

Método
Grupo Transdisciplinario de Investigación
en Ciencias Sociales
www.grupometodo.org

Borradores de Método

Área de Metodología
ISSN: 1692-9667

de-Codificando La Matriz

Eduardo Salcedo Albarán

Documento 33
Julio 25 de 2005

BORRADORES DE MÉTODO es un medio de difusión de las investigaciones del Grupo Método. Estos documentos son de carácter provisional, de responsabilidad exclusiva de sus autores y sus contenidos no comprometen a la institución.

Borradores de Método. No 33. Julio 25 de 2005.
Editor Fundación Método
Colección Metodología.

© Eduardo Salcedo Albarán
de-codificando La Matriz

© Fundación Método. Carrera 8ª. No 37-10. ofi: 501. Telefax: (571) 4005765
Todos los derechos reservados.
Primera edición 2005.
Impreso en Bogotá, Colombia.

de-codificando La Matriz

Eduardo Salcedo Albarán

Resumen

El objetivo de este trabajo es desarrollar un análisis semiológico de la película Matrix. Este análisis consiste en identificar algunos de los posibles signos que pueden constituir distintas lecturas de la obra. Cada uno de estos signos sirve como dispositivo de entrada a alguna lectura específica. No se pretende desarrollar en su totalidad una sola lectura de la obra, ni identificar todos los signos que en cada una de ellas se encuentran; sólo se proponen algunos signos que sirvan como dispositivos de entrada a posibles lecturas. El escrito consta de seis partes. En la primera parte se propone la lectura del sueño, la vigilia, y el sueño inducido. En la segunda parte se propone la lectura de la Inteligencia Artificial (I.A.). En la tercera parte se propone la interpretación del sueño como simulación. En la cuarta parte se propone la lectura del evento evolutivo denominado Evolución Artificial Fuerte (E.A.F.). En la quinta parte se propone la lectura del evento evolutivo denominado Evolución Artificial Débil (E.A.D.). En la sexta parte se propone la lectura acerca de la constitución epistemológica del mundo en el que vivimos. Cada uno de los signos propuestos, que constituyen parte de cada lectura, remiten a unos contenidos teóricos específicos; dichos contenidos serán expuestos en cada caso.

P

iéñese en la obra como un camino y en la interpretación como una posible manera de recorrer el camino. En cada interpretación siempre se recorrerá el mismo camino; no obstante, se recorrerá de distinta forma. Esta analogía se agota en un aspecto: el camino, a diferencia de la obra, nos lleva a un mismo punto y sospecho que la obra no siempre nos lleva al mismo punto; dependiendo de cómo se recorra, la obra nos llevará a diferentes destinos; una obra no siempre produce la misma sensación o idea final en quienes la aprecian.¹ Por lo demás, la obra es el camino y la interpretación es la manera como lo recorreremos. Ahora bien, las formas de recorrer el camino son prácticamente inagotables y sospecho que constituyen la función por excelencia del intelecto humano. La gran característica del pensamiento humano consiste en interpretar los objetos de distintas maneras; por este motivo, reflexionamos de maneras inimaginables acerca de un mismo objeto. La riqueza del pensamiento humano se encuentra en la capacidad de hallar ambigüedad en los objetos y esta ambigüedad se manifiesta de manera ideal en las interpretaciones estéticas.

Ahora bien, la forma de recorrer el camino tiene un comienzo; hay un instante t en que el camino comienza a recorrerse de una determinada manera; pero una forma particular de recorrer el camino no comienza con el principio del camino. Esto quiere decir que la obra no siempre se comienza a leer desde el principio, de una determinada manera; se puede comenzar su lectura según una interpretación y momentos más tarde cambiarla. En el caso del cine, sospecho que una determinada interpretación no comienza en los primeros milímetros de cinta ya editada. Esto permite comenzar a leer la obra desde la interpretación X y, al cabo de algunos minutos, cambiar a la interpretación Y para, finalmente, terminar con la interpretación Z. Ahora bien, el comienzo de cada uno de estos distintos tipos de lectura se marca por un determinado signo que, en el cine, puede consistir en una palabra, un color o cualquier otro elemento de composición. Estos signos pueden entenderse como dispositivos de entrada a cada interpretación. Puede pensarse que, de esta manera, un signo es unívocamente significatorio porque permite la entrada a una sola lectura. No obstante, el flujo sincrónico y diacrónico de significación del signo,² implica la posibilidad de múltiples formas de caminar a partir un único signo. Bajo condiciones controladas, de laboratorio, como la presente, podemos hablar de *un* camino que parte de *un* signo; no obstante, en condiciones reales sería testarudo esperar que diferentes espectadores de la obra observaran algo idéntico en cada signo. Por este motivo, lo máximo que puedo hacer es proponer un posible camino que parte de un posible dispositivo de entrada al camino, es decir, que parte de un signo; lo máximo que puedo hacer es proponer posibles lecturas de la película, que parten de posibles signos.

Las anteriores palabras, para mostrar que pretendo proponer algunos dispositivos de entrada a posibles lecturas de la película *The Matrix* (1999); no obstante, no pretendo poner en funcionamiento total ninguna de estas lecturas; tampoco pretendo exponer todas las posibles interpretaciones; ello sería tarea de nunca acabar; por el contrario, pretendo mostrar algunos dispositivos de entrada a algunas interpretaciones posibles.

¹ Podría pensarse en la fragmentación de caminos, esto es, aceptar que en cada manera de recorrer el camino se constituye un camino distinto. Esto implica que cada interpretación constituye una obra diferente. Para mi propósito supondré que cada interpretación no constituye una obra distinta, sino que todas se hacen sobre la misma obra antológica.

² Si bien Deleuze no habla de un flujo sincrónico y diacrónico, sí señala que “(...) el signo remite al signo, y remite al signo hasta el infinito”. [Ver Deleuze y Guattari, *Mil Mesetas*, pág. 119]

Acercamiento conceptual

¿Cómo podemos saber que en este preciso instante no vivimos en una simulación, en un mundo virtual, en un sueño o en un simulacro de la realidad? Esta es una de las preguntas más importantes que, en conjunto, plantea la obra. Esto quiere decir que la película puede entenderse, en sí misma, como un signo que remite a esta pregunta. El límite entre lo real y lo irreal es prácticamente irreconocible si lo irreal tiene características idénticas a lo que conocemos como real; colocados dentro de una simulación que sea idéntica a lo que conocemos como real, es muy probable que salvo por alguna falla en la simulación, podamos percatarnos de que vivimos en una simulación; con el sueño la situación es aún más complicada pues sospecho que muy pocas veces, a pesar de las posibles incoherencias presentes en un sueño, podamos percatarnos de que en realidad estamos soñando; por lo general, sólo nos percatamos del sueño durante la vigila. Ahora bien, las posibilidades de percatarnos de que nos encontramos dentro de una simulación disminuyen drásticamente si desde nuestro nacimiento hemos vivido en este mundo virtual. Simulacro, virtualidad, simulación; todos estos conceptos se refieren a la posibilidad de crear un modelo de lo real que pueda engañarnos, sea cual sea el motivo y propósito de este engaño. En este orden de ideas, simulacro, simulación y virtualidad son sólo modelos de lo real; es decir, no son lo real. Pero cuando el simulacro, la virtualidad y la simulación son idénticas al escenario real, y es lo único que conocemos como real, entonces el límite entre lo real y lo que no lo es se desdibuja; incluso, surge la posibilidad de que el mundo simulado pueda interpretarse como real. Si lo único que conocemos es la simulación, entonces puede que la simulación sea para nosotros lo real; si lo único que vemos son sombras, entonces esas sombras serán lo que conocemos como realidad.³

Lo primero que se requiere para la realización de la posibilidad planteada por la película es una concepción cognitiva fisicalista; es decir, se requiere que los estados mentales puedan reducirse y, por lo tanto, pervivir en cualquier tipo material e instancia física. En principio, algunas interpretaciones de la obra pueden entenderse como una alusión al escepticismo cartesiano, razón por la cual puede pensarse en la posibilidad de dualismo de sustancias; no obstante, si bien la duda metódica juega un papel importante, el contenido de la obra no puede desarrollarse completamente desde un dualismo de sustancias. Es común afirmar que, dada la posibilidad de separación entre la mente y el cuerpo –separación determinante en la película-, entonces puede inferirse un dualismo; no obstante, en la obra queda claro que el cuerpo no puede vivir sin la mente.⁴ Esto, a su vez, no quiere decir que no se pueda distinguir la mente

³ La alusión platónica, en este caso, es inevitable: ¿acaso no concebían como reales las sombras que las personas encadenadas en la caverna veían? Efectivamente así era porque el lenguaje de aquellas personas encadenadas solamente permitía representarse aquello que ellos veían y que concebían como real. Lo más real que conocían era las sombras y por este motivo dichas sombras eran la realidad.

⁴ La posibilidad de distinguir el cuerpo de la mente, pero no separarlos como sustancias, puede entenderse desde el funcionalismo, para una explicación de por qué el funcionalismo no es una forma de dualismo, Cfr. Salcedo, *El experimento mental de la habitación china: máquinas entre la semántica y la sintaxis*.

del cuerpo, tan solo quiere decir que la mente siempre tiene que existir en algún tipo de sistema físico. La mente no puede separarse del cuerpo y continuar existiendo.

1. El sueño, la vigilia y el sueño inducido

- 1.1. “¿Has tenido esa sensación en la que no estás seguro si estás despierto o dormido?”

Este signo remite a Descartes. Una primera lectura de la obra, que considero un tanto incompleta, se refiere a la posibilidad de que estemos sumergidos en un profundo sueño del que nunca podemos ni podremos despertar. Ya Descartes se había planteado esta posibilidad en las *Meditaciones Metafísicas*; en algunos casos sus sueños eran tan vívidos como para confundirlos con la realidad:

[...] ¡Cuántas veces me ha sucedido soñar de noche que estaba en este mismo sitio, vestido, sentado junto al fuego, estando en realidad desnudo y metido en la cama! Bien me parece ahora que, al mirar este papel, no lo hago con los ojos dormidos; que esta cabeza, que nuevo, no está somnolienta; que si alargo la mano y la siento, es de propósito y a sabiendas; lo que en sueño sucede no parece tan claro y tan distinto como todo esto. Pero si pienso en ello con atención, me acuerdo que, muchas veces, ilusiones semejantes me han burlado mientras dormía; y, al detenerme en este pensamiento, veo tan claramente que no hay indicios ciertos para distinguir el sueño de la vigilia, que me quedo atónito y es tal mi extrañeza, que casi es bastante a persuadirme de que estoy durmiendo.⁵

Al respecto, un argumento importante en contra de la posibilidad de que estamos sumergidos en un sueño, es la falta de continuidad de la vida que vivimos cuando estamos soñando. Cuando dormimos y comenzamos a soñar, el desarrollo del sueño no se empalma de manera perfecta con el desarrollo del sueño de la noche anterior; contrario a esto, cuando nos despertamos, el desarrollo de nuestra vida de vigilia sí empalma perfectamente con el desarrollo del día anterior; nuestras vivencias, recuerdos y sensaciones mantienen continuidad entre un día y otro. Ahora bien, nótese que el argumento de la interrupción de la vida de sueño no permite escapar del planteamiento de la obra. En condiciones normales, es de esperarse que la vida de los sueños no mantenga continuidad; esto quiere decir que en condiciones normales, en nuestra vida debemos romper la continuidad de las vivencias de los sueños. Sin embargo, ¿qué sucede si estamos sumergidos en un sueño profundo y constante, mediante algún tipo de narcótico? ¿Qué sucede si vivimos en un estado de coma profundo del que nunca hemos despertado y nunca vamos a despertar? En estos dos casos no habría, necesariamente, discontinuidad en la vida de los sueños y, por lo tanto, no habría razones para pensar que la falta de continuidad nos permitiría percatarnos de cuál es el mundo de los sueños y cuál es el mundo de la vigilia. Este primer signo no nos lleva aún a cuestionamientos tan radicales; sólo

⁵ Ver Descartes, *Meditaciones metafísicas*, pág. 21.

abre la posibilidad de que, en algunos casos podamos confundir el mundo de los sueños con el mundo de la vigilia.

- 1.2. “¿Has tenido un sueño, Neo, del que estuvieras seguro que era real? ¿Qué pasaría si no te pudieras despertar de ese sueño? ¿cómo podrías saber la diferencia entre el mundo de los sueños y el mundo real?”

La continuidad parece otorgarle el calificativo de *real* al mundo que interpretamos como real. La diferencia entre el mundo de los sueños y el mundo de la vigilia, consiste en que el primero se ve interrumpido cada noche, de manera que es muy poco probable que podamos llevar una vida paralela a la vida de la vigilia. Ahora bien, si el mundo de los sueños es idéntico al mundo real, y si además es constante, de manera que no percibimos alteraciones bruscas en el desarrollo de la vida⁶ en los sueños, entonces sería imposible distinguir el mundo de los sueños del mundo de la vigilia. Este signo nos indica que es probable que Neo se encuentre sumergido en un sueño o que algunas de sus vivencias, que él concibe como reales, pertenezcan al mundo de los sueños sin que él mismo lo sepa. Este signo remite, en general, a la posibilidad de que en este preciso instante estemos viviendo un profundo pero constante mundo de sueños y que, por lo tanto, todas las representaciones inmediatas no guarden correspondencia con nuestra situación espacio-temporal.

- 1.3. “tienes la mirada de un hombre que acepta lo que ve porque está esperando despertarse (...)”

Este signo parece indicar que Neo, efectivamente, está viviendo en un mundo de sueños o que, al menos, algunas de sus vivencias pertenecen al mundo de los sueños, que se confunde con las vivencias del mundo de la vigilia.⁷ Este signo nos indica una bifurcación entre

⁶ Es necesario referirse a *alteraciones bruscas*, porque en la vida de la vigilia en muchas ocasiones experimentamos alteraciones leves en la continuidad de nuestras vivencias. Por ejemplo, después de consumir alguna droga fuerte, es muy probable que el estado alterado de conciencia redunde en lo que folclóricamente se conoce como “laguna mental”, es decir, olvidar por completo nuestras vivencias mientras nos encontrábamos en el estado alterado; en este caso, puede haber una interrupción del flujo de vivencias concientes. No obstante, esta interrupción en la continuidad de nuestras vivencias no desdibuja nuestra identidad o las distinciones entre el mundo real y el mundo del estado alterado. Hay otro tipo de discontinuidades, mucho más drásticas, que sí implican confusiones con respecto a la identidad propia y a distinciones entre lo que se sueña y lo que se vive en la vigilia; por ejemplo, una fuerte amnesia puede generar crisis de identidad o confusiones con respecto a si un contenido vago que retorne del lapso olvidado, pertenece al mundo de los sueños o al mundo de la vigilia. Otros tipos de síndromes, generalmente causados por daños en los lóbulos frontales, hacen que las personas sólo puedan retener contenidos en su memoria por lapsos muy cortos; cuando esto sucede, la persona olvida lo que estaba haciendo minutos atrás. Esto implica pérdida de continuidad en los actos, pues se es incapaz de empalmar vivencias pasadas con vivencias actuales y proyecciones. Salvo estos casos, en condiciones normales, nuestra vida de vigilia se caracteriza por ser continua.

⁷ Nótese que hasta este punto he hablado del mundo de la vigilia como el mundo *real*; sin embargo, como ser verá más adelante, lo que es *real* también puede ser cuestionado aún si hablamos del mundo de la vigilia.

dos mundos, aunque aún no nos aclara la distinción o el tipo de vivencias que pertenecen a cada uno. La bifurcación se infiere también de los siguientes signos:

- 1.3.1. En el minuto 17, el agente Smith le dice a Neo que está viviendo dos vidas; una dentro de las reglas de las instituciones y los protocolos sociales, en la que lleva el nombre de Thomas Andersen, y otra por fuera de las reglas institucionales, en el mundo de la computadoras, en la que lleva el nombre de Neo; “una de estas vidas tiene futuro y una de ellas no”
- 1.3.2. En el minuto 21 Neo se despierta de lo que parecía ser un sueño. Había sido encarcelado por tres agentes y, en medio del interrogatorio, su boca se selló; además, un extraño dispositivo con vida propia fue puesto dentro de su cuerpo. Esta extraña vivencia, entre sicodélica y surrealista, empalma con la vivencia de él mismo despertándose de un sueño. De esta manera, parece que dicha vivencia perteneció al mundo de los sueños porque desencadena en un *despertarse*.
- 1.4. “[...] este es el mundo que ha sido puesto ante tus ojos para cegarte de la verdad”.
- 1.5. “[...]has estado viviendo en un mundo de sueños, Neo”
- 1.6. “[...]con la píldora azul [...] la historia termina, te despiertas en tu cama y crees lo que quieras creer”

Los signos 1.4., 1.5. y 1.6. indican que efectivamente hay dos mundos; específicamente el signo 1.4. permite suponer que el mundo de los sueños ha sido inducido por alguien o por algo; este signo implica la acción de una tercera entidad que tiene la capacidad de manipular el mundo de los sueños. Ahora bien, con respecto al signo 1.4., hay de una clara similitud con el escepticismo cartesiano. El signo 1.4. remite a la posibilidad de que alguna entidad nos esté engañando; en Descartes dicha entidad es el genio maligno, mientras que en La Matriz son las máquinas:

Supondré, pues, no que Dios, que es la bondad suma y la fuente suprema de la verdad, me engaña, sino que cierto genio o espíritu maligno, no menos astuto y burlador que poderoso, ha puesto su industria toda en engañarme.⁸

- 1.7. Neo: “mis ojos me duelen”.
Morpheus: “Nunca los has usado antes”.

Este parece ser uno de los signos más importantes en la presente lectura. Finalmente, se aclara el mundo al que pertenecen las vivencias de Neo: lo que ha experimentado durante

⁸ Ver Descartes, *Meditaciones metafísicas*, pág. 24.

toda su vida ha formado parte del mundo de los sueños; toda su vida Neo ha estado durmiendo y esta es la primera vez que comienza a usar sus ojos, es la primera vez que está despierto, en estricto sentido; por lo tanto, sus ojos le dolerán tanto como aquel que intenta mirar al sol después de salir de *la caverna*. Una entidad manipuladora ha inducido a Neo a un profundo y constante sueño del que sólo podía despertar con la ayuda de alguien que ya estuviera despierto.⁹

2. Inteligencia Artificial (I.A.)

- 2.1. “[...] en algún punto del siglo XXI la raza humana se unió en una celebración, [...] nos maravillamos de nuestra propia magnificencia al dar nacimiento a la I.A.”

Este signo revela la naturaleza de la tercera entidad que ha producido el mundo de sueños en el que Neo vive: una Inteligencia Artificial; ¿cuál es la porción de ciencia y cuál es la porción de ficción de esta posibilidad?

La posibilidad teórica de la inteligencia artificial se sustenta en el funcionalismo, que es un tipo de explicación que flexibiliza sus consideraciones acerca de la composición física necesaria para afirmar que un proceso puede ser considerado como mental. ¿Qué es lo que hace que un carburador sea un carburador? Funcionalmente, un carburador se puede definir como aquello que mezcla gasolina y aire en un motor; esto quiere decir que un carburador puede definirse como un concepto funcional en términos de sus relaciones causales. ¿Este tipo de explicación puede hacerse extensivo al caso de la mente? Puede pensarse que sí, y que el cerebro en realidad ejecuta cierto tipo de tareas, definidas en términos de sus relaciones causales, formalizables e instantiables en una máquina, Así, el funcionalismo ha recurrido a las máquinas como punto de referencia de su explicación, porque en ellas se puede distinguir claramente entre las relaciones causales de las funciones desarrolladas únicamente por el soporte físico y las funciones desarrolladas únicamente por el soporte lógico.

Si los procesos mentales pueden ser caracterizados y definidos en términos de sus relaciones causales, entonces éstos pueden ser ejecutados por una máquina que compute una cantidad muy alta de procesos algorítmicos que sean formalizaciones de los procesos mentales; una máquina con esta capacidad es la máquina de estado discreto de Turing.¹⁰ La máquina de estado discreto puede ejecutar un número infinito de estados distintos, mediante unas entradas y salidas grabadas en una cinta con divisiones. En cada división de la cinta hay un símbolo que

⁹ Esto no quiere decir que las personas conectadas a La Matriz solamente pueden despertar gracias a la mediación de una persona que ya está desconectada; es posible que algunos factores mentales permitan a una persona despertar por su cuenta. Esta hipótesis se plantea en el capítulo de animatrix titulado *World Record*.

¹⁰ Cfr. Turing, *La maquinaria de computación y la inteligencia*.

la máquina lee y sobre el cual decide dejar impreso, o borrar y escribir otro. Aunque la máquina sólo puede leer, borrar, marcar y correr la cinta para pasar al siguiente cuadro (o estado), cada estado del programa puede definirse funcionalmente dentro del papel que desempeña en la operación de toda la máquina. Esto, porque la máquina de estado discreto, o máquina universal de Turing, consiste en el concepto de un aparato que determina los estados ejecutados por un número n de otras máquinas de iguales características. Dado que al hablar de la máquina de estado discreto no se hace referencia a la composición física, o a los materiales que se deben usar para que ésta sea eficiente, el funcionalismo puede reconocer la eventual irrelevancia del *hardware* dentro de los procesos mentales. En conclusión, se puede determinar y limitar la definición funcional de un estado psicológico en términos de los estados de los programas ejecutados por la máquina universal de Turing.

Ahora bien, surge una complicación al adoptar el funcionalismo. Si se acepta esta postura, entonces es necesario aceptar que una máquina de café, un dispensador de gaseosas, un abridor de válvulas y, en general, cualquier mecanismo que ejecute programas de estados, podría ser considerado como un aparato pensante. Cuando una máquina de café realiza procesos de estados, no es muy claro que esté en capacidad de *creer, imaginar, soñar* y, en general, de *pensar*. Esto sugiere dudas acerca del carácter intencional de los contenidos mentales pues básicamente surgen dos opciones: (1) Las máquinas pueden poseer intencionalidad; (2) Poseer mente implica algo más que ejecutar estados computacionales; es decir, se puede ejecutar estados mediante un mecanismo y carecer de intencionalidad y capacidad de representación.¹¹

2.2. “(...) una conciencia singular que engendró una raza entera de máquinas”

Ahora bien, la I.A., por sí sola, no representa la posibilidad de máquinas que estén en capacidad de crear mundos de sueño para dominar a los humanos. La I.A. consiste en un programa de investigación que tiene el propósito de esclarecer algunas cuestiones acerca de la naturaleza del pensamiento humano y, en dicho programa, se identifican dos corrientes: la I.A. fuerte (I.A.F.) y la I.A. débil (I.A.D.). El presente signo nos remite a la I.A.F. pues en la I.A.F. se concibe la posibilidad de que las máquinas posean verdadero pensamiento y conciencia.

La principal distinción entre la I.A.D. y la I.A.F. consiste en que, para la primera, una máquina no puede poseer intencionalidad genuina, conciencia o entender algo. Teóricos de la I.A.D. aceptan que una máquina solamente puede simular los estados y contenidos mentales de un ser humano; puede simularlos pero no duplicarlos o poseerlos.¹² Ahora bien, para los teóricos de la I.A.F., una máquina no sólo simula los contenidos y estados mentales sino que los duplica y, por lo tanto, los posee. En este orden de ideas, una máquina puede pensar, ser intencional, ser consciente y entender todo, tal como cualquier persona podría hacerlo. Uno de los argumentos más importantes en contra de la I.A.F. es el experimento mental de la habitación china, de John Searle. Según este argumento, la intencionalidad es un rasgo de nivel superior, causado y realizado en la constitución microbiológica del cerebro. Esto quiere decir

¹¹ Para apoyar la segunda posibilidad, Searle ha propuesto el argumento que se conoce como el experimento mental de la habitación china. Este argumento será ampliado más adelante.

¹² Con respecto a la distinción entre simulación y duplicación *Cfr.* Fodor, *La explicación psicológica*.

que materiales artificiales nunca podrán generar el rasgo subjetivo de la mente humana. Según Searle, (1) la mente humana posee semántica y sintaxis, (2) la sintaxis no genera semántica y (3) como siempre que una máquina está procesando información está haciendo una manipulación sintáctica, computacional, entonces (4) estos procesos nunca podrán ser idénticos al procesamiento de información que realizan los seres humanos, el cuál sí es poseedor de semántica. Según Searle, las máquinas nunca podrán poseer verdadero y genuino pensamiento.¹³ Así pues, la posibilidad propuesta por la obra, se refiere específicamente a la I.A.F.; matemáticos, evolucionistas, ingenieros, filósofos y otros teóricos participan en esta postura argumentando la posibilidad de que una máquina pueda ser tan inteligente como cualquier persona.

Por su parte, para Minsky, la mayoría de argumentos en contra de la I.A.F., incluyendo el experimento mental de la habitación china, consiste en asegurar que a las máquinas les hace falta un elemento misterioso que, aunque existe, no puede percibirse; se supone “la existencia de alguna parte mágica, que no tiene propiedades indetectables”¹⁴ y con base en esta suposición se procede a asegurar que una máquina nunca podrá ser tan consciente como un ser humano. Al respecto, Minsky dice que:

No deberíamos estar buscando algún elemento específico, faltante. La mente humana tiene muchos ingredientes y cada máquina que se ha construido carece de docenas o cientos de ellos.¹⁵

Es claro que la mente humana es mucho más flexible, recursiva y adaptable que una máquina; para Minsky, en estos tres elementos radica la principal distinción actual entre el pensamiento humano y el posible pensamiento de una máquina. No obstante, la presencia de estos elementos en la mente humana no implica que sea imposible duplicarlos en otro tipo de material distinto al orgánico que actualmente nos caracteriza. Cuando una máquina actual se enfrenta a un problema y no encuentra cómo solucionarlo, simplemente se detiene, se atasca; no obstante, cuando los humanos nos encontramos en esta misma situación procedemos a interpretar el problema de distintas maneras hasta que finalmente alguna de las interpretaciones nos permite una solución; de hecho, cuando desistimos de generar interpretaciones alternas, lo hacemos básicamente porque decidimos desistir y no porque nuestra programación se agote. Esta extensa capacidad de posibilidades de interpretaciones, flexible y adaptable, resulta de la gran cantidad de mecanismos instalados en nuestro cerebro. Sabemos que el conocimiento se representa de maneras distintas en nuestro cerebro y que cada una de estas maneras de representación se almacena en distintas regiones; también sabemos que estos distintos almacenamientos se ejecutan mediante distintos procesos. Típicamente, un programa de computador actual utiliza un solo tipo de representación y por este motivo carece de la

¹³ Para los planteamientos originales del experimento mental de la habitación china *Cfr.* Searle, *Mentes y cerebros sin programas*; *Cfr.* Searle, *The Rediscovery of the Mind*. Para una revisión de la naturaleza del argumento y de sus refutaciones *Cfr.* Moural, *The Chinese Room Argument*; *Cfr.* Salcedo, *El experimento mental de la habitación china: máquinas entre la semántica y la sintaxis*.

¹⁴ Ver Minsky, *Will robots Inherit the Earth*.

¹⁵ Ver Minsky, *Will robots Inherit the Earth*, pág. 9

flexibilidad característica de la mente humana; partiendo de este hecho, se supone que los computadores no pueden pensar.

Según Minsky, muchos filósofos piensan que el entendimiento y la conciencia son característicos de una mente incorporada en una biología viva, pero esta idea se fundamenta en el hecho de que las máquinas parecen no manifestar la posibilidad de interpretar y enfrentar los problemas desde diferentes puntos de vista:

Si usted [...] entiende algo de una sola manera, entonces en realidad no ha entendido nada. Esto es porque, si algo sale mal, usted se atasca [...]; esto es porque, cuando alguien aprende algo 'de memoria', decimos que en realidad no está entendiendo. De todas maneras, si usted tiene distintas representaciones entonces, cuando una aproximación falla, usted puede intentar otra. [...] Representaciones bien conectadas le permiten analizar ideas en su mente y prever las cosas desde diferentes perspectivas hasta que encuentra la que le funciona. Y eso es lo que queremos decir con *entender*.¹⁶

El cerebro humano no usa una sola representación, sino que ejecuta distintos procesos en paralelo y supervisiones de nivel superior, con lo cual posibilita la constante reformulación de problemas y la reinterpretación de un mismo suceso:

La razón por la que pensamos tan bien no es porque contenemos partes misteriosas [...], es porque empleamos asociaciones de agencias que trabajan en concierto para mantenernos alejados de un atascamiento. Cuando descubramos cómo funcionan esas asociaciones podremos ponerlas dentro de computadores. Entonces, si una aproximación en un programa se atasca, otro mecanismo puede sugerir una alternativa de aproximación. Si usted viera una máquina hacer cosas como estas, ciertamente [...]pensaría que la máquina es consciente.¹⁷

En general, para los teóricos de la I.A.F., la constitución física no implica restricciones en las posibilidades de procesamiento de información y almacenamiento; así, teniendo el conocimiento necesario acerca de los mecanismos de procesamiento y almacenamiento del cerebro humano, se posibilita la duplicación de contenidos mentales en cualquier tipo de material; incluso, en posturas más radicales de la I.A.F., no es necesario contar con el conocimiento detallado acerca del cerebro humano y no debemos esperar a hacer una duplicación artificial de este órgano, para lograr la posibilidad de que una máquina posea cualidades mentales. Según McCarthy, uno de los pioneros de la Inteligencia Artificial, máquinas simples como termostatos poseen creencias; en algunos casos, la adscripción de cualidades mentales es la mejor manera para hablar acerca del funcionamiento de las máquinas:

Adscribir ciertas creencias, conocimiento, libre voluntad, intenciones, conciencia, habilidades o deseos a una máquina o un programa de computador es legítimo cuando dicha adscripción expresa la misma información acerca de la máquina que la que se expresa acerca de una persona.¹⁸

De manera legítima, se puede asegurar que un termostato tiene, al menos, tres creencias asociadas a tres acciones específicas que él ejecuta:

¹⁶ Ver Minsky, *Why People Think Computers Can't Think*.

¹⁷ Ver Minsky, *Will robots Inherit the Earth*, pág. 11

¹⁸ Ver McCarthy, *Ascribing Mental Qualities to Machines*, pág. 1

‘La habitación está muy fría’, ‘la habitación está muy caliente’, y ‘la habitación tiene la temperatura correcta’. (Además,) le adscribimos el propósito ‘la habitación debería tener una adecuada temperatura’. Cuando el termostato cree que la habitación está muy fría o muy caliente, manda un mensaje diciéndole esto al horno. Un predicado de creencia levemente más complejo podría también usarse, en el que el termostato tiene creencias acerca de la temperatura correcta que debería haber y otra creencia acerca de la temperatura que hay.¹⁹

Lo más importante de las consideraciones hechas por los teóricos de la I.A.F. es que la estructura cognitiva de creencias e intenciones puede estar presente en un robot o en un programa de computación. Si se puede adscribir creencias a máquinas tan simples como termostatos, entonces parece más legítimo aún adscribir creencias a máquinas complejas como programas de conversación, programas de solución de problemas o robots. Para MacCarthy, de la complejidad de la máquina depende la complejidad del tipo de contenido mental que se puede adscribir; así, es legítimo adscribirle contenidos mentales complejos a máquinas complejas como computadores. Así, una máquina, en tanto cúmulo de material tradicionalmente interpretada como *artificial* o cúmulo de algoritmos ejecutables, puede tener creencias, deseos, voluntad y, en general, intenciones; una máquina puede tener pensamiento en el mismo sentido en el que una persona tiene pensamiento.²⁰

3. La simulación y el sueño como simulación

Mediante el signo 1.4. quedó claro que el estado de sueño profundo y continuo ha sido producido por alguna entidad, esto nos lleva a cuestionar la naturaleza de dicha entidad y el mecanismo que usa para producir el mundo de sueño en el que Neo vive. El signo 1.4.1. nos mostraba que una de las vidas del personaje era “vivida en los computadores”; esto parece significar que Neo vive en una simulación computarizada.

Una de las principales características de una simulación y, por lo tanto, de una creación virtual de un modelo real, es el hecho de que las condiciones reales y las condiciones simuladas deben poder engañar al espectador de la simulación:

Se podría decir que una máquina simula con éxito (...) cuando jueces entrenados son incapaces de discriminar la conducta de la máquina de la conducta del organismo en una situación de prueba que sea relevante.²¹

¹⁹ Ver McCarthy, *Ascribing Mental Qualities to Machines*, pág. 14

²⁰ Un clásico criterio para determinar que una máquina tiene pensamiento idéntico al de un ser humano es el test de Turing o “el juego de la imitación”. Según Alan Turing, si una persona A conversa con una persona B y un computador C, y se da el caso en que la persona A no puede diferenciar cuál es la máquina y cuál es la persona, entonces es legítimo asegurar que la máquina C piensa de manera idéntica a como piensa la persona B. *Cfr.* Turing, *La maquinaria de computación y la inteligencia*.

²¹ Ver Fodor, *La explicación psicológica*, pág. 161

Si bien esta definición no especifica lo que es *una situación de prueba relevante*, en el caso planteado por la película la simulación ha estado presente en todos los momentos de la vida de Neo, por lo que se puede asegurar que sea cual sea la situación de prueba relevante, la simulación ha salido favorecida;. En ningún momento Neo, ni alguna otra persona que esté conectada a La Matriz, se percatado de que está viviendo en una simulación y no en la realidad. Este tipo de simulación permite un control total de las posibles variables contextuales, es decir, esta simulación permite crear un ambiente de laboratorio en el que el espectador puede y debe quedar atrapado en los poderes causales de la simulación. Esto sucede incluso para una simulación como el test de Turing, en el que el interés principal de la conversación es hacer que el espectador quede atrapado en los poderes causales de la máquina y la persona, hasta el punto de que no pueda diferenciar cuáles son las respuestas de la máquina y cuales son las respuestas de la persona. En el caso de la simulación planteada por Turing, se puede asegurar que los poderes causales cognitivos de la persona y la computadora tienen el mismo tipo de efectos sobre el espectador. Las respuestas de la persona y la computadora tienen poderes causales sobre los estados psicológicos del espectador que, entre otras cosas, crea una imagen de su interlocutor; gracias a estas respuestas el espectador crea una representación de su interlocutor;²² así pues, cuando estos poderes causales son idénticos en la persona y en la máquina, se puede hablar de una simulación exitosa.²³

De esta manera, se puede asegurar que la simulación y el diseñador de la simulación superan el test de Turing porque no hay fallas o defectos que permitan diferenciar el sistema real del simulacro o simulación. Si Neo pudiera percatarse de que el mundo en que vive no es el mundo real, entonces la simulación no sería exitosa. No obstante, se debe tener en cuenta que el éxito de la simulación no sólo se da gracias que esta posea un diseño casi perfecto, sino a que es el único ambiente que conocen los espectadores de la simulación. Esto hace que su lenguaje esté limitado por la frontera de cognición de su mundo, de lo que para él es la realidad²⁴.

- 3.1. “¿(...) sabía que la primera Matriz estaba diseñada para ser un mundo perfecto, donde no había sufrimiento, donde todo el mundo sería feliz? Fue un desastre. Nadie aceptaba el programa. Cosechas enteras se perdieron.”

²² La imagen que cree el espectador será idéntica aunque esté escuchando las respuestas dadas por la computadora; en la medida en que, frente a las respuestas dadas por la computadora, el espectador se represente una persona, entonces se puede asegurar que los poderes causales de los enunciados de la computadora son prácticamente idénticos a los poderes causales de los enunciados de cualquier ser humano.

²³ Vale la pena advertir que, al aceptar la relevancia de los efectos psicológicos del espectador en la dinámica de interlocución, es necesario aceptar que los actos de habla están determinados, en gran parte, por dichos efectos psicológicos y no, en su totalidad, por el nivel que cualitativamente pueda poseer el emisor, ya sea computador, robot o persona. *Cfr.* Salcedo, *Robots, actos de habla e imposibilidad de verificación intencional*.

²⁴ Aquí la similitud con la *alegoría de la caverna*, de Platón, de nuevo es evidente. Ahora bien, esto no implica un solipsismo porque hay coincidencia entre la cognoscibilidad del mundo de Neo y la cognoscibilidad del mundo de todas las demás personas que están conectadas a La Matriz. Todos viven en una simulación idéntica, con las mismas fronteras y limitaciones de cognición. Este aspecto será desarrollado más adelante.

Este signo nos remite a un importante aspecto técnico de las simulaciones psicológicas, señalado por Fodor:

(...) la investigación psicológica está continua e implícitamente redefiniendo las condiciones necesarias para que se dé una adecuada simulación de los tipos de conducta en cuestión. Intentar enumerar *a priori* tales condiciones equivaldría, por tanto, a intentar anticipar los resultados que se tendrán que seguir de esa investigación.²⁵

Así pues el ensayo y error parece ser el único procedimiento que permite el diseño de simulaciones psicológicas exitosas. La única manera de escapar al método de ensayo y error, a la corrección de los detalles de la simulación sobre la marcha, sería anticipando los resultados de la simulación; ahora bien, como la simulación en este caso se refiere al aspecto psicológico del ser humano, y como dicho aspecto se caracteriza por una intencionalidad y subjetividad que redundan en impredecibilidad,²⁶ entonces es casi imposible establecer *a priori* las condiciones necesarias y suficientes para una simulación psicológica exitosa. Sería necesario anticipar todas las respuestas psicológicas y las consideraciones de los actos que hagan las personas conectadas a la simulación.

¿Cómo es posible predecir los resultados de una simulación psicológica? Por una parte, desde la deducción se intentó mostrar que un correcto cálculo de predicados es suficiente para hacer, y por lo tanto predecir, todas las consideraciones necesarias de los actos humanos. Este enfoque fue inspirado por los métodos de resolución de teoremas y consistió en considerar la posibilidad de formalizar y representar el conocimiento mediante axiomas y lógica común, o cálculo de predicados, para deducir los efectos de las acciones a ejecutar. Para permitir este cálculo, se utilizan unos axiomas antecedentes, llamados *axiomas estructurales*, que definen las condiciones y efectos de cada acto en un sistema particular. Con respecto a estos axiomas estructurales, Dennett dice que:

(...) a este conjunto de axiomas el sistema le aplicaría una acción -postulando la ocurrencia de una acción A en una situación S - y luego deduciría el efecto de A en S para obtener una descripción de la situación resultante S'.²⁷

Ahora bien, se debe considerar el hecho de que nuestra actividad consciente no se asemeja, según lo que podemos percibir, a un constante cálculo de predicados; de esta manera, se puede proceder en este asunto, sobre cualquiera de las siguientes hipótesis: o nosotros no pensamos mediante el cálculo de predicados, aunque un robot pueda hacerlo, o nosotros sí pensamos deductiva pero inconscientemente mediante el cálculo de predicados. Dejando de lado la veracidad de cualquiera de las hipótesis sobre las cuales se sustenta el enfoque deductivo, Dennett dice que lo cierto es que este no ha mostrado funcionar de manera

²⁵ Ver Fodor, *La explicación psicológica*, pág. 162

²⁶ Así, la física, que se encarga de sistemas físicos no intencionales, produce predicciones del comportamiento del sistema de manera exacta, mientras que la sociología, la antropología o la economía no, pues su objeto de análisis son sistemas intencionales.

²⁷ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 181

eficiente, salvo en casos triviales. Una de las principales complicaciones del cálculo de predicados es que cualquier aspecto X de una determinada situación Y puede cambiar en una circunstancia Z, por lo cual se requiere introducir un axioma nuevo y distinto para cada una de las circunstancias Z que pueden resultar de una modificación en el aspecto X dado en la situación Y.²⁸

Esta consideración y formulación de axiomas se presenta como una tarea extremadamente difícil de realizar; dicha explosión axiomática se intentó solucionar mediante la suposición de que nada cambia en una situación salvo aquello que se dice explícitamente que cambia, pero rápidamente los investigadores abandonaron esta suposición al percatarse de la inaplicabilidad de la cláusula *ceteris paribus* al mundo real. Además, la formalización de la cláusula *ceteris paribus* parece bastante complicada porque no implica una definición explícita su significado:

¿Qué quiere decir que las demás cosas permanecen iguales? ¿Exactamente qué disposiciones de cuáles otras cosas se considera que permanecen iguales? Si tuviéramos que responder a esta pregunta sería inútil invocar la cláusula *ceteris paribus*, puesto que la utilizamos precisamente para evitar esa tarea. Si se pudiera responder esta pregunta no habría, en primer lugar, necesidad de invocar la cláusula.²⁹

Una manera de enfrentar el problema desde el enfoque deductivo, por adelantado, sin tener que incluir nuevos axiomas para cada una de las aclaraciones necesarias con respecto a aquello que puede o no puede cambiar en una determinada situación, consiste en considerar la posibilidad de agrupar las variaciones de circunstancias por estereotipos. Así, sería necesario formalizar sólo algunos estereotipos específicos, sin necesidad de hacer explícita cada una de las variaciones. Estos estereotipos son escenarios típicos o paradigmáticos en los que, generalmente, se desarrollan nuestros actos; estos escenarios paradigmáticos son *estructuras*, según Minsky, y *guiones*, según Schank.³⁰ Ahora bien, Schank y Abelson también reconocen la imposibilidad de proveer un único procedimiento que permita enfrentar todas y cada una de

²⁸ Pareciera que en cierto tipo de ejercicio inductivo también se presenta una complicación similar, en la que es necesario repasar todas y cada una de las posibles variaciones que se pueden dar en todas y cada una de las posibles circunstancias posibles.

²⁹ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 180

³⁰ El uso de la metodología de guiones y estructuras no ha sido marginal en la I.A., por el contrario constituyen un pilar de esta investigación; de hecho, según John Searle, el argumento que usa para refutar la I.A. fuerte, conocido como *la habitación china*, se le ocurrió cuando leía acerca de los programas de comprensión de relatos de Shank y Abelson. Searle señala que como *input* se le puede proporcionar al programa un relato del tipo *un hombre fue a un restaurante y pidió una hamburguesa. Cuando le trajeron la hamburguesa estaba quemada. El hombre salió enfurecido del restaurante sin haber pagado la hamburguesa.* Después, se le puede preguntar al programa ¿el hombre se comió la hamburguesa? y este, como *output*, responderá que no. En ninguna parte del relato se dice explícitamente, que el hombre no había comido la hamburguesa y, sin embargo, el programa fue capaz de responder satisfactoriamente la pregunta. Esta capacidad se da gracias a que el programa cuenta con un *guión de restaurante* que indica cómo son los acontecimientos que generalmente suceden en un restaurante. En este guión se señala que, por lo general, cuando una persona se enoja al recibir la comida, y no paga la cuenta, entonces no comió lo que había pedido. De esta manera, el programa aparea la pregunta con el guión de restaurante. En este enfoque no es crucial la cuestión de si las máquinas poseen y experimentan sentimientos y emociones, lo que se pretenden es formular una teoría del conocimiento que dé luces en la manera como se actúa en el mundo de manera exitosa.

las tareas del mundo. Esto quiere decir que si bien se puede proponer un sistema de características deductivas –en la medida en que a partir de unos guiones se introducen variaciones-, es necesario ligar el contenido de la tarea a ejecutar con el método procedimental que permite actuar:

No existe forma de desarrollar un sistema de comprensión adecuado en un ordenador sin proveer al mismo con un conocimiento extenso sobre el mundo concreto con el que tienen que tratar [...]. Debemos ser cautos ante la posibilidad de que el conocimiento en un área pueda estar organizado de acuerdo con principios diferentes del conocimiento en otra área. Quizás no hay un sólo conjunto de reglas y relaciones para construir todas las bases de conocimiento potencial que se requieran, [...] creemos que la forma de representación de conocimiento no debe ser demasiado deferente del contenido. Cuando el contenido cambia drásticamente, la forma también debe cambiar.³¹

No obstante, este enfoque deductivo de estereotipos presenta algunas complicaciones que el mismo Dennett señala:

Mientras que los insectos parecen cargar con esos sistemas de control, la gente, aun cuando parezca que depende de estereotipos, cuenta con sistemas de respaldo del pensamiento con los que puede abordar más eficazmente los problemas que surgen. Además, cuando la gente se vale de estereotipos, depende al menos de estereotipos que ella misma concibe. Hasta la fecha nadie ha sido capaz de proponer alguna idea que funciones acerca de cómo se podría orientar el mecanismo de elaboración de estructuras o de guiones de una persona mediante su experiencia previa.³²

De esta manera, si bien el enfoque deductivo puede orientar la acción eficiente de una máquina, hasta el momento no se ha propuesto un estereotipo, guión o estructura que funcione correctamente en un ambiente que no sea *de juego*. Como La Matriz tiene que procesar la información del sentido común y las condiciones ordinarias de miles de personas, entonces el ambiente que simula no es *de juego* sino de *cotidianidad*; por lo tanto la probabilidad de anticipar **todos** los posibles resultados psicológicos de **todas** las acciones de **todas** las personas conectadas debió haber sido, en principio, bastante baja; por este motivo, es muy probable que el diseñador de La Matriz haya tenido que afinar sus errores en la marcha.

3.2. Oracle: “te pediría que te sentaras, pero de todas maneras no lo vas a hacer.

Y no te preocupes por el jarrón (...)”

Neo: “¿cuál jarrón? (...)”. - Neo tropieza accidentalmente con un jarrón y lo bota con su codo izquierdo-.

Oracle: “ese jarrón”.

Neo: “[...] lo siento”.

Oracle: “te dije que no te preocuparas”

Neo: “¿cómo lo supo?”

Oracle: “[...] lo que en realidad te va a revolver la cabeza después, es si aun lo habrías roto si yo no hubiese dicho nada”

³¹ Ver Schank y Abelson, *Guiones, planes, metas y entendimiento*, pág. 17

³² Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 106

Es muy probable que una vez afinada, descifrada y entendida la estructura de toma de decisiones de las personas que estarían conectadas, se pudiera proveer la sensación de libertad y, con esto, se simularía el libre albedrío, aunque el diseñador de la simulación pudiera realizar una anticipación de todas las respuestas psicológicas.³³ Este desciframiento y comprensión, por parte del diseñador de La Matriz, que redundaba en capacidades predictivas, parece reflejarse en *Oracle*, quien misteriosamente conoce la estructura de decisión y la estructura de causalidad que usa La Matriz para simular la toma de decisiones y el libre albedrío.³⁴ Así pues, el presente signo nos remite al hecho de que es posible conocer la estructura de toma de decisiones y de causalidad epistemológica³⁵ de las personas. Una vez descifrada dicha estructura, es posible hacer predicciones que parezcan profecías; por este motivo, algunos de los enunciados del *Oracle* parecen sentencias. Una de las condiciones más importantes de la vida humana parece ser la posibilidad de decidir libremente. Es bien conocido el debate filosófico acerca del determinismo y el libertarismo, acerca de si somos libres para decidir o si estamos determinados a actuar según la voluntad de un ser superior o según el flujo de causalidad; no obstante, en la cotidianidad las personas funcionamos mediante una constante toma de decisiones aparentemente libres. Sospecho que incluso los partidarios más radicales de la postura determinista tienen que tomar varias decisiones cada día, para vivir normalmente. De esta manera, es de esperarse que la toma de decisiones también tuviera que ser simulada por el diseñador de La Matriz, para que esta representara un ambiente familiar a quienes participaran de la simulación. Así, las personas que están conectadas a La Matriz pueden experimentar la sensación de decidir, aunque esta decisión sea aparente porque solamente están obedeciendo a una estructura causal. Tal vez, en la realidad, nuestra situación no sea muy distinta.

3.3. “Recordarás que no crees en ninguna de esa basura del destino. Tu estás en control de tu propia vida [...]”.

3.4. “[...] La Matriz no puede decirte quién eres”.

3.5. “Te dije que yo sólo podía mostrarte la puerta, tu tienes que caminar a través de ella”.

³³ Esto no parece diferir mucho de la idea de que, aunque en la realidad experimentamos una sensación de toma de decisiones, estamos determinados por el flujo de causalidad o por una entidad de orden superior.

³⁴ En la segunda parte de la obra se aclara la relevancia de *Oracle* en el diseño de la Matriz; este aspecto no será ampliado en el presente documento.

³⁵ La causalidad epistemológica se refiere a la adscripción de relaciones causales. Esta causalidad epistemológica se diferencia de una causalidad ontológica, en la medida en que es el resultado de nuestras atribuciones, las cuales están limitadas por nuestras capacidades perceptivas. A esta causalidad se le puede contrastar la causalidad ontológica, que sería la estructura de relaciones causales de la realidad. Para la discusión acerca de la atribución de relaciones causales epistemológicas, *Cfr.* Scholl y Tromoulet, *Perceptual causality and animacy*.

Al parecer, la estructura causal de La Matriz no consiste en un determinismo absoluto pues, aunque se pueden predecir las consecuencias de las acciones –incluyendo las decisiones de las personas conectadas, pues la toma de decisiones también forma parte del flujo de la relación causal-, también parece claro que es posible alterar el curso de dichas relaciones causales. Esta posibilidad de alteración, reconocida por *Oracle*, se resalta en la segunda parte de la película, razón por la que no será analizada en el presente escrito. No obstante, vale la pena observar que los tres anteriores signos remiten al hecho de que la estructura causal simulada en La Matriz puede alterarse. Al parecer, hay una clara diferencia entre “conocer” la profecía y “vivir” la profecía, aunque, esta distinción también puede ser el resultado de la simulación de la estructura. Tener conocimiento sobre el futuro puede formar parte del mecanismo de persuasión para simular la capacidad de tomar decisiones acerca de ese mismo futuro: Saber que si X, entonces el futuro será Y, y evitar X para evitar Y, puede ser parte del mecanismo que emplea La Matriz para hacer que la elija X y, de esta manera, auto reforzar la aparente capacidad predictiva. Así pues, estos tres signos nos llevan de regreso al problema ontológico y metafísico del libre albedrío: es probable que aunque día a día experimentemos la posibilidad de elegir nuestro propio destino, esta sensación no sea más que el resultado de la misma experiencia determinista. Es probable que en realidad no elijamos nada, o que elijamos lo único que podríamos haber elegido, y no obstante experimentemos una sensación de libertad. Estos tres signos nos remiten a las preguntas: ¿en la realidad hay una estructura de causalidad que nos determina? ¿En la realidad hay una entidad metafísica que nos determina? ¿En la realidad actuamos con plena libertad de elección?

- 3.6 “[...] tiene las mismas reglas básicas, como gravedad. Estas reglas no son distintas a las de un sistema de computador. Algunas de ellas pueden ser violadas, otras pueden ser rotas.”
- 3.7 “[...] he visto agentes golpear a través de paredes de concreto. Hombres que han vaciado proveedores completos sobre ellos impactándole a nada más que al aire. No obstante, su fuerza y velocidad están aún basadas en un mundo construido en reglas. Por ese motivo, ellos nunca serán tan fuertes o tan rápidos como tu puedes serlo”.

Los dos anteriores signos remiten al hecho de que los humanos parecen tener “alguna” una ventaja sobre las máquinas, pues estas últimas están regidas por reglas de programación; al parecer, a diferencia de las máquinas, los humanos poseen cierta intuición inimitable. Si bien la posibilidad de programas intuitivos se amplía en la segunda parte de la película, parece claro el hecho de que los humanos tienen una capacidad cualitativa insuperable al momento de violar y romper reglas;³⁶ esta capacidad no puede estar presente, en igual

³⁶ En la tercera parte de la película se resalta la diferencia que hay entre un rasgo puramente sintáctico y un rasgo semántico de, tal vez, el concepto más intuitivo y más difícil de formalizar: el amor. En la obra de Barthes, *Fragmentos de un discurso amoroso*, se plantea la necesidad de lectura en primera persona de todos y cada uno de los fragmentos que pueden referirse a situaciones específicas del amor. Esto implica que cualquier aproximación conceptual al amor parece estar atrapada en un fuerte subjetivismo, atrapada en una lectura de sujeto en primera persona; esto, a su vez, implica que el amor parece ser uno de los

medida, para el caso de las máquinas pues ellas están definidas en términos de dichas reglas: el fundamento constitutivo de toda máquina es la programación y, en esta medida, es difícil pensar en la constante ruptura de dicha programación. Ahora bien, debe tenerse en cuenta que esta es, en el fondo, la pregunta más importante del debate de inteligencia artificial. Si aceptamos la idea de que las personas son o poseen algo más que pura sintaxis, entonces, salvo en el caso de que la sintaxis sea suficiente para generar semántica, las máquinas nunca podrán imitar la capacidad conciente de las personas, ni la capacidad de romper las reglas de programación; no obstante, si se reconoce y acepta que las personas consistimos en pura sintaxis de interacción entre microniveles y macroniveles de nuestra constitución biológica, entonces, después de todo, no tendríamos ninguna diferencia radical cualitativa con una máquinas. No obstante, pareciera que, entre más entendemos nuestra constitución biológica, más nos asemejamos a una máquina que ejecuta reglas de interacción de micro elementos pues, recuérdese que nuestra constitución biológica, incluida parte de la estructura de nuestro cerebro³⁷ que determina nuestras conductas, son el resultado de las instrucciones consignadas en nuestro ADN y, por lo tanto, de algoritmos ejecutados al momento de la realización de nuestro código genético. Esto requiere una interpretación informacional de los genes y de la molécula de ADN. Dicha interpretación no representa una excentricidad; por el contrario, ha sido aceptada en el cuerpo teórico de la biología molecular.³⁸

Ahora bien, lo que parece ser una capacidad intuitiva de los humanos para romper las reglas de programación, puede entenderse de una manera “tosca” o de una manera “elegante”, en función de sus efectos. Por su parte, la manera “tosca” consiste en resaltar la capacidad que tienen las personas para escapar a los límites impuestos por las instituciones que imparten control y regular la conducta de individuos y agregados sociales. A pesar de la fuerza regulativa y coercitiva que se imponga en una sociedad, las personas siempre encuentran la manera de escapar a dichas regulaciones. Por otra parte, la manera “elegante” consiste en remitirnos a la genialidad y a la creatividad. La capacidad para interpretar los objetos del mundo de maneras múltiples permite la dinámica de creación artística y creación científica; esta lógica de la creación, que parece difícil de formalizar y simular, es una importante característica del intelecto humano. Con respecto a la creación artística, debe tenerse en cuenta el caso de robots que, en cierto sentido, crean imágenes visuales a partir de una programación básica y de un limitado conjunto de reglas de comportamiento.³⁹

Al parecer, es claro que los humanos estamos en capacidad de interpretar de manera casi inagotable cualquier objeto del mundo, y esto no es distinto para el caso de las reglas que,

conceptos más intuitivos y subjetivos del lenguaje de lo mental. Esto no impide que una máquina pueda emitir un sonido o una marca en los que se exprese el enunciado “te amo”; no obstante, parece haber cierto rasgo semántico, representacional, que escapa de las posibilidades sintácticas de una máquina.

³⁷ La totalidad de la estructura de nuestro cerebro no es, únicamente, el resultado de las predisposiciones dictaminadas por nuestros genes pues, gracias a la capacidad de plasticidad del cerebro, la estructura cambia como resultado de nuestras experiencias. Así, la totalidad de nuestra estructura cerebral es el resultado de predisposiciones genéticas desencadenadas por la interacción con el entorno.

³⁸ Cfr. Williams, *A package of information*, pág. 42.

³⁹ Al respecto, puede observarse el caso de Aaron, “un robot computarizado, construido y programado por Harold Cohen”. [Ver Kurzweil, *The age of spiritual machines*, pág. 167].

se suponen, dictaminan nuestro comportamiento.⁴⁰ Si estamos determinados por algún conjunto de reglas, entonces tenemos la capacidad de interpretar dicho conjunto de reglas de distintas maneras. Esta idea puede entenderse si se acepta que lo que somos es el resultado de la ejecución de algunos algoritmos genéticos expuestos a distintos factores desencadenantes del entorno. Aunque, en principio, los resultados de las combinaciones de estos algoritmos genéticos son predecibles, la intencionalidad humana continúa caracterizándose por un fuerte rasgo de “dificultad de predicción” y, por este motivo, continuamos diferenciando entre sistemas físicos y sistemas intencionales. Lo mismo sucede con la cantidad total de posibles estructuras sinápticas. Si bien, en principio, la cantidad total de posibles estructuras sinápticas es contable, predecir la aparición de estas estructuras es una tarea bastante difícil de llevar a cabo.

Así pues, llegamos a la cuestión central: la presencia y estipulación de un conjunto inicial de reglas impuestas al comportamiento de un sistema, no garantiza la predictibilidad práctica de dicho sistema.⁴¹ Esto quiere decir que, en muchas ocasiones, incluso sin romper las reglas, estamos en capacidad de escapar a sus determinaciones inmediatamente predecibles. Si se acepta la idea de que podemos interpretar casi inagotablemente un conjunto de reglas, entonces sería necesario aceptar que, en principio, un robot también debería estar en capacidad de interpretar de manera insospechable un conjunto de reglas y, por lo tanto, escapar a las determinación inmediatamente predecibles; esto quiere decir que, en principio, un robot podría encontrar una interpretación o combinación estrictamente predecible pero insospechada. Esto nos lleva a la posibilidad de que, sin necesidad de romper las reglas se den dilemas, por ejemplo, morales. Isaac Asimov presentó lo que actualmente se conoce como las “leyes de la robótica”⁴², según las cuales se debe regir la conducta de un robot. En muchas historias de Asimov se plantean situaciones en que los robots, sin quebrantar dichas reglas, entran en conflictos de decisión; en otros casos, dispositivos o algoritmos alternativos permiten una combinatoria insospechada que lleva a efectos igualmente insospechados.⁴³ Tal vez nuestro cerebro, cumpla la función de aquellos poderosos dispositivos capaces de hallar combinaciones completamente imprevistas e insospechadas entre reglas inicialmente estipuladas.

⁴⁰ A riesgo de caer en vaguedad, acepto la posibilidad de entender dichas reglas como las reglas estipuladas por la programación genética, que predisponen en gran medida nuestras tendencias conductuales, y las reglas de regulación básicas de socialización que, aunque en un sentido distinto, también regulan nuestra conducta.

⁴¹ Si bien esta predictibilidad puede ser teórica, en términos prácticos es una tarea prácticamente imposible.

⁴² (1.) Un robot no puede lastimar a un ser humano o, a través de sus acciones, permitir que un ser humano sea lastimado. (2.) Un robot debe obedecer las órdenes dadas por un ser humano, excepto en el caso de que estas órdenes entren en conflicto con la primera ley. (3.) Un robot debe proteger su propia existencia en la medida en que esa protección no entre en conflicto con la primera o la segunda ley.

⁴³ Ahora bien, con respecto a las referencias cartesianas de la obra, se debe tener en cuenta que Descartes también había concebido la posibilidad de que, aunque un genio maligno lo estuviera engañando, este no le pudiera imponer ningún tipo de pensamiento; esto quiere decir que Descartes también puede escapar de las reglas impuestas a su pensamiento: “[...] con gran cuidado procuraré no dar crédito a ninguna falsedad, y prepararé mi ingenio tan bien contra las astucias de ese gran burlador, que, por muy poderoso y astuto que sea, nunca podrá imponerme nada.” [Ver Descartes, *Meditaciones metafísicas*, pág. 24]

- 3.8 “Oracle me dijo que me enamoraría, y que el hombre que yo amara sería el elegido. Así que ya ves, no puedes estar muerto; no puedes estarlo porque te amo.”

Este signo plantea la siguiente cuestión: ¿Cómo es posible que Oracle sepa acerca del enamoramiento de Trinity si este sucede por fuera de La Matriz? Este enamoramiento se dio por fuera de La Matriz y, por lo tanto, por fuera de la estructura causal de La Matriz. Esto quiere decir que la estructura de causalidad y toma de decisiones en La Matriz es idéntica a la estructura de causalidad y toma de decisiones que usan las personas que no están conectadas a ella; esto prueba que las máquinas y, específicamente, el diseñador de La Matriz, han logrado descifrar los procesos de toma de decisiones y los mecanismos de acción de los seres humanos. Ya señalé en páginas anteriores que intentar conocer esta estructura por adelantado, mediante cálculo de proposiciones y métodos deductivos, es prácticamente imposible. Por este motivo, el presente signo remite a la otra posibilidad de comprensión de la estructura de causalidad y toma de decisiones, a saber, mediante los métodos de inferencia inductiva.

Suponiendo que las personas diseñemos planes antes de ejecutar cualquier acción, entonces, al parecer la inducción sirve para:

- (a) Calcular la probabilidad de que, al ejecutar la acción X , *suceda, se obtenga o se realice* Y ⁴⁴; teniendo en cuenta que Y es la realización exitosa de la acción X . De esta manera, Y sirve de criterio de éxito de la acción X .
- (b) Calcular la probabilidad de que, al ejecutar la acción X , *no suceda, no se obtenga o no se realice* W ⁴⁵.

Parece evidente que a y b presentan dos posibles funciones que podría desempeñar la inducción en la acción humana; sin embargo, es interesante preguntar si a puede separarse y considerarse como una función distinta de b o si a y b forman parte de un mismo mecanismo o proceso de diseño de planes de acción. A mi parecer, no hay implicación lógica entre a y b , luego estas se pueden entender como dos funciones distintas. Calcular la probabilidad a es un proceso supremamente complejo, en la medida en que requiere todo el conocimiento del sentido común adquirido en experiencias pasadas y un proceso eficiente de segundo nivel que permita asociar todo este conocimiento; sin embargo, el cálculo de la probabilidad a no implica las complicaciones computacionales que implica el cálculo de la probabilidad b . Si se acepta que la cantidad de mundos posibles es infinita⁴⁶, entonces el cálculo de la probabilidad b no solamente es difícil, sino que es computacionalmente imposible. Esto último quiere decir que es básicamente imposible que el cerebro, incluso mediante redes distribuidas en paralelo, pueda

⁴⁴ Entiéndase Y en tanto objeto en un sentido amplio como una acción, un conjunto de acciones, o un simple objeto tecnológico

⁴⁵ W puede definirse de la siguiente manera:

- $W = \{S / S \text{ es un objeto distinto de } Y\}$, esto es,
- W es el conjunto de cada uno de los objetos resultantes en cada uno de los mundos posibles distintos al mundo posible en el que de la acción X resulta Y .

⁴⁶ Cf. La investigación social de las decisiones morales y la filosofía moral, proposiciones 16 -19.

realizar este tipo de cálculo. Así, la proporción de resultados de cada tipo de cálculo⁴⁷ nos permite identificar la función que, probablemente, cumple la inducción en nuestro proceso de toma de decisión

Ahora bien, si se pueden separar los cálculo *a* y *b*, entonces el diseño de un sistema de planificación inductiva no implicaría una contradicción lógica; sin embargo, si estos dos cálculos no se pueden separar entonces sí sería lógicamente imposible este tipo de diseño.⁴⁸ Esta consideración permite entender la manera en que las personas toman decisiones y cómo dichas decisiones están inmersas en un flujo causalidad:

Recordemos una vez más al pobre R1D1 y supongamos, para una mejor claridad de la discusión, que tenía un conocimiento empírico perfecto de las probabilidades de todos los efectos de todas sus acciones que él mismo podía detectar. Por tanto, cree con probabilidad de 0.7864 que la ejecución de JALAR (CARRETA, HABITACIÓN) causará que las ruedas de la carreta hagan un ruido audible; y con probabilidad de 0.5 que la puerta de la habitación se abrirá en este caso hacia adentro y no hacia fuera; y con probabilidad de 0.999996 que no habrá elefantes vivos en la habitación, y con probabilidad de 0.997 que la bomba seguirá en la carreta cuando la mueva. ¿Cómo podría R1D1 encontrar esta última aguja importante en el pajar del conocimiento empírico?⁴⁹

Desde este punto de vista, lo que Dennett presenta en el anterior párrafo es un cálculo del tipo *b*. Si se tiene en cuenta que este es computacionalmente imposible y, por lo tanto, presenta limitaciones lógicas, entonces se puede asegurar que mediante el cálculo de *b* no se pueden diseñar planes de acción; si así fuera, cada plan de acción debería hacer predicciones *ad infinitum* y esto, a su vez, redundaría en un congelamiento de la acción. Sin embargo si, como cuestioné líneas atrás, el cálculo de *a* se puede separar del cálculo de *b* entonces se posibilitaría el diseño de sistemas de planificación de acción mediante inducción. El cálculo de la probabilidad *a* sería algo de este tipo:

Una enciclopedia ambulante se caería en un barranco a pesar de todo su conocimiento sobre barrancos y sobre los efectos de la gravedad, a menos de que estuviera diseñada de tal manera que pudiera encontrar las partes correctas de conocimiento en el momento preciso para que pueda planificar sus compromisos con el mundo real.⁵⁰

En este orden de ideas, la selección y asociación de cierto conocimiento empírico permitiría el cálculo eficiente de, únicamente, la probabilidad tipo *a*. Esto no quiere decir que la posibilidad de este cálculo sea una realidad actual en cuanto a las simulaciones, tan sólo quiere decir que no hay imposibilidad computacional en dicho cálculo. Ahora bien, dado que no hay imposibilidad lógica en el cálculo de la probabilidad tipo *a* –en la medida en que no hay imposibilidad computacional – podría suponerse que si la capacidad de predicción es una función cerebral⁵¹, entonces esta se da mediante un cálculo tipo *a*. Esto, a su vez, permitiría

⁴⁷ Esto, en caso de considerar el resultado del cálculo \square como posible porque se elimine el supuesto de infinitos mundos posibles.

⁴⁸ Siempre que se mantenga el supuesto de infinitud de mundos posibles.

⁴⁹ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 180

⁵⁰ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 180

⁵¹ Cfr. Llinás, *El mito del yo*, pág. 4

que *Oracle* realizara predicciones de sucesos ajenos a La Matriz, mediante la ejecución de este complejo pero fáctico mecanismo inductivo que, posiblemente, guarda proporcionalidad entre lo que sucede en el cerebro de las personas conectadas y desconectadas.

- 3.9. En el minuto 32, después de que Neo ha sido absorbido por el espejo del mundo de los sueños, se despierta dentro de un capullo, conectado con unos cables.

Este signo significa que el mundo de los sueños ha sido producido mediante conexiones de cables que cumplen el papel funcional de lo que puede entenderse como un vínculo vital entre el feto y su madre. Mediante estos cables se proporciona los nutrientes necesarios para mantener en buen estado físico el cuerpo de las personas conectadas y los impulsos eléctricos que alimentan el cerebro para producir la simulación. Con esto, se puntualiza la distinción entre un mundo real y un mundo simulado y se muestra que este último tiene el propósito de esconder algún aspecto del mundo real.

- 3.10. “¿En este momento, estamos dentro de un programa de computador?”

- 3.11. “La Matriz es un mundo de sueño generado por computador”.

Posiblemente, lo que se concebía como un sueño según la primera lectura, ahora se conciba como un programa de simulación ejecutado por computador. Esto implica que los computadores están en capacidad de, por sí mismos, crear simulaciones que permitan dominar a los humanos. Esta posibilidad sólo se puede dar en el marco de la I.A.F., pues se requiere una inteligencia completamente autónoma, con capacidad de prever las posibles necesidades, consideraciones e intenciones de la mente humana. Una inteligencia artificial de este tipo, según Hans Moravec, debe ser una inteligencia de calidad universal.⁵²

4. Máquinas, Evolución Artificial Fuerte (E.A.F.) y cyborgs

Esta lectura consiste en examinar las posibles consecuencias evolutivas de la I.A.F. Si llega el día en que las máquinas estén en capacidad de duplicar el pensamiento humano y por lo tanto posean una inteligencia flexible e intuitiva como la humana, es muy posible que se dé un escenario de exterminio por convivencia de especies que luchan por un mismo recurso de supervivencia; no obstante, también puede pensarse en el caso de un paso evolutivo. La

⁵² Ver Moravec, *The Universal Robot*.

presente lectura se fundamente en el hecho de que hay una guerra entre hombres y máquinas; a su vez, esta guerra tiene rasgos de supervivencia pues la subordinación de una de las especies implica la supervivencia de la otra:

4.1. “(...) no sabemos quién atacó primero, si ellos o nosotros”.⁵³

Para algunos teóricos de la I.A.F., la aparición de máquinas inteligentes puede interpretarse como un paso evolutivo, en la medida en que el material inorgánico que las constituye es mucho más fuerte y resistente que el frágil material orgánico que compone los humanos. De esta manera, el presente signo nos remite a lo que puede denominarse Evolución Artificial Fuerte (E.A.F.): la idea de que las máquinas inteligentes constituyen nuestra futura identidad emergente; esto puede implicar la extinción dramática del Homo sapiens (posiblemente, en los términos planteados por la película), o puede implicar la conversión paulatina del Homo sapiens en Robo sapiens.

La constitución biológica de nuestro cuerpo parece estar expuesta a la mortalidad sorpresiva; además, la capacidad de procesamiento y almacenamiento de nuestro cerebro, en algunos sentidos, es bastante limitada, pues tardamos mucho tiempo en aprender pequeñas piezas de información que, con el paso del tiempo, son fácilmente olvidadas. Estas falencias ponen en riesgo dos de las más grandes ambiciones culturales: la salud y la sabiduría; “todo el mundo quiere sabiduría y bienestar”⁵⁴. La medicina y desarrollos tecnológicos como la terapia genética y la nanotecnología parecen iluminar el camino para posibles superaciones de estas limitaciones; sin embargo, se mantiene la pregunta de si estamos dispuestos a adoptar una radical transformación cultural, al aceptar la robotización de nuestra maquinaria biológica. Teóricos de la I.A.F. aseguran que esta transformación es inevitable; para ellos, el futuro de la vida sólo es posible a través de la aparición de cyborgs. Las máquinas constituyen nuestro futuro evolutivo en la medida en que la maquinaria artificial es mucho más resistente que la débil maquinaria biológica que actualmente es portadora de vida; nuestra maquinaria biológica puede ser frágil, por lo que no satisface nuestras expectativas de longevidad y bienestar. Por este motivo, con el paso del tiempo y los desarrollos tecnológicos, según Minsky, será necesario reemplazar las partes del cuerpo que son más frágiles:

Una vez liberados de las limitaciones de la biología, podremos decidir la duración de nuestras vidas –con la opción de inmortalidad- y elegir, entre otras capacidades inimaginadas.⁵⁵

Minsky está convencido de que la superación de las limitaciones de mortalidad y procesamiento de información no puede darse a través de la biología; así, la obtención de longevidad sólo es posible mediante la implantación de artefactos artificiales en el cuerpo humano. Si bien la medicina ha permitido combatir algunas enfermedades mortales, al parecer,

⁵³ Esta cuestión es resuelta, en alguna medida, en los animatrix titulados The Second Reinnaisance I y The Second Reinnaisance II.

⁵⁴ Ver Minsky, *Will Robots Inherit the Earth?*, pág.1

⁵⁵ Ver Minsky, *Will Robots Inherit the Earth?*, pág.1

no se ha alcanzado la máxima expectativa de vida posible.⁵⁶ Pareciera que la biología alcanza ciertos límites que se pueden superar con la implantación de artefactos:

Ahora bien, aunque el promedio de vida varió en cada uno de estos periodos históricos, la expectativa máxima de vida, según estimativos, era casi idéntica en todos ellos: 116 años aproximadamente. Esto quiere decir que si bien la medicina ha permitido aumentar el promedio de vida, no ha tenido efectos drásticos sobre la expectativa máxima. Aún estamos expuestos a infecciones y daños irreparables en los tejidos vivos y, de esta manera, el paso del tiempo implica disminución de bienestar, debilitamiento de nuestra salud y disminución de nuestras capacidades intelectuales pierden agudeza. Este cuadro de mortalidad, en el que nuestras probabilidades de vida disminuyen drásticamente de manera proporcional al paso de los años, se puede contrastar con el caso de animales cuya dinámica biológica es distinta; esto, para ilustrar el hecho de que no todos los seres de la naturaleza tienden a disminuir sus capacidades por el paso del tiempo:

(...) ciertas especies (incluyendo algunas variedades de peces, tortugas y langostas) no parecen mostrar ningún incremento sistemático de la tasa de mortalidad con la edad. Estos animales parecen morir principalmente por causas externas, como depredadores o falta de comida.⁵⁷

Según Minsky, el primer tipo de variación en nuestra biología tendrá que darse a nivel celular:

El funcionamiento normal de cada célula envuelve miles de procesos químicos, cada uno de los cuales algunas veces comete errores aleatorios (...); estos errores aleatorios ocurren de maneras tan diversas que ningún esquema de bajo nivel puede corregirlos.⁵⁸

En la actualidad los computadores pueden hacer revisiones periódicas de sus sistemas, los que están en uso y los que no, para corregir posibles errores y así mantener su integridad inicial; este mecanismo de supervisión y autocorrección, comúnmente falla en el caso de los humanos. Esto implica la necesidad de hacer variaciones genéticas reemplazando e instalando nuevos genes que garanticen el funcionamiento constante de las supervisiones y correcciones.

Otra técnica para combatir la mortalidad consiste en reemplazar los órganos deficientes con prótesis artificiales; aunque algunos de estos implantes ya se han constituido en una rutina médica,⁵⁹ otros son aún una meta por alcanzar. Por ejemplo, el sistema digestivo cumple la principal función de ser un reactor químico y por este motivo, pensar en su prótesis implica contar con trasplantes a nivel de tejidos y a nivel farmacológico; esto, aún, constituye una dificultad por resolver.

⁵⁶ De acuerdo a Roy Walford, profesor de patología de la Escuela de Medicina de UCLA, el promedio de vida en la Roma antigua era de 22 años, en los países desarrollados era de 50 años para 1900, y hoy es de 75 años. Ver Minsky, *Will Robots Inherit the Earth?*, pág.1

⁵⁷ Ver Minsky, *Will Robots Inherit the Earth?*, pág.3

⁵⁸ Ver Minsky, *Will Robots Inherit the Earth?*, pág.4

⁵⁹ Por su parte, Gregory Stock asegura que los implantes sólo se aceptan por las personas cuando es estrictamente necesario, esto es, cuando está en peligro la vida. Esto plantea la cuestión de si, en el futuro, estaremos dispuestos a aceptar implantes con fines distintos a la estricta supervivencia. *Cfr.* Stock, *Redesigning humans*.

Ahora bien, tal vez la principal limitación en la técnica de los trasplantes se presenta en el caso del cerebro. El trasplante de corazón consiste en un procedimiento en el que, dada cierta compatibilidad entre el donante y el receptor, se procede a reemplazar el órgano; no obstante, esta posibilidad de “reemplazar” no se puede considerar a la ligera en el caso del cerebro. El intercambio de cerebros inevitablemente implica pérdida de conocimiento que redundará en pérdidas y variaciones irreversibles de la personalidad y la identidad individual.⁶⁰ Una posible solución a este problema consiste en reemplazar partes específicas del cerebro con material fibriloso muy fino, que haya sido cultivado en células fetales, para no afectar todo el conjunto cerebral. Al respecto, lo que parece claro es que las prótesis cerebrales plantean un reto para la tecnología actual y, dado que no se tiene aún noticia de su realización exitosa, se tiende a creer que nunca llegará el día en que esto suceda. Parece poco plausible que haya, en el futuro, un reemplazo de áreas del cerebro; se tiende a creer que nunca llegará el día en que los materiales inorgánicos puedan poseer estructuras cognitivas idénticas o superiores a las humanas.

Minsky está convencido de que el primer tipo de limitaciones que nos interesa superar está en el cerebro, pues a nivel de nuestro desarrollo intelectual parece que hemos alcanzado una meseta: “(...) no hay signos de que nos estemos volviendo más inteligentes”⁶¹ con el paso del tiempo. En efecto, es probable que nuestro código genético haya recogido y almacenado conductas relacionadas con información relevante de nuestro entorno para fines de supervivencia y que, en esta medida, nuestro cerebro se haya adaptado para responder a las condiciones del entorno; sin embargo, las capacidades cerebrales no parecen haberse alterado drásticamente en los últimos miles de años. Es posible que no haya pasado aún un periodo de tiempo suficientemente extenso para que la constitución fenotípica del *Homo sapiens* se haya alterado drásticamente y, por este motivo, es posible que no podamos observar aún alteraciones en nuestras capacidades cerebrales; si este es el caso, la tecnología parece ofrecernos un camino evolutivo más corto que el de la típica variación genotípica por vía de mutaciones. Así pues, los transgénicos –animales y alimentos- son una muestra de cómo la modificación genética (sumada a hibridaciones en nanotecnología fundamentada en una hibridación de materiales orgánicos e inorgánicos) ofrecen un camino más ágil para la modificación genotípica de los organismos, que sirve de alternativa a la alteración de características por vía de manipulación de pedigríes. Implantes, hibridaciones de materiales y programaciones de activación y desactivación en genes insertados directamente al genoma, pueden interpretarse como distintas herramientas del mismo mecanismo evolutivo que, mediante cualquier tipo de procedimientos, ha modificado los organismos para que estos respondan a los cambios del entorno.

Incluso los genios más prodigiosos aprenden tan sólo al doble de velocidad que el resto de las personas; esto quiere decir que no hay diferencias drásticas en la velocidad de aprendizaje de una persona promedio y una persona prodigiosa. Nos cuesta trabajo aprender

⁶⁰ La totalidad de lo que somos, de nuestra identidad, está determinada por contenidos de memoria declarativa y no declarativa. Daños en la estructura del hipocampo pueden causar pérdida de los contenidos de ambos tipos de memoria; por ejemplo, se sabe de pacientes que, tras este tipo de daños, dejan de reconocer su propia fotografía. *Cfr. Memory and Brain Systems*, pág. 56.

⁶¹ Minsky, *Will Robots Inherit the Earth?*, pág.5

mucho en poco tiempo porque nuestro cerebro es lento para fijar contenidos específicos de manera conciente; ciertamente el cerebro es mucho más versátil que cualquier máquina computadora al momento de ejecutar otro tipo de cálculos, como los necesarios para trazar planes de acción instantáneos; sin embargo, es más lento para almacenar y fijar contenidos, que una computadora escribiendo en un medio magnético de almacenamiento. “Ciertamente nos ayudaría tener más tiempo, pero la longevidad no es suficiente”⁶² para aprender todo lo que deseamos aprender; por este motivo, la longevidad no sería la solución para el problema de las limitaciones intelectuales. Se puede suponer que el cerebro tiene ciertos límites de almacenamiento,⁶³ por lo cual, aunque tuviéramos siglos para aprender todo lo que quisiéramos aprender, es muy probable que alcancemos un punto en que sería necesario manipular su capacidad. Si bien sabemos muy poco acerca de la naturaleza del cerebro, sabemos que está constituido de muchas áreas y que cada una de ellas cumple funciones distintas y específicas. Por esto, a medida que nuestro conocimiento sobre las funciones de cada área aumente, podremos manipularlas para, probablemente, mejorar drásticamente sus capacidades.

4.2. “Los seres humanos son una enfermedad (...) y nosotros, somos la cura”

En este punto de la manipulación cerebral es que, según Minsky, comenzamos a convertirnos en máquinas y estas comienzan a constituirse en nuestro siguiente paso evolutivo. Como señalé, implantes cerebrales y modificaciones programadas de nuestro genoma, son procedimientos que probablemente comienzan a convertirnos en máquinas. Los humanos nos concebimos como la cúspide del proceso evolutivo aunque este no se haya detenido;⁶⁴ de hecho, es muy probable que la evolución continúe su marcha y, por lo tanto, facilite modificaciones a nuestra especie. Además, las manipulaciones directas de nuestra biología, alternas a las tradicionales pero cumpliendo los mismos fines de la evolución, son procedimientos prácticamente inevitables, por lo cual es necesario comenzar a pensar acerca de “nuestras identidades emergentes”⁶⁵, que serán el resultado de la manipulación sobre nuestro cerebro y nuestro genoma. El presente signo nos remite al hecho de que solamente en la sustitución de parte o de la totalidad de nuestra maquinaria biológica, podremos superar las limitaciones de nuestra estructura. Para Minsky, este futuro consiste en la aparición de cyborgs. Al respecto, lo más importante es reconocer la posibilidad fáctica de este tipo de manipulaciones sobre nuestra biología. Según sabemos, la mayoría de nuestro conocimiento se

⁶² Minsky, *Will Robots Inherit the Earth?*, pág. 5

⁶³ Vale la pena tener en cuenta la diferencia entre capacidad de procesamiento y capacidad de almacenamiento cerebral. La primera se refiere al tipo de información que estamos en capacidad de comprender y utilizar. La segunda se refiere a la capacidad de almacenamiento, más no de procesamiento. Aun si nuestro cerebro tiene una capacidad limitada de almacenamiento, es posible que una capacidad muy limitada de procesamiento nos cree la ilusión de que tenemos una gran capacidad de almacenamiento, porque nos toma mucho tiempo alcanzar el límite de dicha capacidad.

⁶⁴ “Sabemos que el Homo sapiens no es la última palabra en la evolución de primates (...)”. Ver Stock, *Redesigning Humans*, pág.1

⁶⁵ Minsky, *Will Robots Inherit the Earth?*, pág. 6

distribuye en diferentes redes a través de nuestro cerebro; a su vez, estas redes están compuestas de diminutas células nerviosas y de unas estructuras llamadas sinapsis, en cuya configuración se dan los índices de disparo entre neuronas. De estos índices de disparo dependen las posibilidades de procesamiento de las percepciones. Ahora bien, para reemplazar el cerebro de una persona sin que se produzcan mayores traumas, sería necesario reestablecer los tipos específicos de sinapsis establecidas entre dos neuronas. De igual manera, sería necesario conocer y mantener las respuestas específicas ante los neurotransmisores. Estos dos elementos plantean, principalmente, dos retos: el tamaño – la diminuta escala de las reacciones electroquímicas- y la cantidad de sinapsis.

Es muy difícil reconstruir detalladamente un trillón de elementos, sobretodo si estos elementos son microscópicos; no obstante, si se tiene un millón de máquinas trabajando por separado, y cada una de ellas con capacidad de construir miles de partes por segundo, entonces la tarea podría completarse en cuestión de minutos. Este tipo de máquinas, con dimensiones atómicas en su estructura y en sus posibilidades de construcción, son lo que se conoce como nanotecnología⁶⁶:

La mayor parte de manufactura actual está basada en la formación de materiales a gran escala. En contraste, el campo llamado ‘nanotecnología’ apunta a la construcción de materiales y maquinaria mediante la colocación de cada átomo y molécula precisamente donde la queremos.⁶⁷

Mediante este tipo de tecnología se pueden construir copias exactas e idénticas a una pieza original. Según Minsky, este tipo de artefactos implicarán, de manera inevitable, un cambio radical a nivel cultural, pues las posibilidades de máquinas atómicas con programaciones de auto-replicación permitirán la manipulación de cualquier tejido y material biológico.⁶⁸ Esta manipulación, a nivel cerebral, será necesaria porque las tasas de velocidad de aprendizaje, como se señaló líneas atrás, en la actualidad son bastante bajas. Una persona promedio puede aprender un máximo de dos bits por segundo;⁶⁹ si esta tasa se mantiene durante doce horas al día, de manera constante durante cien años, el total de datos aprendidos sería de tres billones de bits, lo cual es menor a la cantidad de información que se puede almacenar actualmente en un disco compacto. Incluso aquellas personas que dicen tener memoria fotográfica, no parecen haber superado estos límites de almacenamiento y procesamiento de información; por ejemplo, las obras completas de Shakespeare abarcan alrededor de 130 millones de bits; esto implica, según las estimaciones de Landauer, un periodo de cuatro años de memorización. En la actualidad ningún cerebro humano ha alcanzado estas capacidades. En síntesis, se puede señalar que la robotización de la biología constituye un camino de superación de nuestras limitaciones y, en este sentido, las máquinas son la cura para las limitaciones y falencias encarnadas en la biología orgánica.

⁶⁶ El término “nanotecnología molecular” se puede usar para describir “sistemas moleculares que pueden construir dispositivos, incluyendo copias de ellos mismos”. Ver Mulhall, *Our Molecular Future*, pág. 38.

⁶⁷ Minsky, *Will Robots Inherit the Earth?*, pág. 7

⁶⁸ Cfr. Mulhall, *Our Molecular Future*.

⁶⁹ Estas estimaciones se hacen con base en las observaciones de Thomas K. Landauer, del centro de investigaciones de Bell Communication. Citado Minsky, *Will Robots Inherit the Earth?*, pág. 8.

- 4.3. “Zión (...) la última ciudad humana. Es el único lugar que nos queda”.
- 4.4. “ (...) la cúspide de su civilización. Digo *su civilización* porque cuando comenzamos a pensar por ustedes, se convirtió en *nuestra* civilización, lo cual es, por supuesto, de lo que esto se trata. Evolución, Morpheus, evolución... como el dinosaurio. Mire por la ventana, tuvieron su tiempo. El futuro es nuestro mundo, Morpheus. El futuro es nuestro tiempo ”

Al pensar en el tipo de máquinas planteado por la E.A.F. –nanomáquinas que permiten la construcción de cyborgs-, es necesario reconsiderar la concepción tradicional de “máquina”. Por lo general pensamos en aparatos o artefactos de acero de grandes dimensiones; por esto, cuando se piensa en “robotización” viene a la mente la idea de robots similares a los que aparecen en obras cinematográficas de ciencia ficción; puede que esto suceda, sin embargo, el futuro inmediato plantea una robotización sin diferencias radicales de apariencia física. Esto quiere decir que en este proceso de robotización podremos ser personas con la misma apariencia actual, pero portadoras de millones de máquinas diminutas en nuestro cuerpo, que nos permitirán manipular y potencializar nuestras propias capacidades.⁷⁰ Con el paso del tiempo esta maquinaria de tamaño atómico continuará reemplazando la maquinaria biológica hasta que, posiblemente, la totalidad de nuestro fenotipo esté constituido de ellas. El presente signo nos remite a la situación en que el Robo sapiens se impone ante la frágil naturaleza del Homo sapiens. En la evolución, las especies más fuertes se imponen ante las más débiles y, casi siempre, hay un escenario de exterminio, sobretodo si hay recursos vitales perseguidos por ambas especies; si la creación de la I.A.F. representa un paso evolutivo, es muy probable que también tengamos que enfrentar dicha dinámica evolucionista de exterminio a manos de nuestra identidad emergente.

⁷⁰ De hecho, desde hace algunos años se han venido incorporando artefactos artificiales al cuerpo humano y la cultura los ha asimilado como *parte de nosotros*. Desde hace varios años somos personas portadoras de máquinas internas; no obstante, las personas que portan implantes artificiales de corazón o de oído, continúan siendo interpretadas como personas normales – no se interpretan como máquinas-; esto muestra el hecho de que, aunque aumente la cantidad y la calidad de los implantes artificiales, hasta que la maquinaria artificial reemplace totalmente la maquinaria biológica, continuaremos siendo personas. Puede asegurarse que mientras estos implantes no alcancen la posibilidad de duplicación del cerebro, no hay motivos para que una persona experimente una ruptura radical en su propia interpretación de *persona* natural; no obstante, sospecho que cuando la nanotecnología posibilite este tipo de implantes cerebrales, estaremos dispuestos a seguir aceptando a sus portadores como personas conscientes e intencionales. No hay razones para trazar una radical diferencia entre “ellos –las máquinas-“ y “nosotros-los humanos”. Esta postura es defendida por Hans Moravec. Al respecto, *Cfr.* Moravec, *Rise of the Robots y Moravec, The Universal Robot.*

5. Máquinas, Evolución Artificial Débil (E.A.D.) y cyborgs

- 5.1. “ (...) a través de la historia humana, hemos dependido de las máquinas para sobrevivir”

Una de las características más importantes del Homo sapiens es su capacidad para interpretar los objetos de distintas maneras. Por este motivo, una simple roca o un trozo de madera pueden interpretarse como una herramienta que, a su vez, puede tener tantos usos como se desee. Estas herramientas, a su vez, en el momento en que se automatizan comienzan a interpretarse como máquinas que están en capacidad de realizar tareas específicas según nuestro antojo. Sin que se asegure que las máquinas nos van a reemplazar, en tanto especie, puede asegurarse que el futuro y la superación de nuestras limitaciones humanas sólo es posible mediante ellas. De esta manera, el presente signo resulta y remite a lo que denomino Evolucion Artificial Débil (E.A.D.).

El punto de partida de la defensa de la E.A.D. es común al de la E.A.F.: los humanos no somos el último estadio de la evolución; esto quiere decir que, de cierta manera, parte o la totalidad de la estructura cognitiva del Homo sapiens tendrá que trascender en el futuro. Como señalé, no hemos observado variaciones drásticas en el fenotipo de nuestra especie; esta es la primera razón para pensar que en el futuro nuestra maquinaria biológica será preservada y que los cambios que la evolución traiga sólo se reflejarán a nivel de nuestra cognición. Nunca antes en la historia habíamos tenido la posibilidad de manipular nuestra biología, de manera tan directa como hoy podemos hacerlo; la terapia génica y la nanotecnología permiten la idea de modificar las características estructurales de nuestra especie. Esto quiere decir que es posible manipular el genoma con el fin de obtener resultados específicos de nuestros descendientes. Básicamente, estas posibilidades parecen abrir la puerta a mejores humanos: más sanos; con mejoras en su aspecto físico y en sus capacidades intelectuales. No obstante, el hecho de que estos procedimientos estén disponibles no quiere decir que van a ser usados y aceptados por todo el mundo. Por ejemplo, el uso de la terapia génica ha sido fuertemente regulado con el propósito de evitar nefastas consecuencias éticas y sociales:

Es ilegal en la mayor parte del mundo probar fetos para el propósito de selección de sexo, pero su práctica es común (...), solo el 32 por ciento de los doctores en los Estados Unidos pensaban que la práctica debería ser ilegal.⁷¹

Si bien se puede tener en cuenta todas las consideraciones religiosas, sociales y morales de la selección extrema de nuestros descendientes, se pueden examinar estas posibilidades a la luz de un contexto más amplio: la evolución. Gregory Stock está convencido de que las posibilidades planteadas por la E.A.F deben observarse con más cautela: “Antes de examinar

⁷¹ Ver Stock, *Redesigning Humans*, pág. 14. Por supuesto, el hecho de que se restrinja este tipo de prácticas no quiere decir que se elimine completamente su uso, al igual que con las drogas ilegales y el aborto.

de manera profunda las mejoras biológicas de los humanos y lo que esto puede traernos, debemos considerar qué verdad se contiene en este sueño de máquinas.”⁷²

Stock reconoce que, indiscutiblemente, los desarrollos en I.A. en bio-informática y en diseño de computadores, producirán efectos en el rediseño de nuestra cognición; no obstante, la cuestión importante es determinar si estos cambios producidos por dichas tecnologías serán incorporados a nuestra naturaleza.

La primera posibilidad a examinar es la de los implantes artificiales. Por ejemplo, en la película *Johnny Mnemonic*, de 1995, se plantea la posibilidad de que mediante un implante cerebral se pueda potencializar drásticamente la capacidad de almacenamiento de una persona –esta posibilidad también fue examinada por Minsky en el apartado anterior. Al respecto, Stock está convencido de que este tipo de cambios no será adoptado:

La gente puede soñar con mejorar sus mentes mediante la implementación de chips en sus sistemas, pero una interfase sofisticada entre nuestro sistema nervioso y el silicio, sería increíblemente compleja.⁷³

Según Stock, gran parte de lo que se concibe como “posibilidades reales de robotización” tiene fundamento en la ciencia ficción de Hollywood y, dado que se ignoran las complejidades reales, entonces las verdaderas posibilidades son absurdas. Toda esta tecnosexuberancia, en términos de Stock, que supuestamente redundará en una transformación completa de los individuos de nuestra especie, es nada más que el resultado de la fe ciega en la ciencia que, en algunos casos, no se logra distinguir de la ficción. La propia y única configuración sináptica, peculiar de cada persona, implica que, a diferencia de lo que supone Minsky, nunca lleguemos a recolectar la información completa de cada configuración individual y, por lo tanto, nunca se pueda duplicar la totalidad de la configuración de un individuo; esto también elimina la posibilidad implantes o trasplantes cerebrales. Una cuestión es contar con nanomáquinas que permitan la construcción de tejidos a escalas minúsculas y otra, muy distinta y aparentemente insuperable, es recolectar la información acerca de la estructura que dicha nanomáquina debe construir. Esto quiere decir que si llega el día en que contemos con nanomáquinas que nos permitan operar en la escala molecular de las sinapsis, necesitaremos saber qué tipo específico de conexiones interneuronales se requieren para una determinada expresión cognitiva. De esta manera, la imposibilidad de la duplicación cerebral, para Stock, está fundamentada en la recolección de la información concerniente a la configuración particular de cada persona. Aunque se conozca la consistencia y función de cada una de las áreas del cerebro, nunca llegaremos a recolectar la información de la configuración sináptica detallada de cada cerebro; por este motivo, la posibilidad de aumentar las capacidades cerebrales en armonía con la estructura cerebral preexistente en cada individuo, es casi nula.

Así pues, la posibilidad de crear extensiones artificiales del cerebro, implicaría conocer la configuración detallada de las conexiones de la persona a la que se implantará el artefacto. Como esta información es prácticamente imposible de recolectar, entonces la fantasía propuesta en *Johnny Mnemonic* está condenada a ser sólo eso, una fantasía. El hecho de contar

⁷² Ver Stock, *Redesigning Humans*, pág. 18

⁷³ Ver Stock, *Redesigning Humans*, pág. 20

con algunos implantes exitosos de electrodos en áreas específicas del cerebro, no quiere decir que estamos ante la posibilidad verdadera de lograr una conexión total de empalme entre el sistema nervioso y extensiones artificiales de materiales inorgánicos: “este tipo de empalmes no es similar a hacer un empalme de líneas telefónicas”.⁷⁴

Ahora bien, piénsese en el uso de corazones artificiales. En la actualidad se cuenta con aparatos artificiales que pueden suprimir las funciones del corazón biológico; además, este tipo de aparatos presenta posibilidades de funcionamiento durante periodos de tiempo más extensos que los periodos de funcionamiento del corazón biológico; no obstante, hasta donde se sabe, nadie ha decidido implantarse un corazón artificial si su propio corazón biológico está funcionando correctamente:

No puedo pensar en ningún aparato que nos sirvieran mejor que nuestro propio corazón, el cual responde tan perfectamente a nuestros cambios de actividades y emociones y está tan bien empalmado al resto de actividades de nuestro sistema circulatorio. Un corazón humano saludable representa el ideal al que cualquier reemplazo debe aspirar, y excepto por un poco de más durabilidad, ninguna mejora es posible (...)⁷⁵.

En general, para Stock, cada uno de los órganos humanos, en condiciones saludables, representa el ideal de funcionamiento de una máquina. Esto quiere decir que un órgano artificial no representa mejoras radicales con respecto al órgano natural saludable; por este motivo, las personas solamente acuden a los implantes artificiales en casos extremos, esto es, cuando se cumple un fin de estricta supervivencia. De esta manera, cuando las personas están saludables prefieren mantener sus órganos biológicos; así, es muy probable que aunque esté disponible tecnología de implante de órganos artificiales, las personas no acudan a ellos si su salud no lo requiere. Es muy probable que lo mismo suceda con un implante de cerebro o de una parte del cerebro. Así, Stock asegura que:

Solamente un verdadero creyente podría imaginar que estamos ante el umbral de obtener (...) nuevos poderes sensoriales y computacionales (...). Aún la más seria conceptualización de una red funcional de electrodos, capaz de generar una suficiente flexibilidad (...) y predictibilidad de empalme con su órgano, es una posibilidad distante. Dada la incompatibilidad natural de la carne y los microchips, desarrollar implantes de cerebro que mejoren el pensamiento será mucho más difícil que construir una inteligencia superhumana que sea puro computador.⁷⁶

Nuestro contacto directo con el mundo se da a través de nuestros sentidos; por este motivo, Stock supone que las primeras mejoras –no necesariamente alteraciones genotípicas o fenotípicas directas- se darán sobre nuestros mecanismos de percepción. Es más probable pensar en aparatos que nos permitan tener super-oídos, super-ojos o super-narices, que en conexiones de banda ancha instaladas en nuestro cerebro. Así pues, la postura defendida por los creyentes en la I.A.F. consiste en una visión futura de *cyborgs*; no obstante, es más posible contar con un cyborg funcional: un *fyborg*. La principal diferencia entre el *cyborg* y el *fyborg* es que

⁷⁴ Cfr Stock, *Redesigning Humans*, pág. 21

⁷⁵ Ver Stock, *Redesigning Humans*, pág. 22

⁷⁶ Ver Stock, *Redesigning Humans*, pág. 22

el primero consiste en la incorporación de artefactos artificiales en nuestro cuerpo; mientras que el segundo consiste en la combinación funcional, en lugar de física, con las máquinas:

Ayudas auditivas, gafas, ropa y teléfonos, aunque físicamente fuera de nosotros, funcionalmente forman parte de nosotros (...), la mayoría de personas con problemas auditivos, por ejemplo, tienen una ayuda auditiva estilo *fyborg*, en lugar de un implante estilo *cyborg*. (...) La gente puede tener un asistente personal digital que le susurre a su oído para recordarle el nombre de las personas, en lugar de un implante aumentador de memoria en su cerebro.⁷⁷

Así pues, en la E.A.D. se asegura que, en la medida en que nuestra naturaleza es biológica, queremos mantenernos de esa manera. No habría un paso evolutivo de lo inorgánico a orgánico. Además, de hecho, ya somos *fyborgs*; no tenemos que esperar al futuro para observar cómo la íntima colaboración con las máquinas permitirá mejorar nuestra forma de vida. En la actualidad usamos ropa, usamos automóviles, teléfonos y otro tipo de aparatos de vanguardia. Lo característico de todos estos aparatos es que podemos suspender su uso cuando queramos; lo mismo sucede con los lentes de contacto o con los audífonos para sordos: si cambiamos de idea con respecto a su uso, tan solo lo suspendemos; por el contrario, un implante colocado mediante cirugía en nuestro cerebro no sería tan fácil de remover:

Que una persona saludable se adhiera un implante cerebral para recibir llamadas (...) es irreal. Los beneficios son muy pocos, los retos son muchos y las alternativas muy numerosas.⁷⁸

Es claro que en la actualidad establecemos una relación de largo plazo con el silicio, sin embargo, según Stock, esto no debe nublar nuestras ideas acerca de nuestro futuro y el de nuestra descendencia pues hay motivos para esperar con cautela la venida de superpoderes de procesamiento en las computadoras. Según Moravec, en la década de los años noventa la capacidad de procesamiento se ha duplicado cada año y esta tendencia se mantendrá en el futuro; sin embargo, algunos investigadores creen lo contrario: en quince años, cuando los elementos computacionales de los chips se aproximen a la miniatura de cinco átomos y los efectos de la física cuántica detengan esta disminución de tamaño, el crecimiento de poder computacional parará o disminuirá drásticamente. La cuestión importante es determinar si en el futuro el crecimiento de las capacidades computacionales continuará aumentando o en algún momento se detendrá.

Ahora bien, con respecto a las consideraciones acerca de la inmortalidad que hace Minsky, Stock asegura que:

Nada es más importante en la trayectoria y significado de nuestras vidas que nuestra mortalidad. A los optimistas de la tecnología les gusta enfocarse en el aumento de la expectativa de vida durante el pasado siglo para asegurar que este continuará ascendiendo en el próximo siglo. Pero en la mayor parte de los anteriores treinta años, este aumento resultó de medidas de salud pública, mejor nutrición y la introducción de antibióticos, y no de la conquista de la biología humana.⁷⁹

⁷⁷ Ver Stock, *Redesigning Humans*, pág. 25

⁷⁸ Ver Stock, *Redesigning Humans*, pág. 27

⁷⁹ Ver Stock, *Redesigning Humans*, pág. 33

Como sea, hay un elemento indudable de evolución que interfiere en lo que somos y seremos como individuos pertenecientes a una especie. Es muy posible que en este camino de evolución estén presentes las máquinas y que distintas hibridaciones de nuestra naturaleza sean consecuencias inevitables de las variaciones genóticas y fenotípicas que nos esperan.

6. ¿Dónde está la verdadera Matriz?

La posibilidad de que en este preciso instante estemos conectados a una simulación se mantiene y es una consecuencia inmediata de la obra; de hecho, no parece haber prueba o argumento alguno para refutar definitivamente esta posibilidad. Ahora bien, el hecho de que estemos desconectados de la simulación no implica necesariamente que estamos en pleno contacto con la realidad primera o última del mundo; aun si estamos desconectados de La Matriz, en este preciso instante no estamos conociendo e interactuando con la ontología del mundo que habitamos.

6.1. Neo: “¿esto no es real?”

Morpheus: “¿qué es real? ¿Cómo defines lo real?”

6.2. “Si estás hablando de lo que puedes sentir, de lo que puedes oler, saborear y ver, entonces lo *real* son simples señales eléctricas interpretadas por tu cerebro”

Todo lo que percibimos, lo percibimos gracias a nuestra constitución biológica. Gracias a los aparatos sensoriales que poseemos podemos ver, oler, escuchar, saborear y sentir texturas. Cada una de estas percepciones es transportada en forma de impulso eléctrico por nuestro sistema nervioso, y analizada por neuronas en áreas específicas de nuestro cerebro. Cuando estos impulsos llegan a nuestro sistema nervioso central, son interpretados mediante comunicaciones que envuelven procesos químicos y eléctricos entre neuronas:

La señal generada por una neurona y transmitida a través de su axón toma la forma de un impulso eléctrico cargado de iones de sodio y potasio, pero es transmitida de célula a célula a través de sinapsis químicas, en la forma de moléculas y neurotransmisores.⁸⁰

Gracias a este proceso ejecutado en la microconstitución del cerebro, se puede asegurar que la interpretación de todo lo que nuestros sentidos nos reportan con respecto a la constitución del mundo en que vivimos, consiste en un procesamiento químico y eléctrico. La riqueza de percepciones, esto es, la riqueza de colores, olores, sonidos, texturas y sabores, supuestamente presentes en el mundo en que vivimos, en realidad es el resultado de los impulsos eléctricos producidos por nuestros aparatos de percepción, que luego son reducidos a índices de disparo sinápticos. La simulación de La Matriz consiste en proporcionar estos impulsos eléctricos directamente al cerebro, de manera que todas estas percepciones sean tan

⁸⁰ Ver Parnavelas, *The Human Brain: 100 Billion Connected Cells*, pág. 19

vívidas como si se recibieran a través de nuestros aparatos de percepción;⁸¹ en realidad, el proceso de La Matriz posibilita un procesamiento de impulsos eléctricos casi idéntico al que sucede en condiciones normales.

Estos dos signos no sólo remiten a la clara posibilidad de que los impulsos que resultarían de nuestras percepciones fueran proporcionados directamente a nuestro cerebro; en realidad estos dos signos remiten al hecho de que *lo real* es el resultado de interpretaciones de estos impulsos, hechas en nuestro cerebro. Aunque estemos desconectados de La Matriz, la constitución del mundo en que vivimos no es más que una construcción epistemológica; lo que conocemos como “constitución del mundo” es el resultado de procesos cerebrales; la constitución del mundo en que vivimos está limitada por nuestras capacidades perceptivas y cerebrales. Esto quiere decir que si tuviéramos distintas capacidades perceptivas y cerebrales, viviríamos en un mundo con otros colores, otros olores, otras texturas, otros sabores y otros sonidos. Así pues, no hay radical distinción entre la realidad simulada, creada por La Matriz, y la realidad creada por nuestras percepciones porque, en últimas, ambas realidades son el resultado de nuestros procesos cerebrales; en el primer caso nuestros aparatos de percepción arrojan impulsos para que el cerebro los interprete y, en el segundo caso, estos impulsos son producidos de manera artificial pero igualmente interpretados en el cerebro. Esta posibilidad se da según las consideraciones acerca de lo que es un *input* y un *output* del sistema intencional que constituye la identidad de los organismos:

La diferencia entre Funcionalismo y Psicofuncionalismo da origen a una diferencia en la especificación de los *inputs* y *outputs*. Los Funcionalistas están limitados a especificar *inputs* y *outputs* que sean una parte plausible del conocimiento de sentido común; los Psicofuncionalistas no tienen tal limitación. Si bien ambos grupos insisten en la especificación física – o al menos no mental- de *inputs* y *outputs*, los Psicofuncionalistas tienen la opción de especificar *inputs* y *outputs* en términos de parámetros internos tales como señales en las neuronas de *input* y de *output*.⁸²

Desde el Psicofuncionalismo se concibe la posibilidad teórica de que no haya radical distinción entre la realidad vivida en La Matriz y la realidad vivida fuera de ella; en últimas, una simulación con poderes causales idénticos a las percepciones, esto es, impulsos eléctricos idénticos a los producidos por nuestros aparatos de percepción, podrían constituir una realidad como la que estamos experimentando en este preciso instante pues la realidad presente tiene su origen en el choque entre la materia y nuestros órganos de percepción;

Es bien sabido que en el mundo hay muchas formas de materia y ninguna de ellas es en sí misma un color, un sonido, una textura, un aroma o un sabor; estas son sensaciones que no solamente dependen de las propiedades de las partículas físicas que intervienen sino de la constitución biológica del aparato de percepción. Así pues, las partículas físicas que hay en el universo son un componente de nuestras sensaciones, pero no constituyen la totalidad de estas. La materia que hay en el mundo adquiere su calidad de color, sonido, textura, aroma o sabor, en la medida en que se pone en contacto con alguno de nuestros órganos de percepción.⁸³

81 En la actualidad ya son comunes los experimentos acerca de estimulación artificial en nuestro cerebro, y creaciones de percepciones como olores y colores, mediante manipulación en áreas específicas.

82 Ver Block, *Las dificultades del funcionalismo*, pág. 112

83 Ver Salcedo, *La constitución epistemológica del universo*, pág. 11.

Así pues, los humanos, al igual que los demás organismos cefalizados y poseedores de un órgano centralizado de procesamiento, vivimos en una versión inventada de lo que en realidad es el universo; vivimos en una versión virtual, en un sueño, en una hipótesis o en una representación de la ontología del mundo,⁸⁴ de la cual no podemos escapar. Lo importante es que aquella versión inventada es lo que solemos llamar *realidad*, porque en tanto especie, es esta la máxima realidad a la que tenemos acceso. La máxima ontología del mundo, acerca de la cual podemos ser conscientes, es una elaboración cerebral de lo que el universo en realidad es. Solemos hablar de lo que sucede en nuestras mentes y de lo que sucede *en la realidad*, distinguiendo y diferenciando ambos tipos de experiencia; solemos hacer una radical distinción entre estos dos escenarios, sin darnos cuenta de que no hay más que una diferencia cualitativa pues ambos escenarios son el resultado de nuestra mente; son elaboraciones de nuestro cerebro. Por ejemplo, decimos que una roca tiene una existencia independiente y no relativa a observador alguno, mientras que los unicornios no tienen existencia independiente sino que solamente existen en la medida en que los inventamos y “habitan” en nuestras mentes. No obstante, aquella roca que existe independiente de nuestra mente, en sentido estricto, no es más que un cúmulo de materia que nunca podremos conocer; un cúmulo de materia desprovista de color, aroma o sabor. Aquella roca, que existe *en la realidad*, supuestamente de manera independiente de nuestra mente, tal como la conocemos es una invención de nuestro cerebro. Cuando la materia que constituye la roca choca con nuestros aparatos perceptivos, y cuando nuestro cerebro procesa dicha información para crear una hipótesis o representación de dicha materia, en realidad está constituyendo una versión epistemológica de aquella materia ontológica que se puso en contacto con nosotros; en aquel momento, el cerebro está inventando y constituyendo un objeto que denominamos *roca*; el

⁸⁴ Por otra parte, si bien digo que el cerebro hace una *hipótesis* o una *representación* de lo que hay en la ontología del mundo, no quiero decir que todas las representaciones mentales que poseemos son idénticas. Esto quiere decir que no son idénticas la representación o hipótesis que mi cerebro hace del computador que tengo frente a mí, en este preciso instante, por una parte, y la representación que tengo en este instante de algún lugar que visité en las últimas vacaciones. Si bien ambas son representaciones que realiza mi cerebro, la primera se da porque *aquí y ahora* hay un objeto de percepción que genera el estímulo; es decir, hay un objeto inmediato de percepción. Es importante aclarar que en estricto sentido, dicho objeto inmediato de percepción no es el computador sino el cúmulo de materia de se encuentra frente a mí.

Para aclarar esta distinción, me parece útil acudir a la diferencia que Searle hace entre *presentaciones* y *representaciones*. Si se desea, para la presente argumentación, puede resultar aclarador denominar a aquellas representaciones que resultan de un objeto de percepción inmediato, como *presentaciones*, y aquellas representaciones que crea mi cerebro por estímulos pasados, como *representaciones*. De esta manera, una *presentación* sería una representación cerebral causada por el estímulo de un objeto inmediato y una *representación* sería una representación cerebral causada por un estímulo pasado. Así pues, las *presentaciones* serían aquellas representaciones o hipótesis cerebrales que se realizan en tiempo real y que resultan de la interacción entre la materia constitutiva del universo y nuestros órganos de percepción, y las *representaciones* serían aquellas representaciones que el cerebro crea con base en estímulos anteriores. Ahora bien, esto no quiere decir, como lo asegura Searle, que las presentaciones nos permitan “acceso directo” al objeto, ni el verdadero “estado de cosas”, pues ni las *presentaciones* ni las *representaciones*, posibilitan cognición alguna de la ontología del universo. [Ver Searle, *Intencionalidad*, pág. 58] Si bien tiene que haber un primer estado del estímulo que posea información acerca de la ontología del universo, en el que se portan las recepciones externas brutas, dicha información, de existir, se procesa a un nivel completamente inconciente, de manera que el único estado del estímulo que es cognoscible es aquel en que la información bruta ha sido interpretada y representada.

cerebro está constituyendo e inventado un objeto que no hay en ningún lugar del universo, porque en el universo sólo hay partículas físicas y energía, carentes de propiedades. Lo mismo sucede para los demás objetos que nos rodean; aunque puede que sean parte de la naturaleza que conocemos y tengan una existencia independiente a nuestra mente, no hay ningún lugar en la ontología del universo en que, en estricto sentido, puedan encontrarse.

Los planetas, incluso aquellos lejanos y de los cuales no tenemos conocimiento de su existencia, el sol, la luna, las estrellas y todos los objetos de percepción, tal como los conocemos, no existen en ningún lugar en la ontología del universo, sólo existen en la constitución epistemológica del universo, en la versión que el Homo sapiens ha creado del universo, gracias a su morfología de especie; sólo existen en la versión en que vivimos y en la que nos movemos día a día;⁸⁵ he aquí los límites de cualquier cognición, representación y pronunciamiento sobre las características del universo. Ahora bien, lo verdaderamente interesante con respecto a la constitución de nuestra “versión” del universo es la incansable capacidad de expansión; vivimos en un mundo con objetos acerca de los cuales nunca tendremos noticia por medio de nuestros órganos de percepción y que resultan de complicadas elaboraciones mentales. En nuestro universo elaborado contamos con objetos físicos -aquellos que podemos percibir de manera directa-, pero también contamos con objetos cuasiteóricos, teóricos⁸⁶ y, gestándose en el límite de la conoscibilidad, objetos estéticos. Hablamos y contamos como objetos de análisis situaciones, sucesos, personajes e incluso mundos que nunca han existido ni existirán y, con esto, nuestro universo cognoscible se expande espacial y temporalmente. Si bien toda la dinámica de creación e interpretación cerebral no nos diferencia de las demás especies de la naturaleza, la capacidad de expansión cognitiva sí parece ser peculiar; así como la incansable asignación signica y la capacidad de interpretaciones y meta-interpretaciones de todo aquello que contamos como objetos.

Lo anterior me lleva a una idea final. Dado que cada individuo, en estricto sentido, cuenta con una configuración sináptica completamente peculiar, así como unas capacidades de percepción peculiares –a pesar de las regularidades morfológicas-, entonces la versión creada del universo, en estricto sentido, es del orden individual. No podemos estar seguros de que todos vemos el color *rojo* de manera idéntica y, de hecho, es muy probable que haya variaciones estrictas, de un individuo a otro, en la percepción de un color y de cualquier sensación. De esta manera, la creación estética sincera, en su realización matérica, constituye un acto de manifestación de esa interpretación individual acerca del mundo. Variaciones de trazos, variaciones en colores e, incluso, variaciones en formas y figuras son claros reflejos de variaciones cognitivas en cuanto a la constitución de las distintas interpretaciones de la realidad, permitidas por distintos cerebros. La creación plástica constituye la manifestación de un rasgo epistemológico del universo que, aunque con cierta diferencia cualitativa, no dista mucho de la creación epistemológica de la realidad actual y presente; obra y realidad son creaciones gestadas en nuestro cerebro.

⁸⁵ Esto, a su vez, quiere decir que cada especie de ser vivo crea una constitución distinta del universo, en la que hay distintos objetos, con propiedades distintas a las que poseen los objetos del universo epistemológico del Homo sapiens.

⁸⁶ Cfr. Salcedo, *La constitución epistemológica del universo*, pág. 12.

Referencias bibliográficas

- Barthes, Roland. *Fragmentos de un discurso amoroso*. México: siglo XXI, 1982.
- Block, Ned. “Las dificultades del funcionalismo”. En *Filosofía de la mente y ciencia Cognitiva*. Compilador: Eduardo Rabossi. España: Paidós, 1995. pp. 105 – 142.
- Deleuze, Gilles y Guattari, Felix. *Mil Mesets. Capitalismo y esquizofrenia*. Valencia: Pre-textos, 1994.
- Dennett, Daniel. “Las Ruedas del conocimiento: el problema estructural de la I.A.”. En *Filosofía de la inteligencia Artificial*. Compiladora: Margarte Boden. México: fondo de cultura económica, 1990. 167 – 192
- Descartes, René. *Meditaciones metafísicas*. Bogotá: Ediciones Universales.
- Fodor, Jerry. *La explicación psicológica*. Madrid: Cátedra, 1991. pp. 159-192.
- Llinaás, Rodolfo. *El cerebro y el mito del yo*. Bogotá: Editorial Norma. 2002.
- McCarthy, John. “Ascribing Mental Qualities to Machines”. Stanfor Artificial Intelligence Lab, Memo AIM-326, pág. 2, *Computer Science Report*, No. STAN-CS-79-725, marzo, 1979.
- Minsky, Marvin L. “Why People Think Computers Can't Think?”. First published in AI Magazine, vol. 3 no. 4, Fall 1982. Reprinted in Technology Review, Nov/Dec 1983, and in The Computer Culture, (Donnelly, Ed.) Associated Univ. Presses, Cranbury NJ, 1985.
- Minsky, Marvin L. “Will Robots Inherit the Earth?”. En *Scientific American*, Oct, 1994.
- Moravec, Hans. *The Universal Robot*. En *Ars Electronica: Facing the Future*, Timothy Druckrey, ed, MIT Press, 1999. pp. 116-123.
- Moravec, Hans. Rise of the Robots. En *Scientific American*, Diciembre, 1999. pp.124-135.
- Mulhall, Douglas. *Our Molecular Future*. New York: Prometheus Books. 2002.
- Parnavelas, John. “the Human Brain: 100 Billion Connected Cells”. En *From Brain to Consciousness?*. Rose, Steve (Ed). New Jersey: Princeton University Press, 1998. pp. 18-32.
- Salcedo Albarán, Eduardo. *El experimento mental de la habitación china: máquinas entre la semántica y la sintaxis*, Borradores de Método, No. 30, Editorial Método, Bogotá, 2005.
- Salcedo Albarán, Eduardo. *La constitución epistemológica del universo*, Borradores de Método, No. 19, Editorial Método, Bogotá, 2004.
- Scholl, B. y Tremoulet, P. “Perceptual causality and animacy”. En *Trends in Cognitive Sciences*, Vol. , No. 8 (2000): pp. 299-309.
- Searle, John. *The Rediscovery of the Mind*. Cambridge: MIT Press, 1994.
- Searle, John. “Mentes y cerebros sin programas”. En *Filosofía de la mente y ciencia Cognitiva*. Compilador: Eduardo Rabossi. España: Paidós, 1995. pp. 413 – 443.
- Schank, Roger y Abelson, Robert. *Guiones, planes, metas y entendimiento. Un estudio de las estructuras del conocimiento humano*. Barcelona: Paidós, 1987.
- Stock, Gregory. *Redesigning Humans. Our inevitable Genetic Future*. New York: Houghton Mifflin Company. 2002.
- Turing, Alan. “La maquinaria de computación y la inteligencia”. En *Filosofía de la inteligencia Artificial*. Compiladora: Margarte Boden. México: Fondo de cultura económica, 1990. pp. 53 – 80.
- Williams, George C. “A Package of Information”. En *The Third Culture*. John Brockman, editor. New York: Touchstone Books, 1996. pp. 38-50.g