

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE LETRAS Y CIENCIAS HUMANAS



**LÓGICA TEMPORAL MÉTRICA APLICADA AL PROBLEMA DE MARCO
[FRAME PROBLEM] EN INTELIGENCIA ARTIFICIAL**

Tesis para obtener el título de **LICENCIADO EN FILOSOFÍA**

AUTOR:

Luis Gonzalo García García

ASESOR:

Pablo Hernando José Quintanilla Pérez-Wicht

Lima, octubre del 2021

Agradecimientos	3
Introducción	4
1. Presentación del problema	5
1.1. Problema lógico inicial	5
1.2. Planteamiento filosófico	8
1.3. Análisis y replanteamiento	13
1.3.1. Problema de la ramificación	14
1.3.2. Problema técnico — axioma de marco	15
1.3.3. Problema de la relevancia	16
1.3.4. Problema del holismo	16
1.3.5. Propuesta de planteamiento	17
1.4. Delimitaciones	20
1.4.1. Complejidad computacional	20
1.4.2. Infalibilidad del sistema	22
1.4.3. Inducción	23
1.4.4. Problema de la parada	25
1.4.5. Abducción	25
1.4.6. Revisión de creencias	26
1.4.7. Problema de la calificación	27
1.5. Criterios de solución	28
1.5.1. Análisis temporal	29
1.5.2. Premisa intuitiva del sentido común	30
1.5.3. Principios rigurosos	30
1.5.4. Generalidad	31
1.5.5. Precisión	31
1.5.6. Atención epistemológica	32
1.5.7. Simpleza	32
2. Lógica temporal métrica	33
2.1. Desarrollo del sistema	33
2.1.1. Enfoque filosófico	34
2.1.2. Presentación	35
2.1.3. Con propiedades	38
2.1.4. Con acciones	40
2.2. Aplicación al problema de marco	41
3. Reflexión y contrastes	42
3.1. Teoría computacional de la mente	43
3.2. Consideraciones sobre el cambio	44

Conclusiones	44
Bibliografía	45



Agradecimientos

Principalmente a mi familia más cercana, mis padres, mi hermana, mis tíos, mis primos, mis abuelos. Cada uno de ellos ha aportado algo con las conversaciones y consejos que les pedía sobre los aspectos que, por muy teóricos que parezcan, siempre pueden mencionarse en una sobremesa.

Por otro lado, agradezco a mi asesor por su interés en verme progresar y consolidar los conocimientos que intento plasmar en estas páginas.

Sinceramente, a cada una de las personas que me han dado ánimos y consejos de variada índole, les agradezco infinitamente.

Resumen

Este trabajo consta de tres capítulos que desarrollan paulatinamente los siguientes puntos. Primero, los acercamientos filosóficos y posturas que difieren en el planteamiento del problema de marco epistemológicamente además de distinguirlo de problemas cercanos con los que se dan posibles confusiones y también delimitar en qué no consiste el problema en cuestión. Segundo, teniendo en cuenta el criterio temporal como fundamental para tratar el problema y dar una solución, presento un sistema de lógica temporal métrica inspirada en el trabajo de Arthur Prior. Tercero, ofrezco algunas últimas consideraciones acerca de cómo los presupuestos de la teoría computacional de la mente nos permiten entender más aguda y ampliamente el problema de marco; esto considerando que el problema despierta un fuerte interés epistemológico y lógico, y, por tanto, mi propuesta está presta para recibir críticas constructivas.

Palabras clave: problema de marco, epistemología, lógica, inteligencia artificial, filosofía de la mente

Introducción

En la literatura teórica de la inteligencia artificial se acuñó un problema concerniente a lo que ordinariamente llamamos sentido común. Cuando nos referimos al sentido común hablamos de los razonamientos que realizamos en situaciones cotidianas. Estos razonamientos del sentido común son estudiados por la inteligencia artificial para modelar y reproducir los mecanismos que empleamos los humanos para procesar la información rápida y eficientemente para resolver problemas del día a día. A lo largo de los años, como podemos imaginar, se plantearon variadas versiones acerca de lo que concebimos cuando apelamos al sentido común, ya sea lógica o intuitivamente. Cada una de estas versiones dieron distintos nombres a la problemática al compararla y analizarla. Y dentro de este amplio espectro del estudio del sentido común, en esta tesis trato una cuestión llamada originalmente problema de marco [*frame problem*].

En el primer capítulo, empiezo narrando históricamente el problema que surgió de una dificultad computacional. Luego, muestro algunos de los planteamientos que se propusieron del problema y analizo las posibles confusiones que se pueden tener al reconocerlo y diferenciarlo de otras problemáticas. Como notaremos, los filósofos se interesaron en el problema ya que vieron propiamente un problema epistemológico al efectuar ciertos procesos prácticos como destapar una botella para beber agua o tender una cama. En tales procesos, la pregunta yace en cómo sabemos qué cosa no se ve afectada por la acción de destapar una botella o tender una cama. Esta situación permite una visión más amplia, y también aguda, de cómo funciona el razonamiento del sentido común. Específicamente, respecto del cambio; es decir, cómo las representaciones del mundo se modifican en cuanto un agente inteligente está en un mundo complejo de cambios constantes. De ese modo, cierro este primer capítulo para establecer cuáles serían algunos criterios por satisfacer para efectuar alguna solución exitosa del problema de marco en vista de la importancia filosófica del mismo o si este efectivamente es imposible de resolver.

Con este propósito, en el segundo capítulo enfatizo el criterio de la temporalidad para analizar y potencialmente solucionar el problema de marco. Procuro aplicar una lógica temporal métrica inspirada en el trabajo de Arthur Prior (1967a) a la ley de inercia del sentido común [*commonsense law of inertia*] postulada por Murray Shanahan (1997) en sus investigaciones formales sobre el problema en cuestión. Por motivos didácticos, adecuó la notación empleada en esta literatura para sintonizar con los intereses analíticos y explicativos de este texto; y así obtener una nueva formalización de dicha ley de inercia. Lo que podemos

adelantar al respecto es que partimos de ciertos presupuestos epistemológicos que tratamos en el primer capítulo, principalmente, el del criterio de la temporalidad. Según este, la resolución del problema de marco requiere, sea cual fuere la postulación formal, dedicar atención especial a la observación del tiempo respecto del cambio.

En el tercer capítulo, ofrezco una reflexión filosófica que permita comprender las posturas y contrastes más resaltantes que el problema de marco ha generado durante décadas tanto en el rubro teórico del razonamiento del sentido común [*commonsense reasoning*] en inteligencia artificial, como en nuestro entendimiento del razonamiento acerca del cambio en las ciencias cognitivas. Mi énfasis se encuentra en ver la importancia que este asunto (aparentemente solo computacional y lógico) tiene para la filosofía de la mente en la actualidad y los posibles futuros debates epistemológicos que cada nuevo desarrollo tecnológico pueda ocasionar. Lo que me propongo hacer en este capítulo es conectar la problemática original con la formalización propuesta desde un punto de vista cognitivo, quiero decir, evaluar los presupuestos filosóficos que podemos tener en mente para ver el problema de marco desde la perspectiva de la temporalidad. De esta manera, tomaré como base ciertas concepciones de la mente o, mejor dicho, cómo entender la mente para el motivo de esta tesis.

Por último, teniendo en cuenta las tres grandes secciones de esta tesis, culmino la presentación y discusión de estas ideas con algunos comentarios, a modo de conclusiones, acerca de lo que esta investigación me ha llevado a pensar sobre el razonamiento acerca del cambio, del sentido común y, ciertamente, el problema de marco.

Capítulo 1. Presentación del problema

En este capítulo abordo el problema de marco teniendo en cuenta tres aspectos importantes: las más importantes perspectivas filosóficas formuladas, los detalles técnicos de la lógica en la que se acuñó y los criterios para enfrentar este problema. De esta manera, presento un amplio panorama de todo lo que implica prestarle atención teórica al problema de marco.

1.1. Problema lógico inicial

Nos remontamos al *cálculo de situaciones* elaborado por McCarthy y Hayes (1969) en su intento de diseñar una lógica de primer orden que permita describir los efectos de las acciones. Entre los estudiosos del tema, se entiende este momento como la tentativa de formalizar el razonamiento acerca del cambio en un agente artificial. En breve, el problema

de marco fue visto inicialmente como la dificultad con la que uno se topa al programar los efectos de una acción sin tener que programar los no-efectos de esta. Esto dado que un cambio en alguna propiedad no necesariamente implica el cambio en otra. Por ejemplo, si programamos en un agente artificial que la acción de presionar el timbre tiene como efecto la producción de un sonido, también tendríamos que indicar que dicha acción eventualmente no cambiará el color del cielo o el material del que está hecha la puerta... Algo que nosotros los seres humanos notamos con inmensa facilidad. En otras palabras, el problema de marco radica en cómo especificar aquello que no cambia cuando ocurre un evento (McCarthy en Brown 1987)¹. En esta línea de pensamiento, y por motivos de rigurosidad teórica, requerimos entender los conceptos de situación, fuente y acción (Silenzi 2015b):

Las *situaciones* (s) pueden entenderse como fotos instantáneas del mundo (Shanahan 1997: 2) o, dicho de otra manera, como un completo estado del universo en un momento dado del tiempo (McCarthy / Hayes 1969: 481). Los *fluentes* (f) pueden ser entendidos como una función cuyo dominio es el espacio *Sit* de situaciones (McCarthy / Hayes 1969: 482), tomando estos distintos valores en diferentes situaciones. Estos pueden también ser pensados como ‘propiedades variables en el tiempo’ (Shanahan 1993: 2). Por último, las *acciones* (a) son un medio por medio de las cuales, al cambiar el valor de algún fuente, se generan nuevas situaciones basadas en las anteriores. (2015b: 63-64)

Parafraseando esta contundente cita, lo que amerita la comprensión formal del problema inicial invita a imaginar un mundo en el que hay momentos puntuales (situaciones) en los que las distintas propiedades (fluentes) se ven afectadas en cuanto un agente interviene (acciones).

En este sentido, tomando en cuenta la notación sugerida por Shanahan (2016), empecemos por simbolizar un ejemplo, y así aclarar el funcionamiento del formalismo propuesto por McCarthy y Hayes: el cálculo de situaciones [*situation calculus*].

Entonces, tenemos:

- un objeto x cualquiera;
- las propiedades o fluentes COLOR (x, a) y POSICIÓN (x, a);
- en las situaciones sucesivas a, b y c .

Y el estado inicial de nuestro mundo es descrito por los siguientes enunciados:

- (1) COLOR-1 (x, a)
- (2) POSICIÓN-1 (x, a)

¹“The frame problem is that of specifying what does not change when an event occurs” (*Traducción propia*)

Que se leen: (1) el objeto x tiene la propiedad COLOR-1 en la situación a ; y (2) un objeto x tiene la propiedad POSICIÓN-1 en la situación a .

Luego, ejecutamos una acción, por decir, PINTAR el objeto x . Por lo tanto, el color cambia de esta primera situación a hacia una segunda situación b :

$$(3) \text{ PINTAR}(x, a) \rightarrow \text{COLOR-2}(x, b)$$

De esta manera, (3) se lee: de PINTAR el objeto x en la situación a se sigue que el objeto x toma la propiedad COLOR-2 en la siguiente situación inmediata b . Sin embargo, de tal acción no se sigue que POSICIÓN-1 (x, a) se mantenga, es decir, que el objeto x permanezca en la misma posición de a en b .

Por lo tanto, es necesario agregar una condición que haga explícito el hecho de que la acción PINTAR, si bien cambia la propiedad COLOR-1, no afecta a POSICIÓN-1:

$$(4) \text{ POSICIÓN-1}(x, a) \text{ permanece como POSICIÓN-1}(x, b) \text{ luego de PINTAR}(x, a)$$

Y así queda establecido que la acción PINTAR no afecta a la posición del objeto x , que queda como POSICIÓN-1 en la situación b , tal como inicialmente estaba en la situación a .

Este tipo de enunciado que evita inferir la influencia de una acción precisa en todas las demás propiedades que no han sido incluidas en la notación de la ejecución de dicha acción fue denominado axioma de marco [*frame axiom*] en el texto que presenta el cálculo de situaciones y, por consiguiente, el problema de marco.

Por motivos explicativos, este ejemplo puede ampliarse ejecutando una segunda acción: MOVER. Entonces, movemos el objeto:

$$(5) \text{ MOVER}(x, b) \rightarrow \text{POSICIÓN-2}(x, c)$$

Con la misma forma que (3), (5) se lee: de MOVER el objeto x en la situación b se sigue que el objeto x toma la propiedad POSICIÓN-2 en la siguiente situación inmediata c . No obstante, ahora se requeriría de una segunda condición o axioma de marco:

$$(6) \text{ COLOR-2}(x, b) \text{ permanece como COLOR-2}(x, c) \text{ luego de MOVER}(x, b)$$

Y así queda especificado que un cambio en la posición no afecta al color del objeto x .

Como es de notar, este ejemplo toma un mundo en el que hay solo un objeto con dos propiedades y ordenadamente se ha decidido ejecutar dos acciones, una después de la otra. Empero, este mundo es inmensamente más simple que los escenarios a los que se enfrentaría un agente (artificial o humano) en la vida real. Algo importante para tener en cuenta es que

si se tuviera un gran número de acciones a ejecutar en secuencia se tendría una gran cantidad de condiciones o axiomas de marco debido a que estos determinan aquello que no cambia, ya que ciertas acciones no afectan a algunas propiedades o funciones [*fluents*]. De hecho, con n acciones y m funciones deberíamos escribir mn condiciones de marco (McCarthy, J. y Hayes, P. 1969: 31). Por esta razón, cuanto más propiedades y acciones haya, mayor será el número de axiomas de marco a programar. Y tal es la dificultad en cuestión. El procesamiento de tanta información para determinar aquello que no cambia en el mundo es un problema implícito en el problema de marco, que trato en la sección 1.3.

1.2. Planteamiento filosófico

Partiendo del problema computacional, donde yace la cuestión es en la dificultad para representar los efectos de las acciones, es decir, cómo señalarlos en un programa o en un sistema. Sin embargo, posteriormente causó interés filosófico el hecho de que se revelaba un profundo e interesante problema epistemológico, según Dennett, pasado desapercibido por generaciones de filósofos (1987). Dicho así, cada vez que alguien aborda el problema de marco, se topa con la primera dificultad de su planteamiento (Sprevak 2005, Silenzi 2015b), ya que no hay un consenso acerca de cuál es el problema fundamental. En esta sección veremos algunas maneras de definir el problema tomando en cuenta los criterios que cada versión aporta a la discusión.

Algo interesante a tener en mente es que “la primera mención significativa del problema de marco en la literatura filosófica fue hecha por Dennett (1978)” (Shanahan 2016)². Luego, un gran número de filósofos abordaron el tema (Fodor 1983, Brown 1987, Pylyshyn 1987, Dennett 1987, Ford y Hayes 1991, Ford y Pylyshyn 1996, entre otros). Cada uno de ellos remarcaban distintos aspectos del problema, no obstante, no lograban estar de acuerdo sobre el núcleo teórico que se necesitaba. Para empezar, a pesar de que la mayoría de los filósofos mantenían una postura precedida por la teoría computacional de la mente, había otros como Dreyfus (1972) o Fodor (1987) que argumentaban contra la versión clásica de esta teoría. Cabe especificar que la versión clásica de la teoría computacional de la mente nos dice que la mente *es* literalmente una computadora (Rescorla 2020). La crítica mencionada por parte de Dreyfus y también por Fodor se manifiesta fundamentada en que la comprensión de los fenómenos más llamativos y curiosos de la mente ocurren al manejar la información de un

² “The first significant mention of the frame problem in the philosophical literature was made by Dennett” (*Tr. p.*)

modo holista que no se podría comparar con los procesos computacionales desarrollados hasta el momento (en que ambos publicaron sus ideas)³.

Cada autor aporta una luz más acerca del problema que en un inicio no parecía que generaría tanta discusión más allá de la programación. Conforme los avances en la tecnología lo permitían, nuevas lógicas eran diseñadas para tratar el problema. De este modo, una tendencia común es que muchos teóricos lo plantearon “como un problema de [la] relevancia (Glymour 1987, Crockett 1994, Dreyfus 1979, Schneider 2007, Pinker 2005, Shanahan 2009, Ludwig y Schneider 2008, entre otros)” (Silenzi 2015b: 56). Dicho en términos simples, un primer planteamiento es el siguiente:

Problema de marco 1: Dada una cantidad de información y una tarea a ejecutar, ¿qué información es relevante para llevar a cabo tal tarea?

Esta forma de verlo contiene como principal dificultad la determinación de un criterio que indique qué información es la relevante para el caso que se necesite. Dicho criterio debe saber reconocer y relacionar la información que se obtenga frente a la acción a efectuar y al contexto. Empero, evidentemente cualquier cosa puede guardar relación con cualquier otra cosa. Y las relaciones que puede efectuar el programa son puramente inferenciales, a menos que se haya registrado información previa como las condiciones o axiomas de marco [*frame axiom*] descritos en la sección 1.1. Por lo tanto, no habría modo de programar tal criterio desde este primer planteamiento. Y, en caso de que se pudiera, tendría que satisfacer al menos la mínima información necesaria para ejecutar exitosamente una acción en el mundo.

Otra manera de formularlo viene desde un área de inteligencia artificial llamada representación del conocimiento, es decir, cómo circunscribir la información que se tiene del mundo en el sistema (la mente, en nuestro caso). Siguiendo a Dennett, el problema radica en cómo representar la información empírica que se ha obtenido esforzadamente, de tal modo que esta pueda ser eficazmente explicitada cuando se la necesite (1987: 52-53)⁴. A este respecto, tener un sistema que ignore la mayoría de las cosas que sabe y que opere con una sección bien escogida de su conocimiento en cada instante sería conveniente. Que la información sea bien escogida no significa que lo sea por una búsqueda exhaustiva, puesto

³ Ver sección 3.

⁴ “(...) how to represent (so it can be *used*) all the hard-won empirical information –a problem that arises independently of the truth value, probability, warranted assertability, or subjective certainty of any of it. Even if you have excellent *knowledge* (and not mere belief) about the changing world, how can this knowledge be represented so that it can be efficaciously brought to bear?” (*Tr. p.*)

que en ese caso se sobrepasaría el límite temporal requerido para efectuar tal o cual acción adecuada para la situación. Cuando nos referimos al límite temporal estamos hablando del poder computacional de un sistema. En pocas palabras, se tiene en cuenta dos factores para medirlo: la cantidad de computaciones y el tiempo que demanda efectuarlas. Así, un sistema es más poderoso computacionalmente si puede efectuar la mayor cantidad de computaciones en el menor tiempo posible. Dicho de esta manera, ya podemos contar con un segundo planteamiento de la cuestión:

Problema de marco 2: ¿Cómo se le puede dar a un sistema reglas para ignorar lo que debería ignorar bajo una amplia variedad de circunstancias diferentes en un ambiente complejo de acción? (Dennett 1987: 55)⁵

Aparentemente esta manera de definir el problema es casi idéntica a la primera formulación, sin embargo, cuando se lleva a la práctica esta visión, nos topamos con distintos mecanismos de programación para la primera y la segunda. En otras palabras, entre decidir qué es relevante y qué ignorar notamos que en el primer caso el énfasis del criterio está en la información que se utiliza y, en el segundo, en la que no se utiliza. Por este motivo, ambos criterios se representarían o digitarían de manera diferente en un programa. Como vemos, “comprender cómo se llevan a cabo algunas de las actividades cognitivas más interesantes de la mente humana (realizar inferencias, razonar, tomar decisiones, resolver problemas, etc.) es comprender, a fin de cuentas, el problema de marco” (Fodor en Silenzi 2015b: 54).

Manteniendo una postura conciliadora de los planteamientos, es factible considerar que los intereses epistemológicos del problema se confrontan constantemente con su aplicación computacional, ya que este es un problema de inteligencia artificial en todo sentido. Esto es por lo que algunos diferencian ambos aspectos como problemas que poseen distinta aplicación a la realidad, ya que, en el primer caso, sería a la mente humana y, en el segundo caso, a un programa. Al respecto, Shanahan (2016) distingue los siguientes planteamientos⁶:

⁵ “What is needed is a system that genuinely *ignores* most of what it knows and operates with a well-chosen portion of its knowledge at any moment. Well-chosen, but not chosen by exhaustive consideration. How, though, can you give a system *rules* for ignoring –or better, since explicit rule-following is not the problem, how can you design a system that reliably ignores what it ought to ignore under a wide variety of different circumstances in a complex action environment?” (*Tr. p.*)

⁶ “(...) it is possible to discern an epistemological frame problem, and to distinguish it from a computational counterpart. The epistemological problem is this: How is it possible for holistic, open-ended, context-sensitive relevance to be captured by a set of propositional, language-like representations of the sort used in classical AI? The computational counterpart to the epistemological problem is this. How could an inference process tractably be confined to just what is relevant, given that relevance is holistic, open-ended, and context-sensitive?” (*Tr. p.*)

Problema de marco 3a —aspecto epistemológico: ¿Cómo es posible que la relevancia holística, abierta, sensible al contexto, sea capturada por un conjunto de representaciones proposicionales, similares al lenguaje, del tipo usado en la inteligencia artificial clásica?

Problema de marco 3b —aspecto computacional: ¿Cómo puede un proceso inferencial ser reducido manejablemente solo a lo que es relevante, dado que la relevancia es holística, abierta y sensible al contexto?

En ambas situaciones se aborda el problema de la relevancia, tal como vimos con el primer planteamiento del problema de marco, pero esta vez se diferencia el sentido de la pregunta. El aspecto epistemológico se orienta desde las cualidades de la información relevante hacia su potencial modelización. En cambio, el aspecto computacional se dirige de los procesos inferenciales hacia el manejo de información cuyas características la hacen difícil de manejar. Este tercer planteamiento resalta las características o atributos que tiene información considerada relevante. Estas características son el holismo, el estar abierta y ser sensible al contexto. De hecho, programar un criterio que juzgue la información teniendo en cuenta estos atributos es una tarea difícil. Así, pues, cuando nos referimos al holismo, cabe resaltar que se opone a la suma especificación o particularización de la información (mejor dicho, de una propiedad o función), puesto que se logra establecer una generalidad o un patrón según el cual se organiza la información. En tanto hablamos de estar abierta y ser sensible al contexto, nos referimos a que es posible relacionar con gran facilidad lo que nos proporciona una acción específica y la situación general en la que esta se lleva a cabo, y, además, se trata de que la información se vaya actualizando conforme el tiempo pasa y las situaciones se desarrollan en el mundo.

Por ejemplo, cuando vamos a vestirnos para salir a la calle por la mañana. En primera instancia hay ciertas variables para tener en cuenta al decidir, como la temperatura del ambiente, la combinación de los colores e incluso la ligereza de las prendas. Cada una de estas nos proporciona una correlación entre la acción que vamos a efectuar (vestirnos con la ropa) y el contexto en el cual se desarrollará (un día por la mañana en nuestra casa). Si empezamos a colocarnos la ropa fresca debido a que vemos el Sol por la ventana, pero repentinamente vemos un anuncio televisivo de fuerte lluvia por la tarde, nuestras acciones cambiarán de modo que elegiremos ropa más abrigadora para que podamos estar preparados. De pronto, podemos recibir una llamada que nos comunica el fallecimiento de alguien y la

cordial invitación a su velorio. Entonces, cambiaremos nuevamente de decisión y optaremos por colocarnos vestimentas negras para la ocasión. Y así sucesivamente. De esta manera, cada vez que surge nueva información, efectuamos razonamientos que cambian el curso de acción por el que nos habíamos inclinado en vista de que el contexto se modifica. Dicho en otras palabras, si nuestra representación del mundo tiene cambios determinantes para tomar un curso de acción, nos vemos con la cuestión de qué los hace tan importantes o, para emplear el término técnico, cómo es que se tornan relevantes tales cambios. Algo curioso al ver este tipo de situaciones es que corroboran la idea de Fodor de que “cuanto más global sea un proceso cognitivo, menos oportunidad de entenderlo” (en Silenzi 2011a: 434).

Podemos colocar un cuarto planteamiento propuesto por Morgenstern (1996). Como dato interesante, ella expresa que el trabajo en el problema de marco presenta un patrón común entre todos los investigadores. Bajo su punto de vista, este se aborda empezando con el problema de un juguete que se supone ejemplifica una instancia del problema de marco, luego, se propone una solución, sin embargo, una nueva variante del problema del juguete es propuesta y, finalmente, se descubre que la solución anterior no puede resolver el nuevo problema del juguete, y así sucesivamente (*idem*: 4)⁷. El tono irónico critica el hecho de que no se logró progresar mucho en el desarrollo de soluciones al problema. Su planteamiento va de la siguiente manera:

Problema de marco 4: Es el problema de determinar qué permanece igual acerca del mundo en tanto que pasa el tiempo y acciones son ejecutadas sin tener que explicitar todas las cosas que permanecen iguales⁸ (*idem*: 5).

Esta perspectiva no toma en cuenta explícitamente a los efectos de las acciones, sino a los no-efectos de ella, siendo fiel al planteamiento del problema inicial. Por ello, es resaltado el hecho de que las partes que no cambian en el mundo son objeto de atención en el problema. Es decir, lo que importa es cómo determinar lo que no cambia en el mundo mientras suceden acciones.

De estos cuatro planteamientos, probablemente el más aceptable es el último debido a que conserva la preocupación original del problema y especifica que el énfasis se encuentra en

⁷“(…) the work on the frame problem seems Sisyphean in its nature. The general pattern of the research looks something like this: A toy problem, presumably meant to be a characteristic instance of the frame problem, is introduced; a solution is proposed; a new variant on the toy problem is introduced; it is discovered that the old solution cannot solve the new toy problem, and so on” (*Tr. p.*)

⁸“It is the problem of determining what stays the same about the world as time passes and actions are performed – without having to explicitly state all the things that stay the same” (*Tr. p.*)

aquello que no se ve afectado por una acción. Debido a esto, es fácil notar que mucha más información no se encuentra afectada por la acción que aquella que sí se ve afectada. No obstante, los otros tres enfatizan puntos importantes para entender el problema de marco. Estos son: relevancia de información, ignorar información innecesaria y características de la información relevante, respectivamente.

1.3. Análisis y replanteamiento

Teniendo en cuenta las formulaciones propuestas del problema de marco, en esta sección procuro integrar los aspectos resaltados con el fin de ofrecer un nuevo planteamiento (que quizá no tiene nada de nuevo) para poder tener una idea más clara y aguda del problema. De este modo, resulta conveniente dar como primer gran paso definir con precisión en qué consiste este problema para encaminar las soluciones tanto lógica como filosóficamente.

La imagen general del problema inicial involucra cambio, acciones y efectos. Y sumando los conceptos propuestos en los siguientes planteamientos, tenemos relevancia, ignorar, holismo, información abierta, sensibilidad al contexto, permanencia. Además, siendo rigurosos con el lenguaje que empleamos para describir el problema, conviene partir de que en el mundo solo hay objetos y propiedades. Ahora, con fines explicativos, digamos que el problema de marco se encuentra respecto de las propiedades. Entonces, la tarea que efectuaré consiste en evaluar con rigurosidad las condiciones que presenta el problema para el tratamiento de las propiedades.

Cuando hay un cambio en el mundo nos referimos a que una propiedad ha sido reemplazada por otra. Por ejemplo, retomando el caso del *cálculo de situaciones* expuesto en la sección 1.1., si un objeto rojo es pintado de azul, vemos que la propiedad ROJO (x, a) ha sido reemplazada por AZUL (x, b); siendo x el mismo objeto en cuestión; a , la primera situación; y b , la segunda.

Tomando una perspectiva más abstracta, el cambio en el mundo es representado como el paso de COLOR-1 (x, a) a COLOR-2 (x, b). En este ejemplo tomamos en cuenta al color, pero pudo ser la posición, forma, textura, olor, etc. De este modo se representa un cambio en el mundo.

Sin embargo, otro punto a considerar es la acción que causó dicho cambio, es decir, por qué razón la primera propiedad fue reemplazada por la segunda. Continuando con el ejemplo anterior, la acción ejecutada es PINTAR (x, a), siendo x el objeto y a la primera situación.

Y el efecto de dicha acción sobre COLOR-1 (x, a) es COLOR-2 (x, b), siendo b la segunda situación. Aquí notamos que hay una relación causal entre dicha acción y la segunda propiedad. Por lo tanto, decimos que de efectuar la acción PINTAR (x, a) se sigue COLOR-2 (x, b). Tal es una inferencia simple que no requiere de demasiadas computaciones por parte del sistema. Esta es una visión microscópica de un cambio generado en el mundo. No obstante, en este punto surgen las mayores complicaciones.

1.3.1. Problema de la ramificación

No podemos saber con exactitud si la acción PINTAR (x, a) tiene como único efecto COLOR-2 (x, b), puesto que es posible que de ella también se siga un cambio en la posición o la forma del objeto x . Este es conocido como el problema de la ramificación [*ramification problem*], según el cual se hace difícil evaluar los efectos implícitos de una acción como los efectos secundarios o terciarios. Por decir, que hayamos colocado nuestra atención en la propiedad COLOR-1 (x, a) no implica que otra propiedad como POSICIÓN-1 (x, a) no haya sido afectada por la acción y sea reemplazada por POSICIÓN-2 (x, b). Esta dificultad consiste en examinar todos los efectos de la acción ejecutada (McDermott 1987: 106-107), y ello sería bastante exigente en un mundo complejo, ya que es posible realizar una infinidad de inferencias a partir de tan solo una acción y una consecuencia circunscrita en el sistema.

Empero, el problema de marco no se enfoca en esta dificultad en vista de que se interesa, más bien, en las propiedades que no resultan afectadas por una acción. Eso es así debido a que toma como partida la idea de que el número de efectos de una acción es considerablemente menor a los no-efectos, es decir, hay menos propiedades que cambian en el mundo respecto de las que no cambian cuando se realiza una acción.

Por ejemplo, las famosas y curiosas máquinas de Rube-Goldberg, en las que una simple acción que activa la compleja dinámica dispara una secuencia de movimientos que, a su vez, afectan a las distintas secciones o partes de la construcción mecánica. De este modo, es posible discernir entre cuál es el efecto inmediato de aquella primera acción de los efectos subsecuentes tales (como suele pasar en la mayoría de estas máquinas) como freír un huevo, patear una pelota, barrer el piso, entre otras tantas posibles acciones que producen cambios: el huevo que antes estaba entero, ahora está quebrado y sobre la sartén; la pelota que se encontraba estática, ahora se encuentra en movimiento; el polvo que yacía sobre la superficie de la habitación, ahora está arrumado en un pequeño montículo. Entre otros tantos casos

imaginables en donde cualquier agente inteligente lograría notar una diferencia considerable entre un efecto inmediato (o primario) de uno secundario, terciario, y así sucesivamente.

Otro ejemplo más interesante aún es el que incita la compra de una galleta. Si examinamos la logística involucrada, cabe resaltar que la acción de comprar una galleta en una tienda tiene como primera consecuencia la pérdida de un nuevo sol o el equivalente monetario del producto. Luego, es válido inferir que el inventario de dicha tienda tendrá un ítem menos, además de que físicamente el paquete de galletas (es decir, el objeto) ya no estará presente en el establecimiento. Ampliando más este caso, inclusive, en algún momento, quien compra la galleta habrá terminado de consumirla y, por lo tanto, quedará el empaque vacío. Ante la necesidad de desecharlo, algún tacho de basura próximamente incluirá esa envoltura en su interior. Posteriormente, es posible que quien recoja tal basura se vea curiosamente persuadido por el nombre que encuentra escrito sobre esa envoltura y, no solo eso, sino que le traiga recuerdos de su infancia y empiece a llorar, o que le provoque comprar un par de paquetes de esas galletas, y así sucesivamente, entre otras tantas opciones imaginables. De tal modo, la sorpresa que encuentra el recogedor de la basura es una consecuencia no inmediata, o no primaria, de la acción efectuada por quien compró la mencionada galleta al inicio del ejemplo. En este tipo de inferencias consiste el problema de la ramificación, que no es directamente el problema de marco, pero guarda una relación bastante estrecha.

1.3.2. Problema técnico – axioma de marco

Siguiendo y remarcando el punto anterior, la mayor dificultad que le interesa resolver al problema de marco consiste en cómo determinar las propiedades que no se ven influenciadas por una acción. Usando el ejemplo abstracto que nos acompaña desde la sección inicial, sabemos que de ejecutar la acción PINTAR (x, a) sobre la propiedad COLOR-1 (x, a) se infiere que esta es reemplazada por la propiedad COLOR-2 (x, b), es decir, COLOR-2 (x, b) es un efecto de la acción PINTAR (x, a). Sin embargo, tan solo de esta información no se sigue que POSICIÓN-1 (x, a) sea reemplazada por POSICIÓN-2 (x, b). En términos generales, no es válido inferir que un cambio en el color de un objeto directamente afecte a la posición de este a menos que así sea explicitado. Por tanto, es necesario programar en el sistema que la propiedad POSICIÓN-1 (x, a) se mantiene luego de que se ejecute PINTAR (x, a). De esta manera se infiere que luego de PINTAR (x, a), la representación del mundo es COLOR-2 (x, b) y POSICIÓN-1 (x, b). Así es como se le indica al sistema que la propiedad POSICIÓN-1 (x, a) se mantuvo como POSICIÓN-1 (x, b) sin verse afectada por

PINTAR (x, a). La explicitación de que una propiedad no se ve afectada por la ejecución de una acción es el llamado axioma de marco, que fue mencionado en la sección 1.1.

1.3.3. Problema de la relevancia

Una vez que ya encontramos un modo de expresar que una propiedad no resulta afectada por una acción, corresponde examinar qué propiedades, de todas las que hay en el mundo, no se ven afectadas por tal acción. Si el mundo tuviera un número bastante pequeño de propiedades y de acciones, sería más práctico y factible que el sistema logre computar la totalidad de axiomas de marco en un tiempo aceptable para representar las nuevas situaciones que surgen en el mundo en base a las acciones que ocurren en él. Empero, el mundo al que se busca exponer el sistema de inteligencia artificial es complejo, pues presenta una vasta cantidad tanto de propiedades como de acciones (que incluso son simultáneas). Por esta razón muchos teóricos⁹ identifican al problema de marco como el problema de la relevancia, ya que nos topamos con dos grandes dificultades. Una es la de programar la inmensa cantidad de axiomas de marco que indicarían las propiedades que no se ven afectadas por cada acción específicamente. La otra consiste en que, si se superara la primera dificultad, y si fuera posible programar todos los axiomas de marco requeridos para representar el cambio en el mundo, cómo podríamos evitar el tener que evaluar cada uno de los axiomas de marco para representar lo más eficientemente posible el mundo cuando se ejecutan acciones. Esta es la misma razón por la cual hay otros teóricos (Dennett 1987) que creen que es necesario un criterio que permita ignorar la información innecesaria al representar el mundo. En ambos casos, la cantidad de información para procesar es la tarea complicada que requiere simplificarse en cuanto se tengan métodos para ignorar eficazmente lo irrelevante para la situación.

1.3.4. Problema del holismo

Si bien ya analizamos el carácter computacional del problema de marco, ahora toca concentrarnos en la información que se procesa, es decir, conocer cuáles son las características de las propiedades con las que se representa el mundo. Que la información sea holista significa que las relaciones entre las propiedades son interdependientes, ya que es posible encontrar una conexión entre una propiedad y cualquier otra. De este modo, las relaciones lógicas entre las propiedades son múltiples y no sería posible encontrar alguna propiedad que carezca de relación con alguna otra propiedad. La dificultad con esto radica

⁹ Ver sección 1.2.

en que el sistema requeriría evaluar todas las relaciones posibles de una propiedad con otras antes de poder representar la nueva situación en el mundo y así, en efecto, también detectar las relaciones causales entre una acción y todas las propiedades en el mundo.

Por otra parte, coincido con la formulación sobre el holismo aportada por Samuels, ya que lo integra con el problema de la relevancia (2010: 284). De esta forma, siguiendo a este autor, considero que una buena manera de plantear el problema del holismo, *mutatis mutandis*, es la siguiente: dado que casi cualquier información puede ser relevante para la evaluación de otra, entonces cómo determinaríamos qué información es relevante en un instante específico. Siguiendo este razonamiento, puede afirmarse razonablemente que todas las propiedades se verían afectadas de alguna u otra forma por una acción. Pero, de ser así, sería imposible representar siquiera un instante en el mundo luego de que ocurra una sola acción dada la infinidad de propiedades involucradas. No obstante, nosotros, los seres humanos, somos capaces de encontrar con una gran aproximación las relaciones entre una acción y las propiedades que se ven comprometidas con esta. Por consiguiente, cuando se ve a la información como sensible al contexto, también se evoca esta complicación puesto que nos estamos refiriendo a que la información que poseemos *acerca del* mundo se adapta a la situación expuesta *en el* mundo, es decir, las propiedades cambian conforme se presentan nuevas situaciones tan solo por el hecho de estar en dichas situaciones y, además, se toman en cuenta las acciones que afectan a tales propiedades. Por esta razón, personas como Morgenstern (1996) plantean el problema de marco como la determinación de las propiedades que permanecen iguales al ejecutarse una acción en el paso del tiempo.

1.3.5. Propuesta de planteamiento

Luego de haber ampliado estos conceptos, ofrezco una formulación motivada por la primacía del manejo de propiedades en el lenguaje lógico de primer orden que, viéndolo en perspectiva, se adecúa o encaja con los intereses teóricos de este problema. Así, corroboro la idea de que para el problema de marco es crucial entender la manera en que asociamos las propiedades a los objetos, puesto que un tercer elemento protagónico, a saber, las acciones, modifican estas relaciones que los objetos guardan con las propiedades. En otras palabras, la dinámica de cambio de propiedades de un objeto no se da gratuitamente (según el planteamiento original del problema de marco), sino que se debe a la participación de alguna acción o acciones que, a su vez, son ejecutadas por un agente artificial. Como es de conocimiento general, nosotros, los seres humanos, reconocemos, la mayor parte de las

veces, qué cambia, o cómo un objeto cambia alguno de sus atributos, como consecuencia de alguna acción.

Por ejemplo, una persona no necesita demasiada explicación para discernir que una broca gira al presionar el botón de encendido en un taladro debido a la acción de presionar dicho botón. Esta es una trivialidad para un agente humano, sin embargo, en caso se ignore otra información sobre el taladro, la persona podría descubrir que tal instrumento, a su vez, cuenta con un mecanismo llamativo que enciende un ventilador en la habitación contigua. En términos eléctricos, la conexión causal entre presionar el botón y el giro de la broca es evidente para cualquier persona, empero, la conexión igualmente causal entre el accionar de tal botón y el encendido del mencionado ventilador es desconocida por casi cualquier agente que intente accionar ese taladro. Por lo tanto, si entendemos este ejemplo en términos de objetos, propiedades y acciones, podemos obtener lo siguiente. Los objetos centrales son, como ya fueron mencionados, la broca, el taladro, el botón y el ventilador. Las propiedades, en la situación inicial a , son ESTÁTICO (broca, a), APAGADO (taladro, a), SOBRESALIDO (botón, a), APAGADO (ventilador, a). La acción central aquí es PRESIONAR (botón, a). Y las subsecuentes propiedades, en una segunda situación b , son DINÁMICO (broca, b), ENCENDIDO (taladro, b), PRESIONADO (botón, b), ENCENDIDO (ventilador, b). Entonces, tenemos que de PRESIONAR (botón, a) se siguen los cuatro cambios indicados en la situación b (para alguien que conoce todas las relaciones de causa-efecto entre esa acción y cada una de esas cuatro propiedades), sin embargo, un agente humano (a saber, el supuesto programador de un sistema inteligente) difícilmente notaría la relación entre PRESIONAR (botón, a) y ENCENDIDO (ventilador, b), a pesar de que indudablemente existe esta relación causal. Por lo tanto, ¿qué corresponde hacer? Evaluemos nuevamente el ejemplo, pero con otra perspectiva. Un agente que ignora la existencia del ventilador en la habitación contigua no notará de ninguna forma que ENCENDIDO (ventilador, b) es una consecuencia de PRESIONAR (botón, a). Y, de tal forma, no hay ningún problema. No obstante, imaginemos otro caso en el que el agente inteligente conoce la existencia del ventilador en la habitación contigua, pero desconoce la conexión entre PRESIONAR (botón, a) y ENCENDIDO (ventilador, b). Si aplicamos la idea de programar previamente axiomas de marco correspondientes con la situación, tal agente desconocedor podría la siguiente línea:

(7) APAGADO (ventilador, a) permanece como APAGADO (ventilador, b) luego de PRESIONAR (botón, a)

Sin embargo, sabemos que este no es el caso, puesto que APAGADO (ventilador, *a*) no permanece como APAGADO (ventilador, *b*) luego de PRESIONAR (botón, *a*), sino que cambia a ENCENDIDO (ventilador, *b*).

Este error en el sistema se debe a que el agente inteligente tiene en cuenta estas líneas:

(8) PRESIONAR (botón, *a*) → DINÁMICO (broca, *b*)

(9) PRESIONAR (botón, *a*) → ENCENDIDO (taladro, *b*)

(10) PRESIONAR (botón, *a*) → PRESIONADO (botón, *b*)

E ignora esta otra:

(11) PRESIONAR (botón, *a*) → ENCENDIDO (ventilador, *b*)

Si extrapolamos este ejemplo a la cognición humana, naturalmente nos encontraríamos con una respuesta de tipo “¿cómo iba a saber yo que PRESIONAR (botón, *a*) tendría como efecto ENCENDIDO (ventilador, *b*)?”. Del mismo modo puede presentarse esto en el agente artificial: ¿cómo iba a inferir que (7) es falso, si desconocía (11), que es información relevante del mundo para este caso?

En este punto nos preguntamos acerca de cómo entender la mente humana, pues lo que expresa (11) puede ser entendido como una presuposición. Para efectos prácticos, conviene apoyarnos en la tesis de que un sistema simbólico es tanto necesario como suficiente para mostrar inteligencia tal como la mente humana lo hace¹⁰.

Dada esta problemática lógica que espero haber expuesto con el mayor nivel posible de detalle, propongo un planteamiento filosófico del problema de marco con la motivación de escoger los conceptos más apropiados y convenientes para repensar el problema:

Problema de marco 5: ¿Cómo reconocer qué propiedades no se ven afectadas al ejecutarse una acción en un mundo complejo?

Puede que parezca una perspectiva simplista del problema, no obstante, observo que el aspecto más importante es saber representar el mundo en términos de propiedades y las relaciones lógicas que hay entre ellas y las acciones. Entonces, la clave para poder resolver el problema está en identificar rápida y eficazmente dichas relaciones entre una acción y las

¹⁰ Ver sección 3.

propiedades del mundo mediante el mecanismo más sencillo, es decir, que demande el menor poder computacional posible.

1.4. Delimitaciones

En las secciones anteriores encontramos cinco planteamientos sobre el problema de marco a pesar de que hay muchas más versiones, pero, curiosamente, pienso que estos son relevantes para el objetivo de esta tesis, que es comprender la dificultad que presenta el problema de marco y cómo abordarlo. Ya que hubo un amplio debate en los años 1980-90s, pretendo aclarar el problema definitorio puesto que cada filósofo aportó con su propia interpretación, ya sea dividiéndolo en aspectos computacionales, epistemológicos o lógicos (Silenzi 2015b), o discutiendo lo que no es el problema de marco (Sprevak 2005: 3-5). Me apoyo en el trabajo de este último autor, y desarrollo ligeramente las supuestas malas interpretaciones del problema. De este modo, veremos lo que no es el problema de marco.

1.4.1. Complejidad computacional

Este no es un problema de complejidad computacional dado que inclusive las máquinas más inteligentes o poderosas se toparán en algún momento con el problema de representar los efectos de las acciones adecuadamente. En otros términos, el sistema de representación no tiene la dificultad de efectuar muchas operaciones, sino en tener la información que le permita representar con precisión el mundo en el que opera. En vista de que el problema de marco presenta la difícil tarea de programar unos supuestos axiomas de marco para indicar qué cosa específica no cambia en el mundo luego de ejecutarse una acción, puede que en un mundo simple (con menos de 10 acciones y 10 propiedades y que, por tanto, contaría con menos de 100 axiomas de marco) quizá indicar todos estos axiomas sea una tarea menos laboriosa. No obstante, cuando el sistema es sometido a un mundo complejo de cambios constantes, este número de axiomas sería tremendamente grande. El problema de marco es fácil de confundir con el de complejidad computacional ya que, en el caso de tener un número muy grande de axiomas, cualquiera podría entender que lo difícil es procesar esa cantidad de información, empero, estos axiomas tendrían que ser programados uno por uno; y el sistema no podría procesarlos si es que no han sido representados por alguien externo al sistema, por no decir que la figura de un programador vendría a ser necesaria. En síntesis, el problema principal yace en cómo es que un sistema puede efectuar las inferencias de un mundo cambiante sin que un programador tenga que intervenir especificando cada cosa que no cambia cuando ocurre una acción; por otra parte, la complejidad computacional involucra

la sobrecarga del sistema tanto por cantidad de información como por el número de operaciones que requeriría ejecutar.

Para mayor esclarecimiento, tengamos dos ejemplos considerablemente distintos:

En uno, el agente artificial tiene la tarea de resolver un clásico cubo Rubik de 3x3x3; y esta tarea consiste en reorganizar los pequeños cuadrados de colores para que cada una de las caras del cubo terminen siendo del mismo color. Como es de saber, se tiene nueve cuadraditos de colores por cada una de las seis caras del cubo; por lo tanto, tenemos un total de cincuenta y cuatro cuadraditos cuyos lugares necesitan ser ubicados para poder obtener el cubo con caras de color completo y no mezclado. De este modo, el programa debe tener en cuenta que hay un número limitado y estricto de movimientos posibles. Estos son los giros horarios o antihorarios de cualquiera de las tres filas o columnas de cada eje; de este modo, hay un total de dieciocho movimientos posibles; y ningún otro movimiento posible. Entonces, en suma, la dificultad en esta tarea radica en la cantidad de movimientos requeridos para ordenar la totalidad del cubo. Además, este proceso no implica solo procesar la cantidad de movimientos de las distintas placas que conforman al cubo, sino, sobre todo, en el cálculo necesario para que las piezas encajen de acuerdo con el objetivo que se ha trazado: reorganizar el cubo.

En otro caso, podemos tener algún objeto que requiera atención mientras seca la pintura con la que se lo ha decorado. Pongamos un cartel de utilería que recién ha sido bañado de color amarillo. Pronto notamos que la pintura, en un inicio, un poco brillante, va disminuyendo a un tono más opaco. Entonces, el cambio de tonalidad en la pintura no va directamente a la par con la forma del objeto, es decir, del cartel. Ahora, podemos volver a ver que el color no ha cambiado y la forma tampoco, pero algo extraño ha ocurrido. ¿Qué cosa? El matiz de la pintura secando. Si vemos con más precaución, develamos una propiedad de la pintura que no es el color, sino la frescura o, mejor dicho, el cambio de estado líquido a sólido. Entonces, ¿cuál es la dificultad de este ejemplo? Pues el grado de detalle con que miramos la situación. Para el problema de marco, lo que importa es determinar lo que no cambia en una situación en términos de objetos y propiedades. Y si tenemos dos o tres propiedades en juego, no nos percatamos de alguna otra a menos de que nos topemos con un caso como este, el de la pintura amarilla, ya que es evidente que algo anda mal. A decir: hay algo o alguna información que se necesita que no está presente en lo que proporciona el ejemplo. ¿Qué es? Una propiedad más específica.

¿Dónde radica la diferencia considerable entre ambos ejemplos? En primer lugar, en que cada situación requiere un número de comandos u operaciones para realizar. En términos concretos, en el caso del cubo Rubik, hay muchos movimientos requeridos para reordenarlo y si evaluamos lo que no cambia, tan solo hay que indicar qué piezas están fuera del alcance de la plancha que se gira en dicho cubo; por otra parte, en el caso del cartel de utilería, solo tenemos la forma del cartel y el color de la pintura en un principio, no obstante, mientras seca la pintura, el matiz se va opacando (y tan solo eso), así que no hay muchas operaciones para realizar sino, más bien, indicar qué no cambia en la medida en que tal pintura va secando (y ya sabemos que ni la forma ni el color cambian mientras pasa esto). Notablemente, la complejidad computacional se manifiesta en el primer caso casi sin esfuerzo, empero, no hay tanta complejidad al ver el segundo caso a pesar de que el problema de ver todo lo que no cambia mientras la pintura seca tiene sus propias exigencias. En suma, en un caso el problema yace en la cantidad de operaciones y comandos para solucionar el problema; en el otro caso, en cómo indicar que el paso de pintura fresca a seca no afecta el resto de las propiedades en la situación mostrada.

1.4.2. Infalibilidad del sistema

No se trata de que el sistema sea infalible, sino que sea razonablemente confiable y robusto. De este modo, tal sistema debe estar en lo correcto más veces que en lo incorrecto, y no debe fallar catastróficamente cuando hay cambios considerados menores en el mundo. Este punto es aceptable en la medida en que nosotros los humanos somos falibles cuando planificamos o predecimos en el mundo cambiante, por lo tanto, un sistema también podría cometer errores que no tengan una magnitud significativa para una tarea en cuestión. Estas líneas quieren decir que las inferencias generadas por el sistema deben, en la mayor parte de los casos, ser apropiadas o acertadas respecto del mundo en el que el sistema se encuentra sumergido. En breve, es permisible que el sistema cometa errores, sin embargo, tales errores no deben exceder un margen prudente que evite la generación de inferencias peligrosas para la representación del mundo o bien de la futura ejecución de acciones.

Simultáneamente, es necesario que el estudio de los potenciales errores en las inferencias o, mejor dicho, en las inferencias falsas, implique unos mecanismos lógicos que eviten estropear la representación general del mundo complejo al que se expone el sistema.

Por ejemplo, tengamos el caso de un agente artificial que tiene la labor de acomodar equipaje de viaje en la maleta de una camioneta. Conforme este irá colocando los objetos unos

encima de otros, podrá notar que ciertas configuraciones de las posiciones no son adecuadas puesto que los objetos se caen y, en consecuencia, la puerta no puede cerrarse. Entonces, el cuidado radica en cómo colocar todo el equipaje sin que este se caiga y así permita cerrar la puerta de la maleta. Un modo factible de probar el sistema que ejecute tales acciones es por ensayo-error. Por decir, si coloca una maleta a un extremo y luego otra más pesada encima, claramente inferirá que el riesgo de caída es mayor que si coloca la maleta más pesada debajo de la más liviana. Sin embargo, puede que haya maletas que, a pesar de ser más pequeñas, resulten más pesadas de lo esperado (o predicho, dicho en mejores términos) por el sistema. Por lo tanto, una manera de comprobar que la supuesta maleta más pequeña, en efecto, es más pesada es colocarlas de tal forma que la ruma de maletas se caiga. Así, desde el punto de vista del programador, el sistema ha cometido un error. No obstante, tengamos el caso en que no solo se cae la ruma de maletas dentro de la maleta, sino que estas empujan el resto de los objetos fuera de la maleta y, a su vez, sale disparada una botella de gaseosa que súbitamente estalla y moja todos los asientos que, a la vez, ocasionan un corto circuito que malogra el arranque del carro y, como consecuencia, ya no puede encenderse y así el viaje debe de abortarse.

Visto este caso, la primera situación en la que el sistema no se percata de la potencial caída de las maletas constituye ciertamente un error. Empero, la segunda situación, en la que las consecuencias resultan fatales no solo para la maleta, sino para los asientos y, en general, para el destino de los viajeros. Notablemente, el margen de error no fue suficientemente agudo en la consideración de quien programó dicho sistema. Sin embargo, este problema es diferente del problema de marco. ¿Por qué motivo? Porque si un sistema fuera infalible, entonces este sería capaz de anticipar cualquier efecto de cualquier causa o, incluso, toda consecuencia de una sola acción, que notoriamente no es el caso, puesto que esto requeriría de un programador conocedor de todas las leyes que rigen el mundo complejo al que es expuesto tal sistema. Sin embargo, las discusiones sobre un posible determinismo del mundo en el que el agente artificial opera escapan al propósito de esta tesis.

1.4.3. Inducción

No es el problema de la inducción propuesto por Hume (1739, Libro 1, parte III, sección 6) en vista que en este se trata de cómo justificar las inferencias de carácter universal a partir de casos particulares, en la medida de que no se dé algún razonamiento circular para confiar en tales inferencias. El problema de marco no trata de justificar sistemas inferenciales, sino en construirlos. Con otras palabras, quiero decir que la tarea de comprobar o fundamentar el

carácter metodológico de un proceso inferencial inductivo no compete a los intereses que evoca el problema de marco; más bien, en cómo elaborar dicho sistema, en el que se llevan a cabo procesos inferenciales que, en efecto, son de casos particulares, pero que no pretenden elaborar una regla de inferencia general respecto de los cambios consecuentes a una acción.

Para graficar estas ideas, veamos un caso de inducción no matemática y otro que involucre el problema de marco. En primer lugar, imaginemos el clásico escenario de los cisnes blancos, según el cual, contar con uno, dos, tres, cuatro y demás cisnes de color blanco, permitiría inferir que la totalidad de cisnes son de ese color. Sin embargo, bastaría con que un cisne sea de otro color, como marrón, para que tal inferencia sea falsa. Analizando este caso, cabe notar que no hemos mencionado acciones, sino objetos (*cisnes*, en este caso) y la propiedad COLOR. Entonces, ¿dónde está la comparación que se hace entre la inducción y el problema de marco? Veamos con un siguiente ejemplo. Si tenemos una mesa en un almacén (un lugar repleto de objetos muy variados) y necesitamos sacarla de la habitación para llevarla a otro lado, la acción que ejecutaríamos sería EMPUJAR. Previamente, como un intento predictivo, tal acción generaría un cambio en la posición del objeto, mas no tan evidentemente el maullido de un gato o el grito de un vecino. Sin embargo, al empujar dicha mesa, no percibimos que un cable atascado en una de las patas, como consecuencia, jaló una escalera que se vino abajo y aplastó la cola de un gato que lanzó un alarido doloroso y estrepitoso. Hasta este punto, el evidente efecto de empujar fue el cambio de posición de la mesa, pero la caída de la escalera o el maullido del gato no estaban incluidos en las inferencias respecto de esa acción. ¿Qué pasaría si tuviéramos como axiomas de marco que el gato permanece en silencio luego de que se empuja la mesa, o que la escalera permanece en pie luego de que se empuja la mesa? Claramente, ambos axiomas serían falsos al comprobar la situación que he presentado.

Empero, ¿y la inducción? Si evaluamos varios casos en los cuales empujar una mesa no provoca el maullido de un gato, sería absurdo “inducir” que empujar mesas no provoca maullidos de gatos, puesto que hay una gran variedad de información (no solo respecto de las mesas o los gatos, sino del contexto de cada caso) que puede eventualmente provocar que un gato maúlle. En el ejemplo presentado: el cable atascado en la pata y la escalera.

El propósito de solucionar el problema de marco es elaborar un sistema que logre representar los efectos considerables de las acciones de modo que quede claro cómo sería el mundo al ejecutar tal o cual acción, es decir, qué cambia y qué no cambia en ese mundo. En cambio, el problema de la inducción no se enfoca en determinar lo que cambia y lo que no cambia,

sino en fundamentar las inferencias que se realizan sobre lo que cambia y no cambia. En este sentido, considero que el problema de la inducción es más amplio que el problema de marco.

1.4.4. Problema de la parada

El problema de interés en esta tesis no es el problema de cuándo dejar de hacer inferencias, el cual es conocido como el problema de la parada [*halting problem*] en la teoría de la computación. Podemos pensar fácilmente que la presencia de axiomas de marco requiere algún criterio que determine en qué momento detener los procesos inferenciales; no obstante, lo que interesa al enfrentar el problema de marco es aclarar qué información es relevante para determinar lo que no cambia al ejecutarse una acción.

Esta dificultad de la parada indicaría que al pretender realizar cuantas computaciones sean necesarias con tal de que se logre representar el mundo luego de que ocurre una acción, no se contaría con un criterio que detenga dichas computaciones, es decir, algo que indique la suficiencia de la información procesada. Este es un problema bastante antiguo en las ciencias de la computación que históricamente viene desde un artículo publicado en 1936 por Alan Turing, donde indica que el problema de la parada, junto con otros, son irresolubles por una máquina universal de Turing o, en otras palabras, por cualquier computadora que funcione con una lógica de semántica binaria, donde una proposición solo puede tener dos valores de verdad: verdadero (1) o falso (0).

1.4.5. Abducción

El problema de marco no es el problema de la abducción en el sentido de adquirir creencias explicativas, es decir, cómo se generan hipótesis a partir de un conjunto finito de observaciones. Este caso es interesante porque puede confundirse con facilidad la generación de hipótesis plausibles con la determinación de un criterio de relevancia, ya que en ambos casos se busca alguna relación causal entre dos partes. En el caso de la abducción, lo que se genera es una explicación verosímil de algún fenómeno en cuestión; y, por motivos pedagógicos, puedo mencionar que un razonamiento abductivo es del tipo que llevan a cabo los detectives cuando evalúan signos que les permitan reconstruir la escena de un crimen.

Por ejemplo, cuando un perito encuentra huellas de barro arcilloso cerca de la víctima, puede abducir que el asesino pasó por una cancha de tenis húmeda momentos previos al crimen. Y, si además nota que la talla del calzado es pequeña y delgada, con mayor probabilidad puede abducir que quien cometió el crimen es una mujer. Profundizando más en la escena,

conforme el perito encuentre otra señal, conseguirá elaborar una explicación más detallada de la posible escena del crimen, y así sucesivamente. De esta manera, las inferencias realizadas no se siguen deductivamente, como sí lo haría un programa. Por lo tanto, el sistema inteligente que presente el problema de marco no se encargaría de construir explicaciones de qué acciones ocasionaron tal o cual cosa, sino, más bien, qué propiedades no se ven alteradas por tal o cual acción.

A este respecto, la diferencia principal se encuentra en que en la abducción crea una explicación o justificación de la información recogida¹¹ y en el problema de marco se pretende determinar lo que permanece igual luego de que ocurre una acción, donde no necesariamente tiene que dar una explicación.

1.4.6. Revisión de creencias

No se trata del problema de la revisión de creencias, es decir, cómo un sistema debería reaccionar frente a nueva información. No necesariamente un sistema que padezca el problema de marco también padece del problema de revisión de creencias, dado que la cuestión central es cómo se representa el mundo frente a las acciones en términos de lo que cambia y lo que no cambia. De este modo, es posible que el programa no necesite adquirir nueva información si se programan todos los axiomas de marco necesarios (tomando como ejemplo un mundo simple con pocas acciones y eventos, por motivos explicativos).

Por ejemplo, tengamos una mañana en la que un agente cree que ver una polilla indica que el día será soleado. Y, en efecto, ese día es soleado. Pero las siguientes dos semanas continúa viendo polillas eventualmente en las mañanas y los días ya no son soleados como sería de esperarse. Por tanto, tal agente se vería en la situación de cambiar esa creencia ya que no le sirve; en este caso, el agente prefiere creer que ver una polilla no necesariamente indica que el día será soleado. ¿Y el problema de marco? En ningún momento el ejemplo muestra énfasis en determinar qué propiedades cambian y cuáles no cambian en el mundo, sino cómo un agente cambia una creencia por otra en la medida de, en este caso, su utilidad explicativa. Sin embargo, no son requeridos ni axiomas de marco ni algún otro mecanismo que especifique que tal cambio de información (cambio de creencia) permite determinar las propiedades que no se ven afectadas por cierta acción. Por lo tanto, el cambio de creencias no es el problema de marco.

¹¹ Cf. Aliseda-Llera (1997)

1.4.7. Problema de la calificación

No se trata del problema de la cláusula *ceteris paribus*. Este es el llamado problema de la calificación [*qualification problem*] en la literatura de IA, pero la presentación es diferente al problema de marco. Este problema trata de especificar inferencias inductivas de modo que sean anulables, es decir, que las inferencias basadas en la observación se mantengan siempre que no haya circunstancias que las bloqueen o que eviten que sea el caso. En otras palabras, el problema de la calificación consiste en determinar qué circunstancias¹² o precondiciones¹³ son dadas para que una causa pueda tener su efecto. Como puede haber una infinidad de esas circunstancias, no se pueden especificar todas. De esta manera, asumimos que no hay ninguna de estas circunstancias de bloqueo a menos que haya razones contrarias. En otras palabras, si X es verdadero y todo lo demás permanece igual (o nada impide que suceda Y), entonces se concluye Y . Sin embargo, el problema de marco nos dice que concluimos Y , y que todo lo demás permanece igual dado X , a menos que haya una razón para lo contrario (para que Y haga que no todo lo demás permanezca igual). Formalmente, lo podemos ver de la siguiente manera:

(a) Si $X \wedge \dots$, entonces Y (problema de *ceteris paribus*)

(b) Si X , entonces $Y \wedge \dots$ (problema de marco)

Así, profundizando en las explicaciones de Sprevak, notamos qué aspectos son mayormente confundidos con el problema de marco, o que son entendidos como este (2005). Al respecto, este autor comenta que en el problema de marco se asume que una propiedad (o función) sigue igual si no hay alguna regla explícita (o acción) que indique su cambio. Esta es denominada la ley de inercia del sentido común [*commonsense law of inertia*] acuñada y desarrollada ampliamente por Shanahan (1997). Según ella, se asume que todo el mundo permanece igual luego de ejecutarse una acción a menos que alguna restricción indique lo contrario. En este caso, la regla indica qué cambia, es decir, cuáles son los efectos de dicha acción. Por lo tanto, si no se ha especificado que una propiedad se vea afectada por una acción, se asume que permanece igual. Sin embargo, Shanahan resalta que esta ley de inercia debe considerarse como un dispositivo representacional y no como un enunciado acerca del mundo; puesto que en el mundo encontramos que tal o cual objeto instancia alguna propiedad que cambia con una acción, pero no encontramos concretamente la indicación de asumir la

¹² Cf. Thielscher (2001)

¹³ Cf. Shanahan (1997: 22)

permanencia de las propiedades específicamente no afectadas por el cambio. Por ejemplo, al serruchar un tablón de madera veremos que la longitud de este ya no es la misma, pues ahora se tienen dos tablas de menor longitud. La acción de serruchar cambió la longitud del tablón, pero no el tipo de madera ni el tipo de metal del serrucho. Empero, la ley de inercia del sentido común cumple un rol útil al razonar sobre esta situación: si en el mundo no hay indicación de que serruchar un tablón cambie la textura de mi polo, entonces asumiré que esta sigue siendo la misma después de serruchar tal tablón.

Estas delimitaciones evitan que se malinterprete el examen técnico con problemas similares, pero que no son precisamente el objeto del problema de marco. Las distinciones mencionadas ayudan a observar la complejidad en el estudio del razonamiento del sentido común, puesto que involucra variadas formas que procesan la información con diferentes propósitos. En este caso, me interesa saber qué propiedades no se ven afectadas al ejecutarse una acción en el mundo, como ya ha sido mencionado en múltiples ocasiones a lo largo de esta tesis.

1.5. Criterios de solución

En un sentido amplio, y teniendo en mente las consideraciones explicadas, plantear y resolver el problema implica entender cómo es que funcionan los mecanismos de planificación, predicción y acción que poseemos nosotros los seres humanos. Más aún, siguiendo a Shoham en Brown, a pesar de que fue definido en el contexto del *cálculo de situaciones*, un formalismo específico, quedó claro desde el comienzo que este es una manifestación de algún problema fundamental en el razonamiento temporal¹⁴. Y él sostiene que el problema de marco surge de los deseos conflictivos al razonar rigurosa y eficientemente acerca del futuro (1987)¹⁵. Esta perspectiva nos permite notar que la inteligencia de un agente (o de un sistema inteligente) radica en la eficacia con que logra predecir los acontecimientos y, a la vez, conjugarlos con los cursos de acción que lo acercan a sus objetivos (este es el punto en el que se confunde con el problema de la inducción y de la abducción, como vimos en las secciones 1.4.3. y 1.4.5., sin embargo, la diferencia está en la formalización de tales procesos). Para ello, es necesario tener claro cómo tales acontecimientos interactúan con las acciones ejecutadas y los efectos de ellas. Estas

¹⁴ Ver sección 2.

¹⁵ "Although it was defined in the narrow context of the *situation calculus*, a specific temporal formalism, it was clear from the start that it is in fact a manifestation of some fundamental problem in temporal reasoning (...) We argue that the frame problem arises from the conflicting desires to reason both rigorously and efficiently about the future" (*Tr. p.*)

relaciones entre cada rebanada de tiempo son llamadas *leyes de movimiento* y estas describen cómo el agente piensa que el mundo cambia a lo largo del tiempo (Sprevak 2005: 2)¹⁶.

Siguiendo esta línea, estamos de acuerdo con Morgenstern en que resolver el problema de marco es central al razonamiento temporal, y que el razonamiento temporal es central a la inteligencia artificial, puesto que este es necesario para planificar, explicar y diagnosticar (1996: 4). Por esta razón desarrollo siete puntos importantes mencionados por esta autora como criterios que debe cumplir cualquier solución exitosa del problema de marco.

1.5.1. Análisis temporal

Debe ir más allá de cualquier lenguaje temporal restrictivo como el *cálculo de situaciones*¹⁷ y debe ser compatible con una teoría general del razonamiento temporal. De este modo, la solución al problema de marco tiene un grado de generalidad aplicable a la cognición respecto del tiempo. Esto quiere decir que la notación que se emplee para describir el cambio en el mundo debe ser lo suficientemente explicativa para abarcar las relaciones causales entre las acciones y las propiedades. Por este motivo, escojo este criterio como fundamental para cualquier planteamiento sensato y para cualquier posible solución del problema.

Un punto por resaltar es la diferencia entre el universo lógico y el mundo real para los humanos, puesto que el tiempo siempre transcurre para nosotros o, por lo menos, así lo es para una cognición sana y cuerda. El razonamiento acerca del tiempo involucra, a su vez, alguna manera de concordar las impresiones que el cambio produce en nuestra percepción de tales o cuales propiedades. Por ejemplo, imaginemos un osito de peluche que reposa tranquilamente sobre un sillón; el peluche no se mueve a ningún lado, no cambia de color ni de material de confección; sin embargo, si permanece en ese estado, eventualmente caerá polvo sobre él y alguien lo notará; la impresión del peluche habrá cambiado de “LIMPIO (*peluche, a*)” a “SUCIO (*peluche, m*)”. Claramente hubo un cambio y este solo puede explicarse, evaluarse o verse, teniendo en cuenta el tiempo como factor de análisis. En el ejemplo propuesto: si no se examina temporalmente el caso, ¿cuál sería el mejor camino, si solo queda el universo lógico? Por esta razón, el tiempo es lo que enfatizo arduamente en el siguiente capítulo de esta tesis.

1.5.2. Premisa intuitiva del sentido común

¹⁶ “These functions or relations between time-slices are called the ‘laws of motion’: they describe how the agent thinks the world changes over time” (*Tr. p.*)

¹⁷ *Cf.* Shanahan (1993)

Debe seguir las intuiciones de que el desarrollo de una acción causa que ciertas cosas del mundo cambien y de que la mayoría de los otros aspectos del mundo permanezcan iguales. A pesar de que esta no es una regla absoluta porque puede ser falsa en algunos casos, típicamente es verdadera. Teniendo en mente este criterio, como explican Moro y Silenzi, algunos sostienen que las personas utilizan heurísticas para elaborar juicios y tomar decisiones que involucran situaciones de incertidumbre. Por un lado, la ventaja es que las heurísticas reducen el tiempo y el esfuerzo para realizar juicios y tomar decisiones. Sin embargo, por otro lado, la desventaja radica en que tales heurísticas conducen a errores sistemáticos o sesgos bajo determinadas circunstancias (2017: 104).

Si bien, en un principio, el lenguaje empleado para el problema de marco era puramente formal, con los años de amplio debate se devela el hecho de que la suerte de “lógica mental” por modelizar no sigue necesariamente los principios de la lógica matemática o, más específicamente, de la lógica computacional. Por esta razón, este criterio enfatiza que el tipo de lógica o los enunciados con los que se diseña el sistema deben estar basados en un correlato asequible con la cognición humana del llamado sentido común. Notablemente es una ardua tarea encontrar dicha(s) premisa(s), si es que, en buena hora, la(s) haya; sin embargo, las esperanzas de avanzar recaen en la mayor correspondencia entre el modelo lógico y la mente. He ahí el punto importante de este criterio entre la inteligencia artificial y la cognición humana.

1.5.3. Principios rigurosos

Los axiomas en la teoría deben ser verdaderos, ya que, a pesar de que las aproximaciones pueden ser convenientes, una teoría falsa eventualmente llevará a resultados incorrectos. En otras palabras, se debe optar por encontrar los principios adecuados que rijan el modo según el cual razonamos respecto del cambio. Dichos principios, como es de imaginar, necesitan ser lo suficientemente acotados para poder mantener una visión clara sobre el problema en cuestión. Para que pueda lograrse este cometido, cabe resaltar el aspecto computacional inicial del problema de marco, ya que así se conseguirá el rigor que amerita cada principio. Estos necesitan ser encontrados por una ardua labor investigativa sobre la cognición humana, como mencioné en secciones anteriores.

En un caso, si tenemos como principio que: si una acción “HACER (x, a)” es efectuada, al menos una propiedad “PROPIEDAD (x, a)” debe de cambiar para que esta sea considerada una acción. Siguiendo esto, notamos que conforme se realiza tal acción, al menos dicha

propiedad cambia. Entonces encontramos un razonamiento circular que, por defecto, es falaz y, por lo tanto, inválido.

En cambio, en otro caso, si establecemos como principio que: de efectuarse la acción “HACER (x, a) ” en un momento a , al menos una propiedad “PROPIEDAD (x, a) ” debe cambiar en un momento consecutivo inmediato b : PROPIEDAD (x, b) . Entonces tendremos una idea teóricamente aceptable de cómo se realizan las inferencias en el sistema programado.

1.5.4. Generalidad

La teoría debe demostrar soluciones actuales a algún conjunto de problemas concretos similares, dado que nos fuerza a ser intelectualmente honestos y a trabajar las cosas en detalle. Como vimos en distintas secciones anteriores, hay varios problemas relacionados con el problema de marco cuyas diferencias yacen en los detalles formales de sus respectivos planteamientos. Si bien el acercamiento lógico de cada uno de estos problemas es una tarea independiente en términos de programación, estos corresponden a capacidades cognitivas que empleamos en el razonamiento del sentido común. Por lo tanto, evaluarlos por separado permite ahondar con meticulosidad, pero también es adecuado metodológicamente ofrecer una visión integral de todos ellos (o al menos de un número considerable).

1.5.5. Precisión

Debe tener una representación concisa a fin de evitar tener una larga lista de axiomas de marco. Esto quiere decir que es preferible evitar que el sistema requiera demasiadas consideraciones para efectuar un razonamiento frente al cambio. Más bien, sería mejor que tuviera un método preciso que se aplique mecánicamente a cada situación para evitar pasos en falso o gastos innecesarios de energía computacional; es decir, prima la menor cantidad posible de operaciones o líneas para inferir adecuadamente.

Por ejemplo, al destapar una botella para beber algo de agua. La acción DESTAPAR (x, a) es una; y BEBER (y, b) es otra. Tenemos los objetos $x = botella$; $y = agua$. Y las situaciones sucesivas a, b, c, \dots . Doy a entender que se efectúa primero DESTAPAR $(botella, a)$ y, en un momento consecutivo, se efectúa BEBER $(agua, b)$. No obstante, hay un detalle interesante que sale a la luz.

Al contar con dos objetos, indicar sobre cuál se efectúa la acción o sobre cuál recae la acción ejecutada es primordial, puesto que un error de tipo “DESTAPAR $(agua, a)$ ” puede

confundirse con “BEBER (*botella, a*)”. En un terreno lógico, estas expresiones son notablemente distintas. Sin embargo, en el mundo real puede entenderse lo mismo: *alguien que sacia su sed con agua embotellada*. Y, para tal cometido, necesita retirar la tapa de la botella debido a que esta obstruye su paso hacia la boca del agente inteligente. Por esta razón, es casi un epíteto señalar que el factor precisión tanto en la notación como en los procedimientos son indispensables para plantear y, eventualmente, solucionar el problema de marco.

1.5.6. Atención epistemológica

Debe estar basada en la teoría, puesto que muchos intentos de solución son puramente formales en cuanto a procedimientos, pero carecen de un respaldo teórico sobre cómo funciona nuestro razonamiento temporal. Así, volviendo a la diferencia de aspectos computacional y epistemológico del problema, no solo se debe afrontar el primero, sino también el segundo. La primacía de este criterio evita que se construyan sistemas sumamente eficientes pero que carezcan de correspondencia con nuestros procesos atribuidos al sentido común. De esta manera, se evita las llamadas *ruedas cognitivas*¹⁸ por Dennett (1987). Según él, puede que una rueda sea un excelente mecanismo de movilización, sin embargo, no hay manifestación biológica de las ruedas en ninguna especie. Por esta razón, critica los intentos lógicos que pierden correlato con la realidad humana.

1.5.7. Simpleza

Debe ser simple (aplicando la navaja de Ockham) ya que siendo menos trabajoso y más elegante es mejor en las labores formales (y posiblemente también en otras esferas de la vida). Comparable con la generalidad, este criterio también es preferible no solo desde una perspectiva estética, sino desde la economía de la construcción teórica.

Luego de mencionar y explicar estas siete nociones, comparto una pequeña síntesis de este último subcapítulo. Estos criterios surgen como requerimientos fundamentales debido a que detectamos constantemente la contraposición entre el desarrollo lógico-computacional del problema y la investigación científico-filosófica del problema respecto de la mente humana. Ese es un dilema permanente en la inteligencia artificial pues algunos trabajan desarrollando programas independientemente de su correspondencia con los procesos cognitivos y otros,

¹⁸ “*cognitive wheels*” (Tr. p.)

más bien, están interesados en estudiar dichos procesos para reproducirlos computacionalmente.

En todo caso, esta tesis se restringe al alcance de este problema en su tratamiento lógico y filosófico. Por tal motivo, evalúo y contrasto las distintas ideas y posturas que se han presentado con tal de entender y resolver, en la medida de mis recursos, esta cuestión.

Capítulo 2. Lógica temporal métrica

Lo que me propongo hacer en este segundo capítulo es desarrollar la propuesta de Silenzi (2011b) según la cual la lógica temporal métrica de Prior (1967a) es útil para reformular el problema de marco en cuanto a la dimensión temporal en el análisis de las situaciones. Para ello, hay conceptos lógicos básicos que explico brevemente y después empleo una notación sencilla para que sea suficientemente comprensible y expresiva. Enfatizo que la notación lógica de Shanahan (2016), que fue empleada para explicar el origen del problema (sección 1.1.) y su siguiente análisis (sección 1.3.), no la continúo en las próximas secciones. Sin embargo, los nuevos términos lógicos se adecúan a las consideraciones explicadas en el primer capítulo. Tampoco utilizo estrictamente la notación proporcionada por los distintos estudiosos de la lógica en cuyo trabajo me apoyo debido a que encuentro conveniente ajustar los términos a las exigencias del problema de marco. En otras palabras, no vuelvo a la notación del primer capítulo exactamente, puesto que ahora agrego nuevos conceptos lógicos, más específicamente, sobre el tiempo.

Lo que busco hacer en este capítulo es mostrar cómo es posible simbolizar explícitamente el cambio de propiedades a través del tiempo, es decir, la notación que sugeriré servirá de apoyo formal a las consideraciones filosóficas acerca del tiempo y el cambio de propiedades mientras este transcurre. De tal modo, en el mejor de los casos, se conseguirá mayor claridad sobre la atención que merece el problema de marco tanto filosófica como lógicamente.

2.1. Desarrollo del sistema

Teniendo en mente como interés central la adecuación formal de los planteamientos e inconvenientes de la teoría para afrontar el problema de marco, decidí elaborar una notación que facilite la comprensión y asimilación de las dificultades que se nos presentan al modelizar este problema. La empresa de encontrar una solución o, mejor dicho, de reestructurar, replantear y reformular lógica y, por ende, filosóficamente, el problema de marco implica escoger un camino que prometa un mayor y fructífero avance. En esta tesis

opté por elegir la dimensión temporal debido a que considero indispensable tener en cuenta al tiempo para estudiar cualquier problema o proceso inferencial que involucre cambio, puesto que, evidentemente, todo cambio supone el paso de un momento inicial a un momento final; más aún, si tal estudio del cambio está abocado al terreno de la cognición humana con miras a la producción de un agente artificial inteligente. Esta es, pues, la vía que podría mostrar mayor riqueza en cuanto a la investigación del razonamiento del sentido común.

Junto a estas intenciones, coloco el énfasis en el estudio lógico del tiempo para comprender bien qué sucede en el espacio lógico cuando analizamos el paso de un momento a otro, el desvanecimiento de una propiedad y la aparición de otra, o cómo una simple acción produce que más de una propiedad automáticamente cambie en un mundo complejo de múltiples propiedades relacionadas unas con otras. Por ejemplo, al encender una mecha de inmediato se consume un milímetro de la soguilla que lleva hacia la dinamita y claramente la soguilla tiene mucho más que tan solo un milímetro de longitud. Por lo tanto, a cada momento se va consumiendo un milímetro más hasta que el fuego llega a la carga y los explosivos son detonados rebosantemente. En un lenguaje formal bastante simple podemos colocar que la acción de encender la mecha trae como consecuencia la detonación de la carga de explosivos, sin embargo, esta no es una consecuencia inmediata. Explicar e incluir en el modelo lógico tal *inmediatez* o, mejor dicho, la *no inmediatez* es lo que motiva a retomar los estudios de la lógica temporal.

2.1.1 Enfoque filosófico

Respecto al estudio de la *inmediatez* y la *no inmediatez*, nos topamos con el debate acerca del tiempo como un flujo continuo o una concatenación de unidades discretas. Por motivos de practicidad teórica, y teniendo en cuenta el objetivo de esta tesis, cabe resaltar que tomaremos en consideración al tiempo como una magnitud que puede ser medida en cantidades discretas, es decir, si bien puede ser evaluada con mucho más detalle si evocamos el concepto de infinitesimal, tomamos la batuta respecto a la perspectiva del tiempo como una magnitud medible extensamente¹⁹. Por lo tanto, un presupuesto del que partimos es que una lógica sobre el tiempo discreto puede y debería ser suficiente para modelizar el cambio, pues esto es lo que nos interesa al tratar el problema de marco.

2.1.2. Presentación

¹⁹ Cf. Bell (2017)

Si bien me inspiro en la lógica que encontramos en Prior (1967a), tal autor usa la notación polaca propuesta por Łukasiewicz en 1924²⁰, que es poco habitual de encontrar en los estudios contemporáneos de nuestro problema en cuestión. Resulta fácil de observar que la disposición de los símbolos en este tipo de notación carece de paréntesis y diferenciar la jerarquía de los operadores puede ser un poco complicado para las personas no entrenadas en su lectura. Por esta razón empleo una adaptación de la simbolización empleada en Øhrstrøm y Hasle (1995) y en Cresswell (2013), cuyos aportes están basados en el trabajo que movió a los esfuerzos de Prior, no solo en gran parte de su vida intelectual, sino también personal.

Como es de saber común entre los conocedores técnicos, todo sistema lógico debe, antes que cualquier otra cosa, ser delimitado con claras definiciones y, de la misma manera, por reglas de inferencia suficientemente transparentes para que permitan un provechoso futuro empleo para el cual se lo construye: generar un modelo. Y una cita interesante para empezar gentilmente es la siguiente: “[El] llamado sistema minimalista para la lógica temporal métrica”²¹ (Øhrstrøm / Hasle 1995: 231)²², está construido en base a un conjunto de variables proposicionales: $\{p; q; r; \dots\}$, en otro conjunto de variables métricas: $\{n; m; \dots\}$, y en las siguientes definiciones de fórmula bien formada (fbf):

[Def 1] Las variables proposicionales son fbf.

[Def 2] Si α y β son fbf, y n es un número positivo, entonces $\neg\alpha$, $\alpha \supset \beta$, $\alpha \wedge \beta$, $\alpha \vee \beta$, $\forall n: \alpha$, $\exists n: \alpha$, $P(n)\alpha$, $F(n)\alpha$ son todas fbf.

[Def 3] No hay otras fbf.

Siendo así la delimitación del sistema, cabe señalar que las definiciones propugnadas son como las que mayormente encontramos en la lógica clásica de primer orden. Exactamente, por motivos de extensión argumentativa, no ahondaré en lógicas de mayor orden.

Otro punto interesante por remarcar es que hay nuevas simbolizaciones; notoriamente, las que corresponden a las partes temporales de los enunciados. Ahora bien, la novedad yace en que los elementos característicos de la lógica temporal métrica son $P(n)\alpha$ y $F(n)\alpha$, donde P

²⁰ Cf. Simons (2017)

²¹ “(...) the so-called minimal system for metric tense logic” (*Tr. p.*)

²² Denominado sistema *MT* por estos autores

es el operador que refiere al pasado y F , el que refiere al futuro. Además, la variable n indica la distancia temporal entre el *ahora* y el momento descrito por la expresión en cuestión.

Estas expresiones se leen así:

$P(n)\alpha$ “ha sido el caso (hace n unidades de tiempo) que α ”

$F(n)\alpha$ “será el caso (en n unidades de tiempo) que α ”

Y a partir de estos operadores logramos definir los otros operadores temporales H y G , tales que:

[Def H] $H(n)\alpha \equiv_{df} \neg P(n)\neg\alpha$

[Def G] $G(n)\alpha \equiv_{df} \neg F(n)\neg\alpha$

Que, a su vez, se leen:

$H(n)\alpha$ “siempre ha sido el caso hace n unidades de tiempo que α ”

$G(n)\alpha$ “siempre será el caso en n unidades de tiempo que α ”

Posteriormente, Øhrstrøm y Hasle aseveran que prefieren no asumir nada respecto a la estructura del tiempo dado que podemos postular dos versiones o perspectivas. Una primera, en relación con un modelo ramificado del tiempo, en la que $G(n)\alpha$ es verdadera si y solo si α es verdadera en n unidades de tiempo desde el *ahora* hasta en cualquier rama del futuro, es decir, en todo futuro posible. No obstante, una segunda, en un modelo basado en el tiempo lineal, en donde no habría diferencia entre $G(n)\alpha$ y $F(n)\alpha$, dado que $\neg F(n)\alpha \equiv F(n)\neg\alpha$ en tal modelo (1995: 232). Lo más interesante está en cómo se modela el tiempo en el espacio lógico sugerido. Siguiendo así estas ideas, cabe señalar por conveniencia teórica que resulta preferible evitar los debates o discusiones acerca de la metafísica del tiempo, puesto que, si bien tales exámenes pueden traer beneficio para una visión menos incompleta del problema de marco, estos escapan al foco de esta tesis. También viene a la mente la duda acerca de si estas expresiones son las más adecuadas para el problema de marco. Empero, estoy dispuesto a esclarecer un poco más estas nociones en el tercer capítulo de esta tesis.

Para mayor entendimiento, ofrezco algunos gráficos que ejemplifican ambas perspectivas mencionadas acerca de cómo optaron estos autores para ver el tiempo con motivos de la lógica propuesta y desarrollada paso a paso:

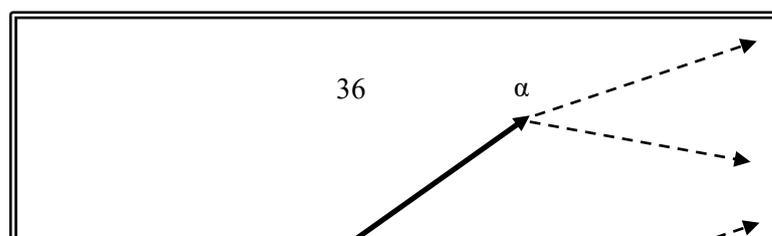
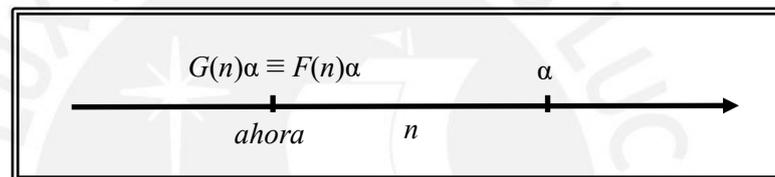


Figura 2.1. Modelo ramificado del tiempo (elaboración propia)**Figura 2.2.** Modelo lineal del tiempo (elaboración propia)

Por motivos prácticos, adopto el modelo ramificado del tiempo debido a que especificar la diferencia entre $G(n)\alpha$ y $F(n)\alpha$ es útil para ser aplicada al problema de marco. Por otra parte, recalco enfáticamente que no es de interés en esta tesis tratar aspectos metafísicos del tiempo tales como dilucidar lo que es, ha sido y será el caso o entrar en el debate del presentismo y el eternalismo. Simplemente abordo la problemática con estas herramientas lógicas en un nivel técnico con el objetivo de dar un fecundo pronóstico de solución. Sin embargo, reconozco que, en medio de algunas enmarañadas discusiones, este desarrollo puede llevar eventualmente a una valiosa interpretación filosófica acerca del sentido común, quizá, en cuanto a nuestra percepción del tiempo propiamente dicho.

Retomando el motivo principal de este segundo capítulo, nos corresponde entender cómo funciona nuestro sistema recién mostrado.

Para notar su expresividad, veamos uno de sus axiomas, denominado MT2' (Øhrstrøm /Hasle 1995: 233), como ejemplo:

(MT2') Si ha sido el caso, hace n unidades de tiempo, que siempre sería el caso, en n unidades de tiempo, que p , entonces p .

A grandes rasgos, este axioma quiere decir lo siguiente: si en el pasado fue verdadero que siempre sería verdadero p , es correcto deducir que en el presente p es verdadero; y la variable n describe la distancia entre el *ahora*, aquel momento en el pasado y el momento futuro de ese momento pasado. Encontramos un ejemplo eficaz en el recuerdo de las promesas. Si Pepe prometió hace exactamente un año que siempre enviaría flores en un día como hoy, entonces es el caso que el día de hoy envió dichas flores.

En breve, la simbolización que corresponde con MT2' es la siguiente:

$$P(n)G(n)p \supset p$$

Y, consecuentemente, esta notación aplicada al ejemplo propuesto resultaría así:

$$P(\text{un año})G(\text{un año}) \text{Pepe envía flores} \supset \text{Pepe envía flores}$$

Donde la medida temporal n corresponde con *un año*, y la variable proposicional p simboliza la frase 'Pepe envía flores'. Este ejemplo muestra cómo cada operador temporal refiere al pasado o al futuro como lo harían los operadores modales de necesidad y posibilidad. ¿Cómo así? Pues los operadores de necesidad serían $H(n)$ para el pasado y $G(n)$ para el futuro; y los operadores de posibilidad serían $P(n)$ para el pasado y $F(n)$ para el futuro. En un modelo ramificado del tiempo, los operadores de necesidad indican que tal o cual propiedad de todas formas será instanciada, ya sea en el pasado o en el futuro; en cambio, los de posibilidad señalan la contingencia de una propiedad, ya sea en el pasado o en el futuro.

2.1.3. Con propiedades

En comparación con el *cálculo de situaciones*, en el sistema presentado se usan variables proposicionales y no propiedades. No obstante, la lógica temporal métrica presentada, explicada y trabajada por Cresswell (2013) toma en cuenta a las propiedades en la simbolización. En este sentido, hago una pequeña variación a dicha notación expuesta en la sección anterior.

Cabe mantener la vigencia de nuestro interés en aplicar la lógica temporal al problema de marco. Sin más palabras, el ejemplo que este autor toma para analizar la lógica heredada de Prior es el siguiente:

(1) Un día todos los que ahora son miserables serán felices.

Y la simbolización que ofrece es esta:

$$\exists n F n \forall x (P n \text{ miserable } x \supset \text{ feliz } x)^{23}$$

Expresado en términos no simbólicos, significa que hay algún intervalo con n unidades de tiempo hacia el futuro en que todo x , si fue miserable en un intervalo con n unidades de tiempo hacia el pasado de ese futuro, es feliz.

Como habrá que observar, tanto las propiedades ‘miserable’ como ‘feliz’ carecen de simbolización y, en comparación con el sistema MT de Øhrstrøm y Hasle, se agrega las variables de objetos: x, y, z, \dots . Entonces, no hay ningún problema con escribir literalmente las propiedades tal como Cresswell lo hace.

Otro ejemplo puede ser el siguiente:

(2) La manzana roja era verde hace un día y en dos días será marrón.

Primero cabe determinar el objeto y las propiedades con las que trabajaremos. Tenemos como objeto a la ‘manzana’ y tenemos tres propiedades: *rojo*, *verde* y *marrón*. Y temporalmente tenemos las distancias de un día entre la manzana verde y la roja; y de dos días entre la manzana roja y la marrón. Pues bien, ya podemos establecer una simbolización. Sería de la siguiente manera:

$$P n \text{ manzana verde} \wedge \text{ manzana roja} \wedge F 2 n \text{ manzana marrón}$$

Lo cual indica que hace una unidad de tiempo la manzana fue verde, en el presente es roja y en dos unidades de tiempo será marrón. Colocamos la conjunción entre esas expresiones, pues no hay ninguna relación condicional explícita en nuestro ejemplo.

Un punto importante que resaltar es que el momento presente de referencia puede ser cambiado de lugar dependiendo del lenguaje en que se haga más simple el análisis de los enunciados. Por decir lo menos, se pudo tomar como presente de referencia el momento en que la manzana era verde y en n unidades hacia el futuro se convertirá en roja, y en $3n$ unidades futuras será marrón. Se entiende de la siguiente forma:

$$\text{ manzana verde} \wedge F n \text{ manzana roja} \wedge F 3 n \text{ manzana marrón}$$

²³ Un detalle que diferencia a Cresswell (2013) de Øhrstrøm y Hasle (1995) es la omisión de los paréntesis “()” para indicar la variable métrica n . De ahora en adelante las variables métricas se pondrán sin paréntesis de la forma $Fn, Pn, PnGn$, etc.

Con este último cambio en la formalización del ejemplo quise mostrar la dinamicidad y facilidad con la que pueden simbolizarse las expresiones con el lenguaje formal temporal propuesto. Dado esto, ya podemos agregar a las acciones como siguiente paso.

2.1.4. Con acciones

Nuevamente, en contraste con el *cálculo de situaciones*, ni Øhrstrøm y Hasle (1995) ni Cresswell (2013) simbolizan las acciones. Empero, el sistema no encuentra dificultades si es que escribimos literalmente las acciones de tipo ‘*mover x*’. Por ejemplo, la siguiente expresión:

(3) En algún momento se moverán todos los objetos rojos.

Puede simbolizarse como:

$$\exists n F n \forall x (P_n \text{ rojo } x \supset \text{ mover } x)$$

Y esto se lee: *existe un n tal que, dentro de n unidades temporales, si fue el caso hace n unidades temporales que x fue rojo, entonces se mueve x.*

Lo que quiero especificar en esta sección es cómo insertar las acciones en la notación de la lógica temporal métrica.

Veamos otro ejemplo con la acción *pintar*, el objeto *x* y las propiedades *azul* y *negro*. Establezcamos que de pintar el objeto *x* cambia de color azul a negro:

$$\exists n F n \exists x ((P_n \text{ azul } x \supset \text{ pintar } x) \supset \text{ negro } x)$$

Que se lee: *existe un n tal que, dentro de n unidades temporales existe un objeto x tal que, si x fue azul hace n unidades temporales, entonces es pintado, entonces es negro.* Lo cual indica cómo procede el sistema para registrar el cambio de un color a otro como resultado de la ejecución de una acción. Un detalle que llama la atención se encuentra en el paréntesis, pues la relación condicional entre el primer color y la acción es inversa a la que habíamos utilizado en la presentación del origen del problema de marco. Anteriormente lo habríamos colocado como “*pintar x* \supset *negro x*”, sin embargo, ahora debemos tener cuidado con los operadores temporales, pues estos expresan los saltos que queremos graficar entre un momento y otro en nuestro modelo ramificado del tiempo. Si interpretamos el hecho de que un operador temporal del pasado esté dentro del rango de alcance de un operador temporal del futuro, y si además ambos tienen las mismas unidades temporales con las que se trabaja, entonces se cancelarían y tendríamos que la propiedad expresada por el antecedente refiere

al momento presente. En otras palabras, si será el caso en n unidades que fue el caso hace n unidades, entonces tenemos que es el caso que tal y cual cosa: la expresión ' $F_n (P_n)$ ' equivaldría al momento presente, pues ambos operadores temporales se cancelan.

Ahora bien, el consecuente dentro del paréntesis no tiene operador temporal, por lo tanto, está bajo el alcance del operador temporal del futuro que rige ese paréntesis. Entonces, tenemos lo siguiente:

Si será el caso en n unidades que el objeto x fue *azul* hace n unidades, entonces se ejecuta *pintar* x .

De esta forma queda claro cómo el paso de un color a otro depende exclusivamente de una acción. En este caso, de *pintar*.

2.2. Aplicación al problema de marco

Un concepto bastante discutido y aplicado en las soluciones lógicas al problema de marco es la ley de inercia del sentido común (*commonsense law of inertia*)²⁴. Para definirla, tomamos la explicación de Mueller (2015) que afirma, *mutatis mutandis*, que los objetos conservan cualquier propiedad a menos que sean afectados por acciones. Por lo tanto, veamos la siguiente oración:

(4) Todo objeto que es rojo seguirá siendo rojo a menos que lo pintemos.

Haciendo uso del lenguaje formal ofrecido en la sección anterior, obtenemos esta simbolización:

$$\forall n \forall x ((\text{rojo } x \supset G_n \text{ rojo } x) \supset \neg \text{pintar } x)$$

De esa manera, la ley de inercia del sentido común corresponde con la expresión (4) en vista de que se asume que siempre se conservará la propiedad '*rojo*' a lo largo del tiempo a menos que se ejecute '*pintar*' (en este caso específico). Ahondando en este ejemplo, cabe mencionar la importancia de la condición para enseñarle al sistema que las propiedades solo cambian cuando explícitamente la ejecución de una acción lo demande o lo requiera. Dicho de forma más general, se asume que las propiedades solo cambian en el tiempo como efecto particular de alguna acción. Si no se ejecuta ninguna acción, entonces es correcto afirmar que la propiedad se mantendrá como tal en el futuro indicado en el sistema. O, en otras palabras, si

²⁴ Ver sección 1.4.

la propiedad no cambia a lo largo del tiempo, entonces no se ha ejecutado ninguna acción respecto de esta.

En síntesis, la formalización de la ley de inercia del sentido común en la notación sugerida de la lógica temporal métrica sería así:

$$\forall n \forall x ((\text{propiedad } x \supset G^n \text{ propiedad } x) \supset \neg \text{acción } x)$$

El alcance de esta tesis radica en presentar, ampliar y contrastar las problemáticas que presenta el tratamiento lógico y filosófico del problema de marco; y este segundo capítulo fue dedicado a establecer una formalización que permita una mejor comprensión del problema, sobre todo, con una visión lógica. Uno de los trabajos que considero valioso para ello es la ley de inercia del sentido común, postulada por Shanahan. Sin embargo, dado que su estudio incluye mucha más matemática, me limité a la adaptación de su criterio y la simbolización que otros autores (ya mencionados en la primera sección de este segundo capítulo) postularon para agudizar la mirada hacia la posible o las posibles soluciones.

Capítulo 3. Reflexión y contrastes

Recapitulando a la cuestión inicial de esta tesis, el problema de marco demanda la formalización de los mecanismos que permiten a un agente inteligente inferir qué propiedades no cambian cuando se ejecuta una acción. El reconocimiento de estos no-efectos de las acciones fue inicialmente explicitado con los axiomas de marco, pero no es una solución factible si se tienen muchas propiedades y acciones debido al problema de la vastedad de la información. Como ya mencioné, Shanahan (1997) profundizó ampliamente en una amplia cantidad de trabajo matemático al respecto sin encontrar alguna solución completamente satisfactoria. Inclusive, cabe tener presente la llamada de atención que efectúa Lifschitz acerca de dicha formalización. Este autor sostiene que, dadas las diferencias entre lo que es esperado y lo que es verdad, los matemáticos se aseguran de verificar todas sus conjeturas, no importa cuán plausibles, mediante el razonamiento. Por ello, él advierte que es crucial para mayor progreso que los estudiantes del conocimiento del sentido común estén preparados para hacer matemáticas difíciles (2000: 267)²⁵. Este es el

²⁵ “Because of the differences between what is expected and what is true, they [mathematicians] make sure to verify all their conjectures, no matter how plausible, by reasoning. It is crucial for further progress that students of commonsense knowledge be prepared to do difficult mathematics.” (Tr. p.)

motivo por el que la formulación que ofrezco de esta ley puede resultar útil para entender la temporalidad al afrontar el problema de marco²⁶.

Sin embargo, los presupuestos de esta tesis pueden ser cuestionados de variada forma. En este tercer capítulo trataré algunas de las bases filosóficas del muy aclamado problema de marco.

3.1. Teoría computacional de la mente

En primer lugar, la perspectiva que tomo del problema de marco está sumergida en la teoría computacional de la mente, según la cual las operaciones cognitivas son equivalentes a computaciones. Y no solo esto, sino que la versión clásica de esta teoría postula que los procesos mentales son literalmente computaciones o, por decirlo de otra manera, manipulaciones de símbolos (Rescorla 2020).

Manteniéndonos en este espíritu, un punto importante es la contraposición que Shanahan (2019) propone entre la tradición simbolista y el conexionismo en la literatura de la inteligencia artificial. En el primer caso se conceptualiza a la mente como un aparato de símbolos formales que opera siguiendo ciertas reglas de inferencia como lo hace Fodor. En cambio, en caso del conexionismo se procura que las estructuras mentales estipuladas correspondan con su correlato neuronal como lo hacen los Churchland; de ahí que venga el nombre ‘redes neuronales’ cuando se habla de cierto tipo de inteligencia artificial basada en las neurociencias. En este contraste de posturas podemos notar que el problema de marco viene precedido por la primera de estas perspectivas, es decir, la del paradigma simbólico. De esta manera, resultaría interesante evaluar la posibilidad de una formulación formal de nuestro problema de interés siguiendo el paradigma conexionista. Sin embargo, esta sería una exploración que escapa al eje central de esta tesis.

Teniendo en cuenta lo mencionado, y ampliando la perspectiva, algunos críticos de esta y de la excesiva racionalidad que implica la teoría computacional de la mente piensan lo siguiente:

La inteligencia humana no debería ser vista desde la perspectiva de la racionalidad para nada. De hecho, la racionalidad debería ser degradada dentro de las ciencias cognitivas. La inteligencia humana debería, en cambio, ser vista biológicamente: la inteligencia humana es solo una respuesta evolutiva al problema de sobrevivir en un ambiente —esta es racional en la medida en que funciona. (...) El problema con la agenda racionalista es que la racionalidad ideal es incompatible con la psicología mecánica, computacional. Cualquier teoría racional,

²⁶ Además, generar una vista formal de algún problema tratado epistemológicamente es bastante útil en materias pedagógicas. (nota del autor)

por lo tanto, está comprometida con la visión de que cuanto más inteligente es el organismo, menos mecánico es y, entonces, su inteligencia es menos explicable. (Dietrich/Fields 1995: 279)

Entonces, podemos notar que estos autores, Dietrich y Fields, apostarían por un mayor estudio de los procesos mentales manteniendo el fuerte y estricto correlato neuronal en cuanto a la aplicación que hacemos de la inteligencia al resolver problemas como los que incluyen al problema de marco. Es decir, ellos tienen una postura que simpatiza con el conexionismo.

3.2. Consideraciones sobre el cambio

Cabe notar que la ley de inercia del sentido común es cuestionable en cierta medida. Tal vez, en primera instancia, asumir que las propiedades no cambian a menos que ocurra una acción parece sensato. No obstante, esta postura tiene otra idea implícita que debería llamar nuestra atención. Asume que el cambio siempre es efecto de una causa externa, es decir, no considera que una propiedad pueda cambiar por sí misma, sino que requiere una acción (externa a ella, o al objeto que la instancia) para que cambie. Este es un punto dispuesto a debate sobre el cambio temporal de las propiedades y la causalidad, pero escapa a los intereses particulares de esta tesis.

Sin embargo, la historia muestra que el problema de marco constituye uno de los mayores problemas epistemológicos que la inteligencia artificial, como disciplina de investigación y desarrollo, ha conseguido mostrar. Poder determinar qué cosas o qué propiedades nos cambian cuando se ejecuta una acción determinada parece simple, pero la realidad es tan compleja y vasta que requerimos más herramientas formales sumamente detallistas para obtener el mayor provecho a los experimentos mentales que hacemos quienes pretendemos abordar la cuestión con holgura.

Por otra parte, y como un agregado, viendo la historia en retrospectiva, “Prior siempre creyó que su lógica temporal podría un día servir de utilidad a otras disciplinas (posiblemente en física matemática). Cuando llegó la demanda exterior para la lógica temporal fue del campo de la lógica computacional” (Copeland 1999: 9). Y, efectivamente, esta tesis también se encuentra en esa premonición de Prior.

Conclusiones

En resumen, dentro del estudio del razonamiento del sentido común, me topé con el problema de marco y lo abordé desde una perspectiva lógica y filosófica, a pesar de que

conocer un poco de su historia permite comprender el desarrollo de los planteamientos, perspectivas y criterios propuestos. Así, analicé las posturas e ideas generadas a partir de la problemática y ofrecí un modo de abordarla. Apoyándome en las distinciones de Sprevak (2005) y los criterios de Morgenstern (1996) pude esclarecer las confusiones comunes al abordar el problema de marco y, asimismo, quise acentuar los requerimientos de alguna efectiva solución, respectivamente. Encontré el uso de la lógica de Prior (1967a) para el problema de marco directamente en el trabajo de Silenzi (2011b), pero el lenguaje formal que preferí usar está basado en Øhrstrøm y Hasle (1995) y Creswell (2013). Finalmente, ofrecí algunas críticas válidas hacia la teoría computacional de la mente y los presupuestos de la aplicación de la lógica temporal métrica a la ley de inercia del sentido común y, más ampliamente, al problema de marco. Y, lo que más claramente queda expuesto en esta tesis es que hay más desconocimiento que conocimiento en cuanto a las herramientas formales apropiadas para modelizar la riqueza del pensamiento y razonamiento humano, por lo tanto, la tarea difícil de encontrarlas y saber aplicarlas a nuestro misteriosamente llamado “sentido común” nos permite ahondar en algunas fecundas ideas filosóficas con motivo del problema de marco.

Bibliografía

Aliseda-Llera, A. *Seeking explanations: Abduction in logic, philosophy of science and artificial intelligence*. PhD thesis on Philosophy and Symbolic Systems. USA: University of Stanford, department of philosophy, 1997.

Bell, John L., "Continuity and Infinitesimals", *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/sum2017/entries/continuity/>, Summer 2017, consultado el 6 de agosto de 2021.

Brown, Frank M., *The Frame Problem in Artificial Intelligence*. Elsevier Inc., 1987.

Copeland, Jack, “Arthur Prior”, *Stanford Encyclopedia of Philosophy* (edición 1999), traducido por Manuel González Riquelme, <https://es.slideshare.net/Filomatic/stanford-encyclopedia-of-philosophy>, consultado el 8 de diciembre de 2019.

Cresswell, Max, “Predicate metric tense logic for ‘Now’ and ‘Then’”. In *Journal of Philosophical Logic*. Vol. 42, N. 1, pp. 1-24. Springer, 2013.

Dennett, Daniel “Cognitive Wheels: The Frame Problem of AI” in *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, edited by Zenon W. Pylyshyn. pp. 41-64, 1987.

Dietrich, E. / Fields, C., “The role of the frame problem in Fodor’s modularity thesis: a case study of rationalist cognitive science”, in *Journal of Experimental & Theoretical Artificial Intelligence*, 7:3, pp. 279-289, <https://philarchive.org/archive/DIEROT>, 1995, consultado el 9 de diciembre de 2019.

Dreyfus, Hubert, *What computers can't do: The limits of artificial intelligence*, New York: Harper Colophon Books, 1972.

Fodor, Jerry, *The modularity of mind*. Cambridge: The MIT Press, 1983.

Fodor, Jerry, "Modules, frames, fridgeons, sleeping dogs, and the music of the spheres", In *The robot's dilemma* (Pylyshyn, ed.), pp. 139-149, Norwood, New Jersey: Ablex, 1987.

Ford, K. / Hayes, P. (eds.), *Reasoning agents in a dynamical world: the frame problem*, London: JAI Press, 1991.

Ford, K. / Pylyshyn, Z. *The robot's dilemma revisited: The frame problem in Artificial Intelligence*, Norwood: Ablex Publishing Corporation, 1996.

Hume, David, *A Treatise of Human Nature*, 1739, Reprinted from the Original Edition in three volumes and edited, with an analytical index, by L.A. SelbyBigge, M.A., Oxford: Clarendon Press, 1896.

Lifschitz, Vladimir, "Book review. M. Shanahan, Solving the Frame Problem." In *Artificial Intelligence*, 123, pp. 265-268. Elsevier, 2000.

McCarthy, J. / Hayes, P., "Some philosophical problems from the standpoint of artificial intelligence". Computer Science Department, University of Stanford, <http://www-formal.stanford.edu/jmc/mcchay69.pdf>, 1969, consultado el 28 de septiembre de 2020.

McDermott, Drew, "AI, logic and the frame problem" in *The Frame Problem in Artificial Intelligence* (Brown, ed.), pp. 105-118, California: Morgan Kaufmann Publishers, Inc., 1987.

Morgenstern, Leora, "The problem with solutions to the frame problem". In *The Robot's Dilemma Revisited: The Frame Problem in AI*. Ford & Pylyshyn (eds.). Ablex, Norwood, 1996.

Moro, Rodrigo / Silenzi, María Inés, "El problema de marco y dos programas rivales en psicología cognitiva" En *Contrastes. Revista Internacional de Filosofía*, vol. XXII, n. 1, pp. 101-117, Universidad de Málaga, España, <http://www.revistas.uma.es/index.php/contrastes/article/view/3418/3128>, 2017, consultado el 16 de octubre de 2019.

Mueller, Erik "Chapter 5: Common sense law of inertia". In *Commonsense Reasoning. An Event Calculus Based Approach*, 2nd edition, pp. 77-90, USA: Elsevier, Morgan Kaufmann, 2015.

Øhrstrøm, Peter / Hasle, Per, *Temporal Logic. From Ancient Ideas to Artificial Intelligence*. Netherlands: Kluwer Academic Publishers, 1995.

Prior, Arthur, *Past, Present and Future*. Oxford: Clarendon Press. https://archive.org/details/PastPresentAndFuture_201805, 1967a, consultado el 8 de diciembre de 2019.

Pylyshyn, Zenon (ed.), *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood: Ablex Publishing Corporation, 1987.

Rescorla, Michael, "The Computational Theory of Mind", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/fall2020/entries/computational-mind/>, Fall 2020 edition, consultado el 2 de agosto de 2021.

Samuels, Richard, "Classical computationalism and the many problems of cognitive relevance", *Studies in History and Philosophy of Science* 3, pp. 280- 293, 2010.

Shanahan, Murray "Explanation in the Situation Calculus", In *Proceedings IJCAI*, 93, pp. 160-165, 1993.

Shanahan, Murray, *Solving the frame problem: a mathematical investigation of the commonsense law of inertia*. Cambridge, Massachusetts: MIT Press, 1997.

Shanahan, Murray, "The frame problem". *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). <https://plato.stanford.edu/entries/frame-problem/>, Spring 2016 edition, consultado el 28 de septiembre de 2020.

Silenzi, María Inés, "El problema de marco y la investigación en las Ciencias Cognitivas" En *Actas de las IV Jornadas de Investigación en Humanidades. Homenaje a Laura Laiseca*, pp. 431-437, Universidad del Sur, 2011a.

Silenzi, María Inés, "El problema de marco: formalización de sistemas dinámicos en agentes artificiales". En *Revista Iberoamericana de Argumentación*, 3, pp. 1-20. Revista electrónica, Universidad Autónoma de Madrid, <https://revistas.uam.es/index.php/ria/article/view/8223/8563>, 2011b, consultado el 28 de septiembre de 2020.

Silenzi, María Inés, "¿En qué consiste el problema de marco? confluencias entre distintas interpretaciones" En *eidos* nº 22, pp. 49-80, 2015b.

Simons, Peter, "Jan Łukasiewicz". In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <https://plato.stanford.edu/entries/lukasiewicz/polish-notation.html>, Spring 2017 edition, consultado el 10 de diciembre de 2019.

Sprevak, Mark, "The frame problem and the treatment of prediction" In *Computing, Philosophy and Cognition*. L. Magnani & R. Dossena (eds.), pp. 349-359, London: King's College Publications, 2005.

Thielscher, Michael, "The Qualification Problem: A solution to the problem of anomalous models". In *Artificial Intelligence* 131, pp. 1-37, 2001.