertas respuestas según las preguntas que se hagan o las entradas que se reciban. Esta manipulación se da con la finalidad de emular las respuestas que daría a esas preguntas un humano nativo hablante del chino. Por ello, se plantea que las respuestas tanto de las máquinas de Turing como las elaboradas en la habitación china resultan ser indistinguibles de las que daría cualquier humano nativo hablante del chino: nadie al ver sólo las respuestas obtenidas sin su proceso de elaboración se atrevería a decir que tales respuestas no son igualmente buenas o propias a las que daría un humano nativo hablante del chino. Se está implementando un programa computacional mediante la ejecución de sucesos que realiza una persona según lo establecido por el libro de reglas, pues se reciben ciertas entradas y se dan ciertas salidas para lograr pasar el test de Turing.

> Me parece bastante ovio en el ejemplo que yo no entiendo ni una palabra de las historias chinas. Tengo entradas y salidas que son indistinguibles de los nativo hablantes del Chino, y puedo tener cualquier programa formal que gustes, pero sigo sin entender nada[3] (Searle, 1980: p. 418).

Sin embargo, con este experimento mental Searle quiere señalar que esta manipulación de la información no conlleva o implica una capacidad de pensamiento como la humana. Probablemente por ello a propósito introduce a un humano en el propio ejemplo pues dirá que el individuo dentro de la habitación, a pesar de tener pensamiento, ser capaz de cumplir con el test de Turing, así como de leer, comprender y él mismo ejecutar lo que el programa dice, aun así no está logrando emplear ese pensamiento en la actividad que está realizando,

[2] Véase el capítulo II: Máquinas de Turing p.41-45 del presente texto
[3] Traducción propia

ya que no está adquiriendo ningún entendimiento del lenguaje chino. Por ende, tampoco logra advertir las ideas que normalmente un nativo hablante del chino asociaría a los símbolos chinos o bien no sabe el significando de los símbolos, ni el propósito de sus acciones. Así, Searle busca llamar la atención sobre ciertas características que para él son parte importante del pensamiento y que en el funcionalismo parece omitirse por completo, a saber: que el pensamiento también tiene que ver con la comprensión o entendimiento de lo que se hace, y no solo con una cierta coherencia de acciones sucesivas, pues la comprensión es mucho más complejo ya que también requiere de creencias, propósitos, intenciones y significados.

Por la manera en que se le está presentando la información al individuo en la habitación, tal no es suficiente para que él logré asociar ideas o cosas a los garabatos de tal manera que vea esas formas como símbolos, es decir, como una representación de ciertas ideas o cosas. Ésta carencia asociativa impide que el individuo aprenda el idioma chino pues, aunque la manipulación formal de símbolos le podría dar alguna noción de dicho lenguaje, a saber, le permite conocer las formas de los símbolos del idioma chino, mediante esa actividad no es capaz de adquirir la semántica del lenguaje. Aquí los símbolos resultan ser algo así como un objeto inerte para el pensamiento, son sólo figuras cuya relevancia no va más allá de su apariencia, pues la manera en que está elaborado el libro de reglas así los establece, pues el libro te dice que hacer con los símbolos según su forma.

Con esto Searle intenta mostrar que, si un humano en estas condiciones no es capaz de emplear su pensamiento respecto de la actividad que realiza como para adquirir el idioma chino, es decir, no es capaz de asociar ideas o cosas a los garabatos o a lo que ésta realizando entonces una máquina bajo estas mismas condiciones tampoco está empleando pensamiento alguno. En estas circunstancias tanto la máquina como el humano no son capaces de entender significativamente lo que hacen, no generan pensamiento respecto de sus procesos, sólo estarían realizando una manipulación de símbolos. No obstante, me gustaría hacer la observación de que, a mi parecer, según la descripción de Turing respecto de sus máquinas el humano de la habitación en realidad está fungiendo como lo que sería el control y la unidad ejecutiva de la máquina de Turing, el humano no es como tal el programa. Esto es importante porque la IA fuerte precisamente apela a que el programa como tal es pensamiento, mas no

la memoria, el control ni la unidad ejecutiva de la máquina que vendrían a ser parte del hardware de la máquina.

Tal vez por ello sea también importante recalcar que el humano es capaz de leer el programa, de la misma manera que lo haría el control de la máquina de Turing, empero como humano si al advertir y entender lo que sería el programa de la máquina, aun así, no es capaz de aprender la semántica del idioma chino entonces cómo sería posible que la máquina tuviese algún entendimiento de lo que está haciendo. El problema es que el libro de reglas en realidad no explica nada, no tiene ningún contenido significativo de los símbolos chinos, simplemente da órdenes, pues como está elaborado a manera de un instructivo de ensamblaje. Por ello para Searle el programa en comparación con el pensamiento humano es mucho más pobre respecto a su contenido. En este sentido, se dirá que los símbolos chinos mencionados por el libro de reglas, las pilas de papel y las respuestas carecen de significado. Esta carencia es consecuencia de la caracterización funcionalista de las máquinas, pues tales reglas solo tienen la finalidad de ordenar un cierto grupo de figuras para dar lugar a otro tipo de figura.

> Los estados mentales como las creencias no se definen por ninguna característica intrínseca sino por sus relaciones causales, y estas constituyen su función. Las creencias, por ejemplo, son causadas por percepciones, y junto con los deseos causan acciones. Las relaciones causales son el único contenido del hecho de tener una creencia (Searle, 2006: p. 86-87).

Un ejemplo de las reglas del libro podría ser: si recibió la figura 56, tome dos figuras del tipo 5d de la primera fila y colóquelos frente a las figuras 8d y 8f de la segunda fila formando una hilera que empiece con 8f, y colóquelos en la salida. Suponiendo que los números y los símbolos alfanuméricos del ejemplo anterior son símbolos chinos, vemos que este tipo de regla solo requiere que el individuo identifique la figura, forma o apariencia del símbolo para poder ejecutar la regla. Fuera de ello no requiere ni dice nada más respecto de los símbolos mismos, es decir, no señala ningún significado, por ende, no implican o representan algo. A lo mucho, este tipo de reglas permitiría al individuo advertir ciertos patrones de relación o maneras de composición entre las figuras, pero sin una noción clara de porque o a que se debe dicha relación, puesto que ésta es simplemente un precepto, más no una relación que permitiría un entendimiento de su composición conforme a una asociación de las figuras respecto de objetos, ideas o eventos.

Así, las reglas no parecen contener ningún conocimiento que pudiera considerarse como algo característico o propio de los garabatos o figuras, pues más bien, las propias reglas son las que estarían dando o imponiendo el tipo de relación que tendrían entre sí los garabatos, que bien podrían considerarse arbitrarias y ajenas a ellos en tanto que no se advierte una explicación o caracterización que permitiera identificar tal relación como propia o relativa a tales figuras. Vemos entonces que para las máquinas tanto las reglas como los símbolos que emplea no tienen sentido propio. Esto no ocurre en el pensamiento humano pues lo que normalmente calificaríamos como pensamiento implicaría que las figuras hicieran referencia a algo y que su ordenamiento se dé conforme a tales referencias y no solo conforme a su forma.

Por ejemplo, supóngase que no sabemos nada de inglés, pero sí español y nos hemos topado con las palabras “esperanza” y “hope” en una pared grafiteada. Cuando vemos la palabra “esperanza”, nosotros asociamos un cierto significado a esta palabra, una cierta idea; empero, esto no ocurriría al ver la palabra “hope”. Pues, en el caso de “esperanza” nuestro conocimiento del español nos permite identificar las implicaciones o asociaciones que dicho símbolo tendría dentro del idioma, por el contrario, no saber inglés provoca que la palabra “hope” no sea significativa para nosotros. Lo mismo ocurriría con cualquier idioma que desconozcamos, y si consideramos la posibilidad de que tales idiomas usen símbolos totalmente extraños nunca los reconoceríamos; ni siquiera diríamos que son símbolos.

También podemos imaginar el caso en que, por alguna razón, se decide oficialmente dentro del idioma español cambiar la palabra “esperanza” por alguna abreviación o algún otro símbolo. ¿Qué ocurre con la palabra ‘esperanza’?, Simplemente pierde el significado que tenía. Además, la decisión sobre cuál será el símbolo que lo reemplace podría ser tomada arbitrariamente, ya que un significado puede ser representado de innumerables formas. Es decir, si bien una forma cualquiera puede ser un símbolo de un lenguaje, por si no implica un significado; más bien, todo símbolo lo es en tanto que es adoptado por un lenguaje o forma parte de un sistema lingüístico.

Asimismo, podríamos imaginar el caso donde dentro de un mismo lenguaje se decide y acepta intercambiar los símbolos correspondientes a dos significados distintos. Esto sería, por ejemplo, intercambiar el significado atribuido a la palabra “verde” por el significado

atribuido a la palabra “ave”, de tal manera que ahora el símbolo o la forma representativa ‘verde’ tendrá el significado al que refería “ave”, mientras que ‘ave’ ahora refiere al significado que antes tenía la palabra “verde”. Lo que se muestra con estos ejemplos es que, dado el manejo de símbolos que lleva a cabo la máquina, se puede advertir que no requiere, que tales símbolos tengan el mismo contenido significativo que pudieran tener en otras circunstancias. Para el programa de la máquina no es necesario que los garabatos tengan un significado para ejecutar sus reglas, éstas solo requieren que tal o cual garabato sea movido u ordenado de cierta manera y para ello basta con reconocerlos por su forma.

A mi parecer aun si el que manipula los símbolos entendiera el chino su entendimiento no tendría ninguna influencia dentro del funcionamiento y realización de las tareas de la máquina, siempre y cuando se apegue a realizar las acciones que le indica el libro de reglas. De esta manera, cualquier entendimiento que pueda tener el humano es ajeno al supuesto pensamiento de la máquina. Tal vez incluso se podría decir que ese supuesto entendimiento del humano en la habitación china no debería ser considerado pensamiento sino simplemente una actividad o reacción extraña del cuerpo de la máquina que no tiene nada que ver con la actividad mental de la misma.

Sin embargo, entiendo que para Searle el individuo al interior de la habitación china sí es parte del software de la máquina. Probablemente por ello en su ejemplo es importante que el individuo no sepa el idioma chino, pues de lo contrario si el humano entiende o genera entendimiento de sus acciones al interior de la habitación se tendría que aceptar que la máquina tiene pensamiento. Pero al carecer de este conocimiento y estar limitado a solo ejecutar las reglas del libro, tanto la máquina como el individuo dentro de la habitación no están generando pensamiento al llevar a cabo tales acciones. Tampoco importaría cuán complejo pudiera ser el sistema empleado para ordenar los símbolos, ya sea que se usen programas computacionales muy sofisticados, libros de reglas muy bien elaborados, una persona o muchas personas, el producto será solo una consecuencia de manipulación de símbolos sin significado.

3.2.2 Intencionalidad

La cuestión que nos ocupará ahora es porqué resulta tan perjudicial para Searle que las máquinas tengan estas carencias en sus productos, es decir, porqué la ocurrencia o no ocurrencia de significado bajo cierto curso de acción permite o determina que algo sea o no pensamiento. Si burdamente advertimos la manera en que se desarrolla en nosotros, los humanos, el pensamiento, podemos darnos cuenta que los eventos del pensamiento humano son significativos o mantiene una relación con el mundo. Éste es el punto de divergencia que propone Searle entre el funcionamiento de la máquina y el funcionamiento del pensamiento humano, y lo ve suficientemente influyente como para no considerar que las acciones llevadas a cabo por la máquina debieran ser propiamente pensamiento. Pues, cosas como estados mentales, percepciones, entendimiento, creencias, deseos, opiniones, ideas en los humanos siempre conllevan cierto contenido, esto es: son sobre algo.

Esta peculiaridad del pensamiento humano Searle la identifica con la intencionalidad y dice que "la intencionalidad es aquella propiedad de muchos estados y eventos mentales en virtud de la cual estos se dirigen a, o son sobre o de, objetos y estados de cosas del mundo" (Searle, 1992: p. 17). Para ver cómo se da la relación de los estados mentales humanos con el mundo, Searle advierte que dicha relación debe ser planteada desde el ámbito lingüístico. Según Searle en el leguaje los estados mentales se presentan como actos de habla, y esto permite presentar o hacer visible las propiedades lógicas del pensamiento. Los actos de habla y los estados intencionales comparten la similitud de que "los estados intencionales representan objetos y estados de cosas en el mismo sentido de 'representar' en el que los actos de habla representan objetos y estados de cosas" (Searle, 1992: p. 20).

Cabe decir que desde este ámbito también se puede hacer un análisis lógico desde el punto de vista funcionalista al identificar las relaciones lógicas de los actos de habla con la finalidad de ver al pensamiento como una relación causal entre entradas, salidas y otros estados mentales, es decir, como una relación meramente funcional entre la mente y el mundo. Empero, aquí no se busca averiguar qué disposiciones y propiedades lógicas podrían tener los elementos de las proposiciones según una función, pues esto no sería distinto a explicar el funcionamiento de los programas de las máquinas. Aquí, lo relevante es mostrar que, si bien las proposiciones tienen propiedades lógicas, no obstante, sus elementos pueden ser

entendidos de dos maneras: 1) una conforme a una estructura lógica de funciones o causas (conforme a una sintaxis); 2) de a cuerdo a una estructura lógica que esté conforme a la manera en que estos son expresados en el leguaje, es decir, por el contenido significativo de su expresión lingüística.

Searle, además, hace uso de las explicaciones lingüísticas sobre los estados mentales para señalar que hay una intencionalidad intrínseca en ellos; el lenguaje permite advertir la ocurrencia de intencionalidad en los mismos por la manera en que los expresa. El lenguaje por si solo no tiene intrínsecamente relación intencional con el mundo: "Dado que las oraciones –sonidos que surgen de la boca o las marcas sobre el papel- son consideradas objetos del mundo lo mismo que otros objetos cualesquiera, su capacidad para representar no es intrínseca, sino que se deriva de la intencionalidad de la mente" (Searle, 1992: p.13)

Veamos, entonces, más detalladamente como ocurre esto. Los actos de habla comparten varias similitudes con los estados intencionales pues tanto unos como otros tienen la particularidad de hacer referencia a algo, en el caso de los actos de habla pueden ser informes sobre algo representan aquello que está siendo informado mediante una proposición. Se dirá que los estados intencionales se presentan con un contenido representativo que en el lenguaje se expresa como el contenido proposicional de las oraciones. De esta manera, se puede entender a los actos de habla como teniendo un contenido representativo en tanto que sus respectivas proposiciones son capaces de representar algo. Los contenidos representativos además tienen un cierto modo psicológico de ocurrencia, y esto señala la manera en que deben ser entendidos los estados intencionales. En el lenguaje esto equivale a que los contenidos proposicionales tienen maneras de ser proferidos en el lenguaje. Los modos psicológicos de los estados intencionales son creer, esperar, temer, dudar, tener ansiedad, desear, avergonzarse, etc.

> Es preciso distinguir el contenido del estado del tipo de estado que se trata. Así, puedo creer que lloverá o esperar, temer o desear que llueva. El contenido es el mismo en los cuatro casos: que va a llover, pero se relaciona con el mundo de modos psicológicos diferentes: creencia, temor, esperanza, deseo, etc. (Searle, 2006: p.212).

Estas dos propiedades de los estados intencionales (el contenido representativo y su modo psicológico) permiten determinar el tipo de relación que un estado intencional mantiene con

el mundo. Los modos psicológicos de los estados intencionales y actos de habla permiten identificar su dirección de ajuste con el mundo. Hay que decir, primeramente, que los modos psicológicos se clasifican en dos tipos, a saber: creencias (aseverativos) y deseos (directivos y comisivos. Tanto el creer como el desear son modos básicos: la creencia puede contener grados de certeza, suposición, descripción, aseveración; en cambio el deseo puede contener grados de anhelo, querer, codicia, ansia, compromiso, etc.

Ahora bien, según sea la clasificación o modo psicológico en que se presente un cierto estado intencional, se determinará, la dirección de ajuste de su contenido proposicional. Si tenemos un estado intencional con un modo psicológico de creencia, entonces, la dirección de ajuste será de mente a mundo, es decir: si los estados intencionales son dichos proposicionalmente como una creencia ya sea como enunciados, descripciones o aseveraciones, el contenido representativo de tal proposición (lo enunciado) debería encajar o corresponder con una ocurrencia en el mundo. En la medida en que haya correspondencia, dicha creencia podrá ser juzgada como verdadera o como falsa. En cambio, si tenemos un estado intencional con un modo psicológico de deseo, la direccionalidad de ajuste será de mundo a mente, pues cuando los contenidos proposicionales son proferidos a manera de ordenes, mandatos, ruegos, promesas, juramentos o compromisos, quiere decir que algo en el mundo debería poder encajar o corresponder con lo expresado por el contenido proposicional, en tanto, que su emisor debería poder producir cambios en el mundo que fuercen la realización de ese contenido.

Hay un tercer tipo de estados intencionales que carecen de una dirección de ajuste. En estos casos solo se busca expresar una emoción como disculparse, insultar o mostrar alegría; es decir: aquí no se trata de lograr alguna correspondencia con el mundo, pues este tipo de estados intencionales de hecho presupone la veracidad o falsedad de sus proferencias. Por ejemplo, si te disculpas por haber chocado con alguien, tú ya estas asumiendo que de hecho chocaste con esa persona.

Otro punto de conexión entre los actos de habla y los estados intencionales es que al ser proferidos los actos de habla expresan un estado intencional. Por ejemplo, si afirmo algo sobre G estaría expresando mi creencia respecto de G. Tal aseveración determina el estado intencional asociado con ella: “El estado intencional expresado no es sólo un

acompañamiento de la realización del acto de habla. La realización del acto de habla es necesariamente una expresión del correspondiente estado intencional, como se muestra por medio de una generalización de la paradoja de Moore. No se puede decir "Está nevando pero no creo que está nevando"" (Searle, 1992: p.24). Incluso en una mentira se estaría expresando un estado intencional, pero en este caso se está realizando un acto de habla insincero.

El contenido proposicional de un acto de habla o estado intencional determinará un conjunto de condiciones de satisfacción, es decir, el contenido representativo expresa aquello que debe ser satisfecho en el mundo, según las características propias de aquello que está siendo representado en el estado intencional y según el modo psicológico en que éste es expresado. De tal modo las condiciones de satisfacción solo se aplicarán a aquellos actos de habla que tengan una dirección de ajuste. Se dirá, entonces, que todo acto de habla o estado intencional con dirección de ajuste expresa un criterio de satisfacción de su contenido proposicional. Lo que nos interesa de esto es la existencia de una condición de satisfacción de los estados intencionales respecto del mundo nos permite afirmar que estos establecen una relación con el mundo. Cuando queremos hablar de la intencionalidad de los estados mentales, dice Searle, tiene que hablarse de ellos desde la estructura lógica de la intencionalidad y por tanto dando a conocer sus propiedades lógicas, a saber, que tienen un contenido representacional y un modo psicológico.

> Los estados intencionales tienen que ser caracterizados en términos intencionales si no queremos perder de vista su intencionalidad intrínseca. Pero si la cuestión es "¿cuál es el modo de existencia de la creencia y otros estados Intencionales?" entonces, dado todo lo que sabemos normalmente sobre cómo funciona el mundo, la respuesta es: Los estados intencionales son tanto causados por, como realizados en, la estructura del cerebro. Y lo más importante al contestar esta segunda cuestión es ver *tanto* el hecho de que los estados intencionales están en relaciones *causales* con los estados neurofisiológicos (así como, por supuesto, en relaciones causales con otros estados intencionales) como el hecho de que los estados intencionales están *realizados en* la neurofisiología de cerebro (Searle, 1992: p.30).

De este modo, para Searle la mente humana tiene al menos otros dos rasgos que el funcionalismo no logra duplicar con sus programas. El primero y más importante es el de la intencionalidad de los estados mentales, pues como se viene diciendo a diferencia de los estados de la máquina los estados mentales humanos tienen la capacidad intrínseca de hacer

referencia a objetos y eventos del mundo, cosa que la máquina no hace. Si bien es cierto que la máquina es capaz de aparentar tener intencionalidad o bien que sus salidas aparentan mantener una cierta coherencia significativa respecto de sus entradas, esto no significa que de hecho sus estados y acciones contengan una intencionalidad intrínseca. En dado caso sería una intencionalidad derivada de la misma manera que el lenguaje obtiene su intencionalidad de la mentalidad humana: si la máquina tuviese algún tipo de intencionalidad ésta le sería dada por los programadores y los humanos que reciben sus productos mas no a que el programa mismo sea capaz de entender o advertir los significados y relaciones que esos garabatos pueden tener con el mundo.

Según Searle, la intencionalidad del pensamiento humano requiere de una infraestructura o red de otros estados intencionales, así como de un trasfondo de soporte para los estados mentales (creencias, experiencias, destrezas, conocimientos, cultura, etc.) que den paso a los mismos. Estos otros estados intencionales, experiencias, creencias o información del mundo deberían estar sirviendo como base o soporte para que tengan lugar los estados intencionales. Esta red y trasfondo de los estados intencionales es lo que influye, guía o incluso determina el tipo de estados mentales que somos capases de tener. Por ejemplo, él menciona que podemos tener diversas interpretaciones de una misma cosa, Searle da como ejemplo un dibujo como este  y nos dice que este dibujo podría verse de diferentes maneras como la palabra TOOT, como un numero 1001, como unos ojos, una mesa, etc. “Pero estas experiencias y sus diferencias son dependientes de nuestro dominio de una serie de destrezas culturales impregnadas lingüísticamente” (Searle, 1992: p.67).

Entonces si bien la máquina logra simular el pensamiento humano, esa simulación no es una duplicación del pensamiento en tanto que no es capaz de contener una intencionalidad intrínseca de sus estados y acciones. Por mucho que sus entradas y salidas logren ser en apariencia idénticas a las que tendría un humano, éstas no consiguen guardar una relación causal significativa con el mundo, así como con otros estados intencionales. Incluso la manera en que una computadora es capaz de seguir reglas difiere de la capacidad humana para seguir regalas en cuanto que el humano se guía por el contenido significativo de la regla para poder seguirla bajo las más diversas situaciones, mientras que la computadora más bien

simplemente obedece una orden o sigue una secuencia de acciones en un determinado momento.

Por ejemplo, una regla podría ser: debes saludar con una reverencia a tus superiores. Supongamos que esto nos lo dijeron dentro del ámbito militar, como una forma de mostrar respeto a cualquier otro militar que tuviese un rango superior al nuestro. Sin embargo, para un humano esto se podría aplicar dentro de otros ámbitos como al saludar a una persona de mayor edad o a alguien que posea más conocimiento o destrezas que tú en algo, etc. Y podemos, de hecho, seguir esa regla fuera de ese ámbito concreto dado que sus implicaciones significativas tienen una causalidad en nuestras acciones. Pero para la máquina esto tendría que ser descrito en términos muy literales y abstractos o bien de una manera gráfica o esquemática y concreta para que sea capaz de llevar acabo la acción correcta, esto sería, por ejemplo: si tienes X, donde X sería el símbolo de superioridad, inclínate hacia adelante en un ángulo de 45 grados.

> En el sentido en que los seres humanos siguen reglas, en este sentido los computadores no siguen reglas en lo absoluto. Solamente actúan de acuerdo con ciertos procedimientos formales. El programa del computador determina los diversos pasos que deber seguir la maquinaria; determina cómo se transformará un estado en un estado subsiguiente. Y podemos hablar metafóricamente como si fuera un asunto de seguir reglas (Searle, 1985: p.55).

3.2.3 Relación mente-cuerpo

La segunda característica por la cual el funcionalismo no logra duplicar con sus programas la intencionalidad humana es debido a que las computadoras no contienen una realización física (cuerpo) especifica como el ser humano. Pues, la propuesta de la IA fuerte contempla que los programas computacionales puedan ser implementados en cuerpos elaborados de diversos materiales siempre y cuando estos permitan ejecutar el programa. Esta peculiaridad, también llamada múltiple realización de las máquinas de Turing se debe a que la IA fuerte sostiene una postura dualista de la mente, es decir, advierte una distinción entre cuerpo y mente, por lo que propone que los programas computacionales son completamente independientes a la maquinaria física de la computadora. Así para la IA fuerte la implementación física de los programas sería, entonces, más un plus para el pensamiento, y no propiamente parte del mismo.

Turing de hecho menciona que este desapego por una constitución física concreta permite a las computadoras ser máquinas universales ya que para la implementación de los programas no es necesario cambiar los materiales de la máquina o la manera en que están organizados sus mecanismos, simplemente basta con que se cambie el programa para poder llevar a cabo otro tipo de actividad. Pues, para Turing y los defensores de la versión fuerte, lo verdaderamente importante respecto a las máquinas y la mentalidad, es solo la descripción de esa capacidad para llevar a cabo ciertas tareas (el programa). Además, no consideran que pudiera haber una correspondencia de una constitución física particular para cada programa según su función, o bien que aquello donde se podría ejecutar el programa pueda influir o tener un papel esencial en la resolución de tareas. Más bien, sin importar la función del programa éste puede ser ejecutado en diversas máquinas físicas, siempre y cuando el mecanismo tenga las bases para correr el programa.

Para Searle, esto es un fallo importante ya que la intencionalidad tiene eficacia causal, la cual nos permite influir sobre los acontecimientos mundanos, y "la causalidad se considera generalmente como una relación natural entre eventos del mundo" (Searle, 1992: p.123). Searle plantea, entonces, que la mente debería tener una cierta realización física pero que no se base simplemente en ser capaz de almacenar, ejecutar y manipular información. Pues el cuerpo no solo implica que los estados del programa tengan una ocurrencia dentro del ámbito de lo físico o la experiencia física, sino que, conllevaría a que la realización de los eventos mentales tenga un poder causal intrínseco que le permitiría poder entrar en interacción con otros objetos físicos e influir o modificar las relaciones causales de los mismos. Por ello, dice que la intencionalidad es parte de un fenómeno biológico y que nuestra constitución física es lo que nos permite interactuar con el mundo, y peculiarmente la constitución humana parece poder generar capacidades intelectuales que otros seres vivos no son capaces de duplicar.

Así, para Searle, la ocurrencia de mentalidad surgiría como respuesta a una realización física que está provista de ciertos mecanismos causales cuya constitución debería permitirle coaccionar, interactuar y reaccionar (influir) de cierta manera ante la percepción o interacción física de otros eventos físicos. Plantea que los sucesos generados por una cierta constitución física responden a su propia capacidad causal. Estas implicaciones de la causalidad de los objetos físicos concederían la posibilidad de que cualquier acción o tarea que lleve a cabo un

cuerpo deberían poder mantener una relación causal, por así decirlo, directa o fiable y viable con su contexto y entorno físico. Por lo cual, también sería una garantía de la intencionalidad de sus acciones.

En el caso de las máquinas, aunque un programa tuviera una realización física esto no implicaría que aquello que es capaz de hacer o producir conlleve intencionalidad, pues el alcance causal de la máquina está delimitado a una causalidad meramente formal o abstracta, de tal manera que cualquier causación que pudiera tener se viera limitada a una simple expresión, descripción o simulación de causalidad, que no está en relación directa o significativa con su entorno. Cuando se habla de simulación se está señalando que hay una ruptura causal entre el planteamiento de la acción y la realización de la acción o que un determinado suceso no se corresponde causalmente con otro suceso. Esto se debe a que la simulación no tiene las capacidades necesarias para entrar en interacción con otros eventos a los cuales aparentemente tendría que relacionarse causalmente y con los cuales debería interactuar. Por ello, los estados de la máquina no garantizan una relación ni correspondencia con el mundo, en tanto, mantienen una relación disociativa con el mundo:

> El cerebro no hace solamente una instanciación de un patrón formal (también hace eso), pero ésta además causa eventos mentales en virtud de procesos neurobiológicos específicos[4] (Searle, 1990: p.28).

Con esto Searle no está queriendo decir que no sea posible crear mentes o que no puedan existir otro tipo de individuos con otro tipo de constitución física diferente a la humana que contengan mentalidad. Más bien está señalando que para que exista mentalidad es necesario un conjunto de características físicas que sean capaces de realizar acciones de tipo intelectual, pero también debe ser capaz de otorgarle intencionalidad a dichas acciones. Esto significa que, tal vez, no necesariamente se requeriría crear un sistema idéntico al cuerpo humano, es decir: duplicar de manera exacta el sistema de tejidos y componentes biológicos humanos. Bastaría que esa constitución sea capaz de generar el mismo tipo de causalidad que genera la constitución física humana respecto de sus acciones: dicha constitución debería poder

[4] Traducción propia

emular, replicar o duplicar la capacidad y causalidad intencional inherente al funcionamiento del sistema biológico humano.

Este aspecto de la postura de Searle parece reconciliarse con la teoría de la identidad, puesto que identifica al humano como el único ser capaz de contener mentalidad dada su constitución física; la diferencia más obvia con la teoría de la identidad es que no niega la posibilidad de que otro tipo de constitución física pudiera ser capaz de generar mentalidad. Se podría decir que Searle da un paso más que Turing, pues al igual que éste también acepta que puede haber una flexibilidad biológica, fisiológica o física de los cuerpos capaces de contener mentes, pero advierte que no puede ser cualquier tipo de cuerpo: en particular, no puede ser un cuerpo cuya constitución resulte completamente ajena a la mentalidad. Existe una condición y es que si bien el cuerpo debe ser capaz de producir las acciones propias de los eventos mentales, su constitución física requiere cierta correspondencia entre el tipo de eventos mentales que posee y sus capacidades y percepciones físicas. Eso de tal forma que las reacciones corpóreas que el cuerpo tenga en interacción con el mundo se vean reflejados en el tipo de eventos mentales que posee y viceversa.

Por ello, sostiene que hasta el momento el único conjunto de características físicas que sabemos produce mentalidad son las características propias al cuerpo humano. Mientras que una máquina de Turing solo estaría señalando la forma en que podría ocurrir tal o cual suceso. Pues, Searle entiende que la causalidad del cuerpo es lo que da paso a los estados mentales y su intencionalidad. Los defensores de la IA fuerte no contemplaron que una realización física determinada para su máquina pudiera tener implicaciones causales para el pensamiento. Cabe decir que esta causalidad que proporciona el cuerpo también da paso a una relación mente-cuerpo aún más estrecha que las propuestas por Turing y Descartes.

Esto explica porqué para Searle los programas no son suficientes para el pensamiento: para él, algo fundamental del pensamiento es que estos sean sobre algo. No es de extrañarse, entonces, que haga una llamada de atención respecto a esta carencia de intencionalidad en el funcionamiento de las máquinas. Tal vez la cuestión planteada por Turing y la crítica que Searle plantea en contra de esa postura, sería más comprensible si fuera replanteada como ¿la manera en que operan los programas es pensamiento? De esta manera, se identificaría

más claramente que Searle centra su crítica en un análisis del funcionamiento de los programas de las máquinas computadoras.

Así, se está analizando el funcionamiento de las máquinas computadoras como una cierta capacidad que, respecto a la capacidad humana para pensar, no se asemeja respecto a la intencionalidad. Al menos en la forma en que la versión fuerte de la IA plantea el funcionamiento de las máquinas como un uso de símbolos que no requieren ni implican por sí mismos algún tipo de entendimiento, comprensión o referencia a algo: "La peculiaridad de la causación intencional es que nosotros experimentamos directamente esa relación en muchos casos dónde hacemos que algo suceda o que alguna otra cosa hace que algo nos suceda… Experimentamos directamente la relación causal, la relación de una cosa que hace que alguna otra cosa suceda. No necesito una ley que los englobe para decirme que el levantar mi brazo causó que mi brazo suba, porque cuando levanté mi brazo experimenté directamente el causar" (Searle, 1992: p.133).

3.3 Conclusión

Podemos decir, entonces, que la problemática que Searle advierte respecto de sostener la idea del pensamiento como un programa reside en que el funcionamiento de los programas no engloba por entero los aspectos del pensamiento humano, sino que más bien solo está tomando en cuenta una parte de las capacidades del pensamiento humano. Pues el funcionalismo de las máquinas de Turing ve al pensamiento como un quehacer abstracto, formal, y completamente a parte de cualquier fenómeno físico, lo cual genera una disociación entre el pensamiento y su realización física. Esta relación en la que el cuerpo no aporta nada a la mente y simplemente se dedica a ejecutar los programas o las actividades que le ordena el programa provoca que las máquinas de Turing carezcan de intencionalidad, una característica que Searle señala como inherente al pensamiento humano. Sin embargo, esta característica del pensamiento humano es producto del funcionamiento de cuerpo humano.

Podemos decir que el funcionamiento de las máquinas de Turing, tal cual han sido planteadas por la versión fuerte de la IA y el propio Turing, no parece ser suficiente para replicar el pensamiento humano. El pensamiento humano parece ser producto de un proceso mucho más complejo que el solo describir una tarea. Como señala Searle, esta capacidad humana implica

significados o referencias de nuestro entorno; así, se está diciendo que la mentalidad habla de cosas del mundo. Las experiencias del cuerpo y los fenómenos del mundo tienen que ver con la producción de significado, entendimiento y mentalidad. Por ello, Searle señala que el pensamiento humano es parte del mundo natural y no solo pertenece a un ámbito simbólico abstracto.

CONCLUSIÓN FINAL

Tanto Searle como Turing parecen ser sucesores de Descartes; sin embargo, cada uno aceptó solo ciertos planteamientos de Descartes y los desarrollo. Turing se mantuvo más fiel a la dualidad mente-cuerpo mientras que Searle se apartó de la existencia de la dualidad mente-cuerpo y más bien la ve como una forma de hablar de ciertos aspectos, habilidades o capacidades humanas. Empero retoma el aspecto de la postura cartesiana de una capacidad semántica de la mente, como lo más característico y propio de la cognición humana.

Turing le da mucha importancia a la mente, sus actividades y su funcionamiento, en cierto sentido parece buscar llenar el vacío explicativo que dejó Descartes respecto de cómo se dan o elaboran los pensamientos en la propia mente. Pues si bien Descartes describe en qué consisten no logra dar una idea clara respecto de cómo funcionan u operan los pensamientos. De esta manera, Turing termina enfocándose en ésta explicación y relegando la explicación de cómo se dan las interacciones entre los pensamientos y las acciones físicas. Curiosamente, Descartes sí parce advertir que el cuerpo sirve a la mente para más que solo ejecutar los pensamientos pues se da cuenta que la mente contiene pensamientos de cosas y experiencias del mundo, empero me parece que no logra explicarlo de manera congruente la transición que sufrirían tanto la transformación de eventos físicos a eventos mentales como la transformación de los eventos mentales a eventos físicos. Para su época considero su explicación bastante buena pero actualmente resulta bastante insatisfactoria.

Turing descuidó otros aspectos cruciales de la mente como es el que nuestros pensamientos son sobre algo o tienen intencionalidad, dejo de lado todo el aspecto significativo de los eventos mentales y se centró en cómo podrían ser éstos descritos y explicados de tal manera que no fueran completamente subjetivos. Al programar sus máquinas logro aislar la influencia significativa del lenguaje y lo llevó a su nivel lógico más abstracto donde aquello de lo que se habla se ve reducido a expresiones que no son propiamente símbolos en tanto que carecen de significado. Logró entablar con ellos otro tipo de relaciones lógicas que dieran los mismos resultados que se obtendrían si tuviesen contenido significativo.

Por otro lado, con su crítica a la IA fuerte Searle nos hace ver que la postura de Turing no da cuenta de la intencionalidad del pensamiento. Searle diría que cualquier suceso producto de un supuesto sistema mental sin importar si sus productos se encuentran en un sistema visible (físico) o no visible (mental) podrá ser considerado como producto de un pensamiento si se puede corroborar que además de haber una coherencia con el contexto, dicho sistema es capaz de imprimir significado a sus acciones.

Pero algo interesante es que tanto Descartes, como Turing y Searle coinciden en que es mediante el lenguaje que podemos decir algo respecto de los eventos mentales, independientemente de cual sea la postura. Por ello me parece que tanto la intencionalidad, como el lenguaje son puntos clave para el entendimiento del pensamiento. Asimismo, al igual que Turing y Searle considero que sí podría existir mentalidad en otros cuerpos: después de todo el pensamiento ya existe por lo cual me parece que debería existir la posibilidad de replicar dicha capacidad en otros seres. Ciertamente las máquinas son la invención más próxima a lograr la creación de metas en seres no humanos y en la actualidad sobre todo hay muchos nuevos descubrimientos científicos y de ingeniería que parecen estar dispuesto a cumplir la meta de recrear la mente humana.

BIBLIOGRAFÍA

Amaya Amaya, Jairo (2009), *Sistemas de información gerenciales: Hardeare, software, redes, internet, diseño*. Bogotá: Ecoe Ediciones.

Banks, Jerry (1999), Introduction to simulation. En Farrington, P. A.; Nembhard, H. B.; Sturrock, D. T. and Evans, G. W. (eds.) *Proceedings of the 1999 Winter Simulation Conference*, pp. 7- 13

Berlanga Llavori, Rafael; García S., Pedro; Gracia L., Isabel; Iñesta, José y Barber, Fernando (2000), *Introducción a la programación con Pascal*. Castelló: Universidad Jaume I.

Bernard William (2012), *Descartes – el proyecto de la investigación pura*.- 2° Edición al español.- Ediciones Catedra. Madrid España.

Bickle, John (1998), Multiple Realizability, *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.). Recuperado de URL = <https://plato.stanford.edu/archives/spr2016/entries/multiple- realizabilidad />.

Block, Ned (1980), What is Functionalism? En Block, N. (Ed.) *Readings in the philosophy of psychology*, vol. 1 (pp. 171-184). Harvard University Press.

Block, Ned (2007), *Consciousness, Function, and Representation. Collected Papers, Volume 1*. Cambridge: MIT Press

Churchland, Paul M. (1999), *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*. Barcelona: Editorial Gedisa.

Coperland, Jack; Bowen, Jonathan; Sprevak, Mark; Wilson, Robin y otros (2017), *The Turing guide*. New York: Oxford University Press

D.C. INCE. (Ed.) (1992), *Collected Works of A. M. Turing, Mechanical Intelligence*. Amsterdam: ELSEVIER SCIENCE PUBLISHERS B.V.

Descartes, Rene (1997), *Las pasiones del alma*. Madrid: Tecnos.

Descartes, Rene (2002), *Discurso del método y Meditaciones metafísicas*. Madrid: Editorial Tecnos.

Descartes, Rene (2011a), Investigación de la verdad por la luz natural. En Cirilo Flórez, Migel (2011), *Descartes. Estudio introductorio de Cirilo Flórez Miguel* (pp. 73-96). Madrid: Gredos.

Descartes, Rene (2011b), Tratado del hombre. En Cirilo Flórez, Migel (2011), *Descartes. Estudio introductorio de Cirilo Flórez Miguel* (pp. 73-96). Madrid: Gredos.

Descartes y Leibniz (1989), *Sobre los principios de la filosofía*. Madrid España: 1° Edición al español, Editorial Gredos.

Graham, George (2000), Behaviorism. *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.). Recuperado de URL = <https://plato.stanford.edu/archives/spr2017/entries/behaviorism/>.

Hilary Putnam (1960), Minds and Machines. En Anderson, A. R. (Ed.) (1964), *Minds and Machines* (pp. 72-97). U.S.A.: PRENTICE-HALL, Inc. Englewood Cliffs, N.J.

Levin, Janet (2004), Functionalism. *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.). Recuperado de URL = <https://plato.stanford.edu/archives/fall2018/entries/functionalism/> .

Lahoz-Beltrá, Rafael. (2005), *Turing. Del primer ordenador a la inteligencia artificial*. España: NIVOLA libros y ediciones.

Nagel, Thomas (1974), Whats is it like to be a Bat? *Philosophical Review*, vol. 83, 435 – 450

Robinson, Howard (2003), Dualism, *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.). Recuperado de URL = <https://plato.stanford.edu/entries/dualism/#MinBod>.

Ryle, Gilbert (2005), *El concepto de lo mental*. Barcelona: Paidós.

Searle, John R. (1980), Minds, brains, and programs. *The Behavioral And Brain Sciences,* vol. 3, 417-457.

Searle, John R. (1985), *Mentes, cerebros y ciencia*. Madrid: Ediciones Cátedra.

Searle, John R. (1990), Is the Brain's Mind a Computer Program? *Scientific American*, 26-31.

Searle, John R. (1992), *Intencionalidad. Un ensayo en la filosofía de la mente.* Madrid: Editorial Tecnos.

Seare, John R. (2006), *La mente. Una breve introducción.* Bogotá: Grupo Editorial Norma.

Smart, J. J. C. (1959), Sensations and Brain Processes. *The philosophical Review*, vol. 68, No. 2, 141-156

Smart, J. J. C. (2000), The Mind/Brain Identity Theory. *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.). Recuperado de URL = <https://plato.stanford.edu/archives/fall2018/entries/mind-identity/>.

Turing, A. M. (1932), *Nature of Spirit*. Recuperado de http://oldshirburnian.org.uk/alan-turing-and-the-nature-of-spirit/

Turing, A. M. (1950a), Computing Machinery and Intelligence. En Anderson, A. R. (Ed.) (1964), *Minds and Machines* (pp. 4-30). U.S.A.: PRENTICE-HALL, Inc. Englewood Cliffs, N.J.

Turing, A. M. (1950b), Maquinas computadoras e inteligencia. En Anderson, A. R. (Ed.) (1970), *Mentes y maquinas* (pp. 10-45). D.F.: Universidad Nacional Autónoma de México.

Villena Saldaña, David. (2017), ¿Qué es el funcionalismo? *Letras*, *88*(127), 129-155. Recuperado en 08 de enero de 2019, de http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S2071-50722017000100006&lng=es&tlng=es.