

Racionalidad, lógica y heurística  
Raymundo Morado  
Instituto de investigaciones filosóficas  
Universidad Nacional Autónoma de México  
México, D. F., México,  
[morado@servidor.unam.mx](mailto:morado@servidor.unam.mx)

Leah Savion  
Departamento de Filosofía  
Universidad de Indiana  
Bloomington, Indiana, Estados Unidos  
[lsavion@indiana.edu](mailto:lsavion@indiana.edu)

Resumen: La noción de racionalidad es crucial para las ciencias de la computación, la inteligencia artificial, economía, leyes, filosofía, psicología, antropología, etc. La mayoría sí no es que todas esas disciplinas presuponen la capacidad de inferir de manera lógica de los agentes.

Las teorías de la racionalidad tienden hacia dos extremos:

O bien presuponen una inalcanzable capacidad lógica o tienden a minimizar el rol de la lógica, debido a una basta ejecución de inferencias falaces. Analizaremos algunos presupuestos en la visión de la lógica clásica y sugeriremos evidencia teórica y empírica para el lugar de la inferencia heurística en la teoría de la racionalidad.

Propondremos (1) bosquejar una nueva teoría de la racionalidad que incluya la noción clave de capacidad lógica como un factor necesario pero realista.

(2) Expandir la noción de inferencia para que incluya la inferencia no deductiva, específicamente la no monotónica, y (3) Enfatizar el rol lógico de heurística inferencial y restringirlo a la economía cognitiva.

Palabras clave: Retractable, lógica, racionalidad, heurística, economía cognoscitiva

1.-Acuidad lógica como factor necesario en la racionalidad.

El concepto de racionalidad es altamente complejo, y a menudo involucra construcciones basadas en principios tomados de la física, ciencias sociales, psicología, evolución, economía, ciencias políticas, filosofía etc. Adoptar una creencia, obtener inferencias de ello, construir un sistema de valuación, y actuar según esas creencias y deseos, puede ser todo juzgado como racional o irracional. Este escrito se enfoca en la racionalidad de inferencia, dejando a parte otros factores en la formación de creencias y acciones. La racionalidad de la inferencia lógica tiene agudeza lógica como componente necesaria.

La agudeza lógica involucra mucho más que la habilidad de extraer conclusiones correctas. El agente lógico hace planes, discierne alternativas, descarta las irrelevantes, arguye, negocia, entiende argumentos desde diferentes puntos de vista, toma parte en razonamientos contra fácticos, evalúa la evidencia y acepta las consecuencias lógicas de esas creencias. La inferencia lógica es una parte del ser lógico (existencia lógica), la cual se torna solo una fracción de lo que es ser racional. Aún, el análisis de la inferencia lógica, su estructura y su contribución al entendimiento de el actual

razonamiento humano es un buen punto de partida para la teoría de racionalidad, por lo que ha sido objeto de rigurosos exámenes por siglos.

Tradicionalmente la lógica fue considerada como una descripción normativa de los trabajos de una mente ideal. Sabemos desde Aristóteles que la gente no razona en concordancia perfecta con cualquier sistema lógico aceptado. Sin embargo, hasta recién, tanto filósofos como psicólogos, sociólogos y antropólogos aceptaron un altamente idealizado modelo de razonamiento humano. Las leyes de la lógica clásica fueron consideradas al menos implícitamente como leyes del pensamiento.

La distancia entre los dictados de teorías lógicas y la actual inferencia humana han sido estudiadas extensamente desde 1960. Los resultados de estos experimentos, no obstante controversiales de interpretación, han sacudido la tradicional visión del hombre como ser racional (Aristóteles) "noble en razón, infinito en facultades (Shakespeare), exhibiendo un retrato desconcertante de las habilidades de inferencia humanas.

La reacción más común fue emplear la distinción de "competencia- desempeño" de acuerdo a la cual tenemos una innata y perfecta habilidad lógica, la cual puede ser capturada en

un conjunto de reglas, solo estropeado por el desempeño imperfecto en la aplicación de esas reglas debido a limitaciones humanas, biológicas, cognitivas y tal vez sociales. La dimensión normativa lógica figura como descripción del nivel de habilidad oculto por los obstáculos de ejecución. La habilidad fue a menudo descrita como algorítmica, así como constituida por un conjunto de reglas que garantizan la validez deductiva de una inferencia. Este modelo es usado por muchas teorías de "lógica mental" que discrepan en el núcleo preciso de las reglas de competencia lógica y su naturaleza (reglas sintácticas y de contenido). Muchos pronósticos hechos por estas teorías chocan con otras observaciones experimentales, que sugieren un nuevo modelo, en el cual la heurística se propone como un sustituto aclaratorio para algoritmos como el aparato central (core devices) y procedimientos de inferencia. Esto ha llevado a algunos autores a sostener que no puede ser un factor necesario en la racionalidad. Puesto que la lógica es considerada como siempre algorítmica y la heurística como no algorítmica.

Queremos resistirnos a la tentación de circunscribir a la lógica al paradigma de inferencia algorítmica. Dada nuestra convicción en la agudeza lógica como una parte esencial de la racionalidad, preferimos usar una noción extendida de lógica como una teoría general de inferencia que incluye a la heurística y permite falibilidad sin comprometer rigor y formalidad. La lógica clásica se torna como ejemplo de método heurístico.

2 Supuestos tradicionales para propiedades meta lógicas de la inferencia racional.

La estructura de la investigación en razonamiento humano en las últimas dos décadas es una consecuencia de el popular paradigma lingüístico, el paradigma lógico formal y el paradigma computacional. Por un tiempo, los lógicos se dejaron deslumbrar por los logros de Frege-Russel-Whitehead. La relación de razonamiento con teorías de lógica formal sugirió un cuadro de pensamiento humano hecho de componentes atómicos discernibles, conteniendo operaciones que funcionan recursivamente en una estructura semántica bien definida. Es común examinar la sintaxis de un lenguaje natural y del pensamiento inferencial independientemente de otros aspectos de la materia en cuestión. El paradigma de la

computación refuerza el cuadro de la mente como un análogo a un programa (base de datos + operaciones), donde las explicaciones pueden ser suministradas en el programa o en los niveles de implementación.

Logicidad, fue igualada a la habilidad de seguir reglas clásicas de inferencia para generar teoremas basados en axiomas. Pero la lógica clásica es insuficiente en muchas áreas de inteligencia artificial y ciencia cognitiva: planear, buscar, reconocer patrones, CWA, esquemas, escritura, marcos (frames), etc.

A continuación algunas de las suposiciones de la naturaleza de la inferencia racional, que se siguen de estos paradigmas:

1. Omniscencia lógica
2. Infalibilidad
3. Consistencia
4. Reglas libres de contexto
5. No tiempo, espacio u otros medios limitantes en la ejecución de una inferencia.

Omniscencia lógica:

Clásicamente, un sistema de creencias racional es cerrado bajo consecuencia lógica. La cerradura bajo consecuencia lógica es la propiedad de un conjunto de proposiciones de ser cerrado bajo consecuencia lógica. Esto es, se incluyen todas las consecuencias lógicas de subconjuntos de proposiciones. A esta propiedad lógica de subconjuntos corresponde una propiedad epistémica de agentes. Un agente es lógicamente omnisciente sí y sólo sí su conjunto de creencias es cerrado bajo consecuencia lógica.

Por ejemplo, en lógica clásica, la omniscencia lógica trae consigo la creencia en todas las verdades lógicas no importando que incomprendibles sean, puesto que por vacuidad, son consecuencia lógica de cualquier creencia.

La omniscencia lógica puede no ser deseable en presencia de inconsistencias. Si usamos lógica clásica, de una contradicción se deriva cualquier cosa y trivializa cualquier sistema de creencias. Un agente racional, al descubrir una contradicción, tendrá que parar de hacer inferencias hasta que la inconsistencia sea restaurada, lo que tal vez nunca ocurra. ¿qué tan racional debería ser tal curso de acción?

Nosotros somos capaces de seguir operando ante la presencia de conflicto en las creencias y solo rehusarnos a sacar conclusiones en pequeñas áreas, tan específicas como sea posible. Hemos

aprendido a vivir con errores, conjuntos de prioridades para la resolución de conflictos y estabilizar mecanismos de emergencia para garantizar una elegante degradación del output si no podemos ocultar el problema.

Como mencionamos antes, se espera que un agente racional obtenga consecuencias obvias, pero aún si la omniscencia lógica pudiera ser parte del modelo implícito de creencias, sea lo bastante fuerte para hacerlas explícitas y eso no es un desideratum racional si el agente es prono a inconsistencias.

#### INFALIBILIDAD

Un sistema inferencial infalible empieza con un conjunto de verdades lógicas y las procesa a través de reglas válidas que preserven la verdad. Desgraciadamente, nosotros no siempre partimos de suposiciones necesarias, nuestra información a menudo es falsa y casi siempre incompleta.

Por otra parte, reglas infalibles son insuficientes para muchas situaciones cotidianas y tienden a ser muy costosas computacionalmente. Como resultado, tenemos que recurrir a aproximaciones, estimaciones y heurística. El ideal de infalibilidad abre camino a la idea de plausibilidad.

#### CONSISTENCIA

En los sistemas de base de datos, mucho esfuerzo se encamina a garantizar y mantener la consistencia (algunas veces mal nombrado mantenimiento de la "verdad").

Esto es de suma importancia porque como se mencionó, en lógica clásica de una contradicción se sigue cualquier cosa. Por consiguiente una contradicción puede volver trivial a un sistema, infiriendo todas las proposiciones como ciertas. Pero trivialmente no es una consecuencia necesaria de inconsistencia. Si solo usamos lógica clásica no podemos escapar de la tarea de mantener la consistencia. Una estrategia más económica podría ser distinguir entre conclusiones racionales e irracionales del mismo conjunto de creencias inconsistentes. Luego, la lógica clásica no puede ser solo una métrica de la racionalidad.

Un sistema que pueda tener contradicción sin trivialidad es llamado para consistente. Dichos sistemas han sido propuestos en [9][10] y [11], para modelar teorías científicas donde conocidas inconsistencias no conllevan trivialidad, y para modelar interacciones de multi-agentes (diálogos, evidencia incompatible, información incompatible, etc.)

Los sistemas de inferencia humana son paraconsistentes en el sentido de que tenemos

creencias contradictorias no obstante el razonamiento continua por medio de el uso de la heurística sin colapso en trivialidad.

#### REGLAS LIBRES DE CONTEXTO.

Clásicamente, la lógica ha evitado el contenido semántico a favor de un contexto libre de tratamientos de inferencia. Pero la lógica puede ser informal y la racionalidad tiene que tratar con propiedades de inferencia semántica, como se ejemplifica en el tratamiento griego de "topoi", falacias y sofismas.

Muchas heurísticas son de contenido- específico o dominio- específico.

Algunas heurísticas se aprende por experiencia y muchas ejecuciones acertadas se deben a la familiaridad con parámetros contextuales. Estos parámetros son importantes si un agente debe reaccionar racionalmente a las altamente contextuales "variables del ambiente", por ejemplo aquellos involucrados en el procesamiento de lenguaje natural.

#### SIN LIMITACIONES DE TIEMPO ESPACIO O RECURSOS

Tradicionalmente en lógica no se toma ninguna previsión (disposición) para la limitación de los recursos de los agentes concernientes a categorización y recuperación, el tiempo necesitado para ejecutar una inferencia, el espacio en memoria para trabajar, o atención selectiva. Como se arguye en [12], una teoría de racionalidad debe tomar todo esto en cuenta. En muchas ocasiones no es racional mantenerse ocupado en cálculos que exceden una prudente asignación de recursos. Gastando tiempo en extensas reflexiones que podrían ser racionales en sí mismas pero a menudo terminan con uno.

#### HEURÍSTICA

Una comprensión realista de la agudeza lógica requiere la inclusión de factores heurísticos. La heurística es intuición, algunas veces precognición, principios o procesos cognitivos, que generalmente fomentan una rápida y eficiente codificación, inferencia, recuperación y producción de información (ver [13]). El término "heurística" se ha vuelto popular en las últimas

décadas como un término genérico para los procesos mentales no algorítmicos (ver [5]). En este escrito el término se extiende para incluir cualquier estrategia inferencial, consciente o inconscientemente adoptada. La inferencia heurística puede verse como una continua expresión de grados de certeza de la conclusión dada por las premisas disponibles.

Bajo esta concepción, la presencia y empleo de heurística cognitiva es inevitable para nuestra vida inteligente. El mundo en sí, nos presenta un arreglo desordenado de bits de datos ambiguo, a menudo no representativo y sin una clara estructura que permita inferencias lógicas correctas. Para sobrevivir necesitamos razonar acerca de nuestro ambiente a partir de información incompleta y fragmentada, limitados por poco tiempo y aún más limitada capacidad computacional. Los métodos algorítmicos para extraer conclusiones que ofrecen las teorías deductivas están dispuestos para entregar resultados correctos cuando se aplican con propiedad, pero son notoriamente lentos y demandantes en términos de trabajo cognitivo y espacio en memoria. La evolutiva necesidad de rápida acumulación de información impuso la existencia de la heurística inferencial. La velocidad de acumulación de conocimiento es tal vez tan importante para nuestra sobre vivencia como la precisión de la información que acumulamos, su interpretación y las inferencias que extraemos.

Este hecho puede explicar porqué somos generalmente tan malos para calcular pero tan brillantes para estimar cantidades, distancia y resultados. El trueque de la velocidad por la precisión es resultado de emplear procesamiento de información heurística y estrategias inferenciales que permiten la selección y simplificación de problemas con una cantidad razonable de tiempo y recursos.

La inferencia heurística, como prototipo (categorización mecánica para clasificar nuevos conceptos por su grado de similaridad a un concepto típico central ya conocido) representabilidad (Aplicación de un criterio de semejanza simple a una nueva tarea) disponibilidad (la regla que dicta razonar con la información disponible) y anclaje. (adherencia a una presentación inicial en el proceso de solución o comprensión de problemas) proveen una herramienta necesaria para la supervivencia de proceso de información mientras se garantiza economía cognitiva.

El modo en que entendemos y aplicamos los conceptos, el principio de economía cognitiva dice que nuestro cerebro está diseñado para enfrentar nuestras necesidades, el mundo y las limitaciones de nuestras herramientas cognitivas, procurando minimizar el trabajo cognitivo mientras acumula la información necesaria rápido, directamente o por inferencia, a menudo perdiendo precisión. Dado que no es racional esperar que un ser finito emplee solo algoritmos que limitarían seriamente la información obtenida en un determinado lapso, una teoría de la racionalidad debe considerar algunas veces racionalizar el uso de inferencias económicas que pierdan precisión a favor de cantidad y velocidad.

El efecto de las reglas heurísticas es probado en un extremo por infalibles métodos lógicos (como inducción completa o silogismos) y en otro por los espantosos dudosos y persistentes medios de brincar a conclusiones lógicamente injustificadas.

A la pérdida de precisión asociada a la heurística a menudo se le llama "tendencia" [Nota de Gaby: no supe si poner tendencia o prejuicio, por eso algunas veces aparece tendencia y otras prejuicio] (ver [13]) una tendencia marca los límites de una aplicación fallida de la aplicación de la relevancia heurística, y una sistemática tendencia al error. Por ejemplo, el principio de conservación representa la suposición de invariabilidad a través de algunas transformaciones. Este principio es usado efectivamente cuando su aplicación es tanto teóricamente correcta (" $p$  o  $q$ " tienen la misma contribución lógica que " $q$  o  $p$ ") u objetivamente correcta (verter líquido en un contenedor de distinta apariencia no afecta su cantidad). Cuando el mismo mecanismo cognoscitivo hace que uno acepte que " $p$  si  $q$ " es lógicamente equivalente que " $q$  si  $p$ ", o que cambiar la temperatura de los líquidos no afecta su cantidad se considera como tendencia. La prototipicidad, que a menudo es una efectiva y necesaria heurística para copiar información con deficiencia, se convierte en una inconveniencia cuando Piolín no vuela.

Los ejemplos más familiares de prejuicio asociados con inferencia heurística aparecieron en la literatura por los años 70: experimentos y observaciones mostraron que la gente crónicamente malinterpreta eventos aleatorios como representativos, admite que evidencia que ratifica no lo hace, comete falacias deductivas

como "afirmación del consecuente" mientras evita la aplicación de reglas válidas como modus tollendo tollens.

La gente interpreta eventos irrelevantes como justificar sus bien fundadas ideas erróneas, basándose en anécdotas que soportan sus creencias mientras pasan por alto evidencia hostil y sustentan sus sorprendentes teorías ingenuas del "sentido común" con adecuadas frágiles explicaciones.

Esas tendencias aparentemente inevitables fueron usadas para fundamentar dos incompatibles y erróneas afirmaciones:

- (1) La heurística no puede ser parte de la racionalidad porque sus tendencias, ponen boca abajo la naturaleza algorítmica.
- (2) La heurística puede ser parte de la racionalidad pero la se sacrifica, puesto que la heurística no puede ser formalizada, por otra parte la prevaecía de tendencias indica que la ilogicidad parece ser inevitable.

Nuestra posición es que nosotros somos más racionales mientras más (y mejor) heurística, y menos (y menos dañinas) tendencias tengamos. La conclusión de la sección previa mostró la inaplicabilidad de los presupuestos clásicos a sistemas de inferencia racional. La heurística no sólo se necesita como herramienta para copiar rápidamente gran cantidad de información, y menos costoso en términos de trabajo cognoscitivo - ello tal vez también evite el costo de la trivialidad discutido arriba. A cambio de eliminar la logicidad de las teorías de la racionalidad (a favor de un discurso resbaladizo de supervivencia, adherencia a reglas sociales, etc.) preferimos acercar la inferencia lógica y la heurística con la ayuda de puentes como la economía cognoscitiva.

#### 4.- Observaciones sobre formalización.

Un agente epistémico capaz de enfrentar aún los mínimos retos del mundo real (sea computadora o humano) debe ser capaz de manejar descripciones incompletas y/o inconsistentes acerca del estado de la cuestión actual. Normalmente, usamos reglas que, no obstante son retractables, garantizan un mínimo de racionalidad en nuestro razonamiento. La lógica clásica provee guía para incrementar información explícita a través de consecuencia lógica, y nos permite retractar información con sus

principios de Reductio ad absurdum y Modus Tollendo Tollens.

Desafortunadamente, la mayoría de las formalizaciones ponen énfasis en un modelo deductivo tradicional de inferencia racional en el cual simplemente agregamos creencias cuando la información se incrementa, pero nunca las perdemos.

Decimos que un operador de consecuencia es monotónico si y sólo si un conjunto de creencias  $X$  puede ser subconjunto de  $Y$  y  $Cn(X) \subseteq Cn(Y)$ . Esto es, añadiendo información a  $X$  podrían eliminarse previas consecuencias. Por ejemplo, las lógicas clásica intuicionista y modal son monotónicas porque la adición de información no afecta la validez de las inferencias obtenidas previamente.

Ejemplos tradicionales de formalismos no monotónicos incluyen aquellos para inducción científica y abducción, probabilidad y estadística. Ejemplos de inferencias no monotónicas en ciencias de la computación son la negación por falla en programación lógica y la suposición del mundo cerrado (SMC) en la administración de bases de datos. La SMC tiene paralelo con la habilidad humana de obtener conclusiones con bases de información insuficiente tratándola como si fuera completa. El razonamiento de la ignorancia es a menudo una buena estrategia, porque muchos hechos son tan vistosos que su ausencia en su reporte cuenta como evidencia contra su ocurrencia. La gente continuamente infiere desde información que pudiera ser en principio incapaz de completez. En esos casos un comportamiento no racional pudiera ser no inferir. Una señal de racionalidad es la habilidad de revisar y unir nuestras conclusiones provisionales sin detener el proceso de inferencia.

Podemos aún hacer la afirmación normativa de que para que un agente con limitaciones cognoscitivas sea racional, algunas de sus conclusiones deben ser retractables. Así, necesitamos modelos que incorporen un estatus provisional de nuestras creencias inferidas. Un modelo de racionalidad que no contiene retractabilidad, un modelo puramente monotónico falla en la respuesta a esta necesidad.

Desde los 80 tenemos formalismos adicionales para modelar diferentes aspectos de la inferencia no monotónica. Por ejemplo, [15] añade a la lógica clásica esquemas de circunscripción que producen el efecto de suposición de mundo cerrado. La circunscripción limita el dominio de la extensión de un predicado, y elige modelos mínimos.

[16] usa una lógica no monotónica modal con un operador lógico M que señala algo como "Es posible hasta donde el sistema conoce". Las alternativas populares incluyen el uso de lógica por default en [17], lógica auto epistémica en [18] y modelos de preferencia en [19].

La heurística a menudo ejemplifica "razonamiento no monotónico" porque en muchos casos produce creencias retractables en el sentido de nueva evidencia, ya que este comportamiento es ya al menos parcialmente formalizado en las antes mencionadas lógicas no monotónicas, la acusación de que la heurística no es formalizable pierde credibilidad.

#### 5 CONSECUENCIAS DE LA NOCIÓN DE RACIONALIDAD.

El proyecto de construir una nueva teoría de la racionalidad debe aspirar a una explicación de los mecanismos de inferencia en términos de múltiples construcciones teóricas. Un modelo tal acomoda el uso de reglas contextuales de inferencia tan bien como reglas sintácticas, permite el empleo de mecanismos pragmáticos (como modelos mentales, imagery, esquemas, escripts) y heurística retractable, toma nota de una larga variedad de limitaciones cognoscitivas (no solo aquellas asociadas con la capacidad de memoria y el tiempo de cómputo), reconoce tendencias generales y provee una explicación para el "razonamiento desviado" en términos de procedimientos no monotónicos.

Es posible discernir dentro de el fenómeno del razonamiento el "código" la "representación" las "estrategias" la "competencia" y el "desempeño". Nosotros presentamos que el origen de la limitación en el proceso total de inferencia es el principio biológico de economía cognoscitiva. El cerebro no simplemente graba y presenta aspectos de la realidad externa y nuestras reacciones ante eso. Categoriza la información, reduce su complejidad dentro

de las estructuras contextuales a una escala manejable y la organiza permitiendo una recuperación efectiva. Estos procesos son posibles (Casi totalmente) gracias al mecanismo costo- eficiencia que llamamos economía cognoscitiva.

El efecto de esta necesidad crucial de economía en el razonamiento no ha sido suficientemente explorado en la literatura. La economía cognoscitiva, creemos juega un significativo papel en la descripción del aspecto de "conocimiento" de los procesos de inferencia (información, representación, conceptualización) y su manipulación. La arquitectura cognoscitiva humana ha diseñado facciones que promueven potencia y velocidad a expensas de algo de fiabilidad. Un sistema racional debería ser capaz de producir un largo número de conclusiones en aras de superar la ignorancia que es nociva para la supervivencia.

#### 6 CONCLUSIONES.

En este escrito sugerimos una visión de la racionalidad que asigna un rol lógico al razonamiento heurístico. Empezamos sosteniendo que cualquier teoría de razonamiento (racional) incluye la noción de habilidad lógica. Tradicionalmente esta noción ha sido confinada a inferencia deductiva algorítmica. Un examen crítico de las suposiciones subyacentes de esta tradición exhibe la ineficiencia de cualquier teoría de racionalidad que limite por sí misma a la lógica clásica. Como resultado expandimos la noción de logicidad para que incluya razonamiento no- deductivo, y mencionamos recientes formalizaciones de relaciones de consecuencia no monotónica. Una teoría de la racionalidad no tiene que dar el ideal de formalización solo porque alberga inferencias heurísticas, aun si no todas las heurísticas son formalizables.

La aun relativamente desnuda estructura de una concepción dinámica de inferencia heurística que la gente emplea junto con el aclaratorio poder de los principios de economía cognoscitiva pueden ser usados como fundamentos para una teoría de inferencia racional.

Traducción de Gabriela Escudero.

Versión: 5 de julio de 2004.