

Parte II

Apéndices

Apéndice A

Emociones y razón

A.1. Damasio: Hipótesis del marcador somático

Para la defensa de “una racionalidad de las emociones”, tal como propone Damasio en [Damasio, 1994], se utilizan conclusiones que provienen del campo de la neurofisiología. Damasio intenta dar explicación a los comportamientos de algunos de sus pacientes que han sufrido daños en el lóbulo central. Estos, se vuelven emocionalmente planos – es decir, no muestran ningún signo de emoción – ante situaciones sensibles o trágicas, y aunque conservan toda su batería de facultades cognitivas, pierden la capacidad para tomar decisiones. Damasio llega a la conclusión de que *“la defectuosa capacidad de toma de decisiones se debe a la ausencia de emociones”*, registrando evidencias experimentales acerca del comportamientos nada exitosos por parte de estos sujetos, en dos aspectos principalmente:

- Emplean, **irracionalmente, cantidades de tiempo desmedidas** en tareas triviales, retrasando innecesariamente la decisión.
- Se **toman malas decisiones** —en vez de retrasarlas, como en el punto anterior— por la falta de motivación que manifiestan sus *representaciones mentales sobre estados futuros*, impidiéndoles preparar respuestas anticipatorias.

Ante estos resultados, podemos enunciar dos formulaciones de la propuesta de Damasio (él utiliza ambas, una más fuerte y otra más débil, en partes diferentes de [Damasio, 1994]):

- Formulación covariante (o débil): *“Los poderes de la razón y la experimentación de la emoción decaen al mismo tiempo”*.
- Formulación causal (o fuerte): *«La reducción de la emoción puede constituir una fuente importante de conducta irracional»*.

Damasio desarrolla el modelo de *“marcadores somáticos”* para explicar, desde un punto de vista neurofisiológico, el porqué del comportamiento de sus pacientes. Su “hipótesis del marcador somático” propone que la participación emocional en el proceso de deliberación práctica y toma de decisiones, lejos de constituir una interferencia perturbadora, es más bien una condición de posibilidad del mismo. Si nuestra deliberación práctica tuviese que desarrollarse según los criterios de la teoría de la decisión, no podríamos tomar decisiones adecuadas a la urgencia, debido a la complejidad del cálculo requerido y al gran número de cursos de acción alternativos a evaluar. Este número, sin embargo, es drásticamente reducido, según Damasio, por la participación de las emociones en la selección de alternativas que serán tenidas en cuenta. ¿Cómo ocurre esto?

A.1.1. Modelo emocional de Damasio

Hemos nacido con cierta maquinaria neuronal innata capaz de generar estados emocionales en relación con ciertas clases de estímulos. Pero además de estas capacidades innatas, poseemos la habilidad de formar conexiones sistemáticas entre categorías de objetos y situaciones, generando un nuevo tipo de reacciones emocionales.

Emociones primarias

Damasio reconoce que estamos programados para responder con una emoción cuando se perciben determinadas características de los estímulos del mundo. Denomina a estas pautas de reacción emocional automática emociones primarias. Este tipo de emociones son innatas, preorganizadas, y pueden conseguir algunos objetivos útiles por sí mismas. Las principales consecuencias de tener emociones primarias, son dos:

- Se producen disparos químicos y reacciones viscerales que dan lugar a la aparición de un cierto estado corporal que nos prepara para reaccionar ante el estímulo. Se tiene así, un mecanismo de respuesta automática.
- Se altera el procesamiento cognitivo de una manera que identifica conscientemente la sensación del estado emocional). ¿Por qué ocurre esto? La tesis de Damasio es que la conciencia consigue un mayor nivel de protección. Ser consciente de nuestras emociones nos ofrece la flexibilidad de una respuesta basada en la historia particular de nuestras interacciones con el ambiente.

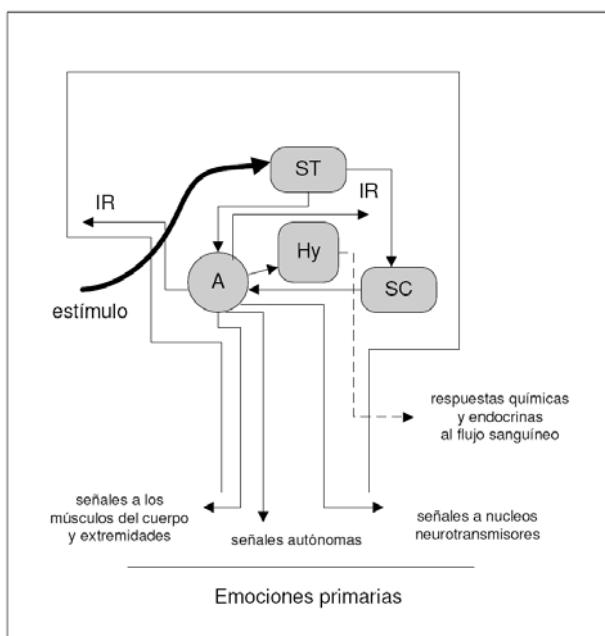


Figura A.1: Emociones primarias. Modelo de Damasio.

En la Figura A.1 se presenta un esquema del proceso de una emoción primaria. El perímetro simula el cerebro y el tronco cerebral. Después de que un estímulo apropiado active la amígdala *A* —vía el tálamo sensorial (*ST*) y el cortex sensorial (*SC*)—, se provoca una serie de respuestas diferentes: respuestas internas (*IR*), respuestas musculares, respuestas a las vísceras (*señales autónomas*), respuestas a núcleos neurotransmisores y respuestas al hipotálamo (*Hy*). El hipotálamo hace que se produzcan respuestas endocrinas y químicas que van al flujo sanguíneo.

Emociones secundarias

Pero el mecanismo de las emociones primarias no describe toda la gama de los comportamientos emocionales. Damasio considera que existen otros tipos de mecanismos, que se apoyan en el mecanismo de las emociones primarias, y que tienen lugar una vez que se han empezado a experimentar sentimientos y a establecer relaciones sistemáticas entre categorías de objetos y situaciones, por un lado, y emociones primarias, por otro. La estructura del sistema límbico no es suficiente para soportar estos procesos y requiere la participación del cortex prefrontal. La emergencia de una emoción secundaria podría describirse así:

- Se inicia a partir de consideraciones conscientes, deliberadas, acerca de una persona o una situación. Estas consideraciones se expresan en forma de imágenes mentales referidas a múltiples aspectos del objeto de análisis. Es estrictamente una evaluación cognitiva del contenido del acontecimiento.
- Algunas de las imágenes que se evocan son verbales (afirmaciones o informaciones referidas a atributos de la situación) y otras son no-verbales (aspectos, contextos). El sustrato neural que les da soporte consiste en diferentes estructuras topográficas que tienen lugar en diferentes cortezas sensoriales (corteza visual, auditiva, y otras).
- Estas imágenes que recogen el aspecto de tales objetos se emparejan, mediante redes de la corteza prefrontal, con determinadas situaciones que ha definido la experiencia individual de cada persona, en forma de estructuras disposicionales (tendencia *a actuar* o tendencia *a huir* según sean de *agrado* o *desagrado*). Es decir, estas representaciones disposicionales son adquiridas, no innatas, y reflejan la experiencia única de dichas relaciones en la vida del sujeto.
- De manera automática e involuntaria, la respuesta disposicional que se produce es comunicada a la amígdala que inicia —utilizando la maquinaria de las emociones primarias— la activación del sistema nervioso autónomo, del sistema motor para dar el aspecto externo de una emoción, los sistemas endocrino y péptido (dando resultado a cambios en estados corporales y cerebrales) y los núcleos neurotransmisores no-específicos en la parte frontal del cerebro y en la posterior basal.

Por tanto, el origen de las emociones secundarias es una consideración consciente y deliberativa, pero las respuestas son un conjunto de cambios en el estado corporal típicos de emociones primarias e innatas.

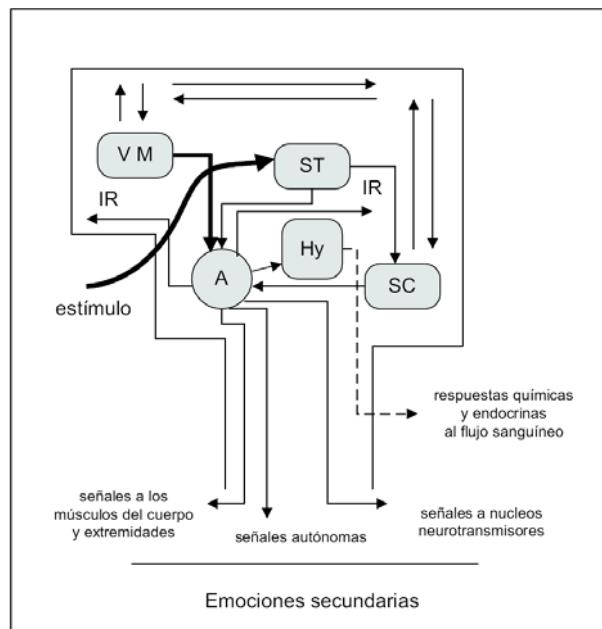


Figura A.2: Emociones secundarias. Modelo de Damasio.

En la Figura A.2 representamos esquemáticamente cómo se produce una emoción secundaria. Los estímulos pueden ser procesados directamente vía la amígdala pero ahora también son analizados por un proceso deliberativo que puede activar las cortezas frontales y prefrontales —simplificando podemos considerar fundamentalmente la *corteza ventromedial (VM)*—. En las emociones secundarias y para expresar su actividad la corteza ventromedial (*VM*) actúa vía la amígdala (*A*). En otras palabras, las emociones secundarias utilizan la maquinaria de las emociones primarias.

Podemos establecer una representación esquemática de la relación entre los mecanismos emocionales secundarios y la maquinaria de las emociones primarias, en el marco del agente como un sistema de control de tres capas, para ver cómo serían las descripciones y los procesos de información.

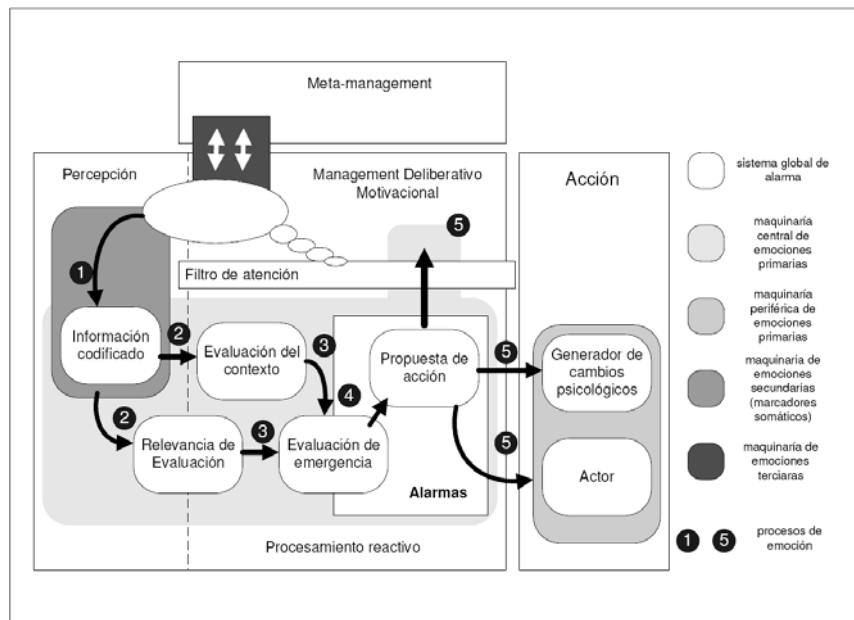


Figura A.3: Procesos emocionales y arquitectura de tres capas.

Conceptualmente, podemos diferenciar la maquinaria de las emociones primarias y los procesos de emociones secundarias, asociadas a:

- La maquinaria central, responsable de detectar y evaluar la relevancia de los estímulos, interrumpiendo procesos de atención, si es necesario.
- La maquinaria periférica, encargada de generar acciones sobre el exterior y cambios psicológicos en el sistema.

Estados corporales “como si”

Por tanto, y según lo hasta ahora expuesto, las emociones, primarias y secundarias, están vinculadas a una situación real, que afectó directamente nuestro estado corporal generando una representación disposicional (de agrado o desagrado) de la experiencia, y que, según Damasio, marcará nuestra tendencia futura cuando se evoquen de nuevo ese tipo de situaciones.

Damasio introduce otra de sus ideas más interesantes con el concepto de **bucle “como si”**. Encontramos en los procesos emocionales, dos tipos estructuras relacionales:

- **Bucles corporales.** Se encuentran en el mecanismo de generación de reacciones emocionales primarias y secundarias: se originan en la mente/cerebro (*señal percibida*) y provocan *efectos corporales* que de nuevo llegan a la mente/cerebro.
- **Bucles “como si”.** En éstos, el cerebro registra una imagen (*más débil*) de un estado corporal emocional, sin que se haya representado realmente en el cuerpo. Es decir, existen estructuras y dispositivos neuronales que nos permiten sentir “*como si*” tuviésemos un estado emocional en el que el cuerpo estuviera siendo activado.

Estos estados emocionales, que no son innatos y por tanto entran en el tipo de secundarios, no surgen de la asociación entre una determinada imagen mental y un estado corporal real, sino entre una imagen mental y un sustituto mental de un estado corporal.

Damasio ve los *bucles “como si”* como las representaciones simbólicas de estados somáticos. La Figura A.4 muestra los caminos de procesamiento de información activos en estado emocionales

que utilizan bucles corporales reales y bucles “como si”, para generar el estado de sentimiento de una experiencia emocional.

Los marcadores somáticos, por tanto, aparecen de dos maneras:

- Como recuerdos con pesos implícitos, que utilizan la amígdala y la maquinaria de los sistemas de emoción primarios en bucles corporales reales.
- Como recuerdos con pesos explícitos sobre el éxito/fracaso de ciertas situaciones, que utilizan las regiones del hipocampo y *bucles “como si”*.

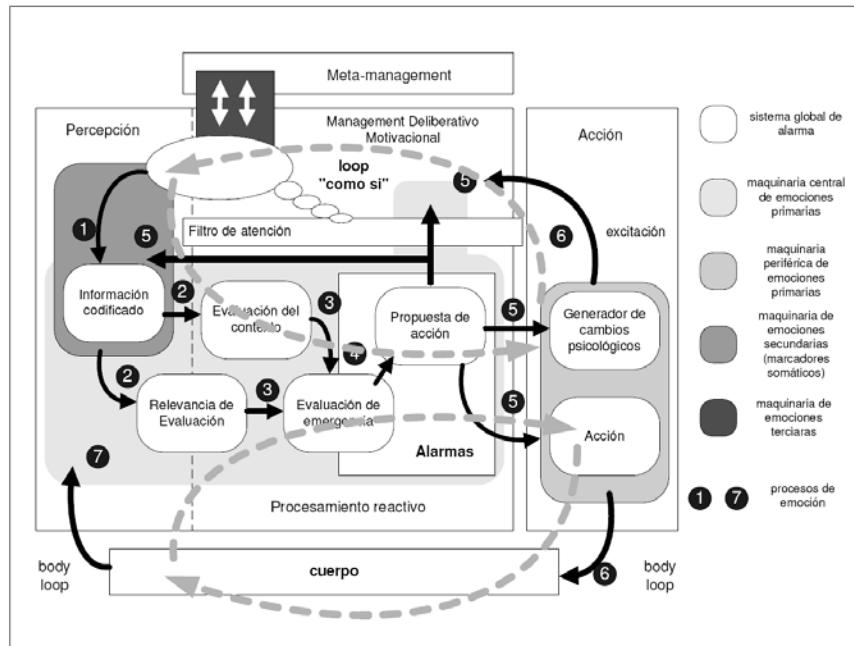


Figura A.4: Esquema de bucles reales y bucles “como si”.

A.1.2. La fría cognición no es suficiente

Consideremos una situación que exige hacer una elección: debemos decidir si seguimos adelante o no con un determinado acuerdo comercial. En el cerebro de un adulto, psicológicamente normal, se generan rápidamente un conjunto de supuestos sobre posibles opciones de respuesta y de resultados relacionados. Para nuestra conciencia, los supuestos están formados por múltiples y variadas imágenes mentales imaginarias. ¿Cómo escogemos? ¿Cómo elegimos?

La primera de las posibles respuestas proviene de una concepción tradicional racionalista de la toma de decisiones. El mecanismo para la correcta elección estaría basado en un análisis lógico de la situación (y por tanto, dejando a parte cualquier contaminación emocional), se seleccionaría la opción que maximizase el provecho subjetivo esperado, para lo cual:

- Se efectuaría un análisis coste/beneficio de cada una de las situaciones posibles.
- Se calcularía la relación entre beneficios inmediatos o recompensas a largo plazo.
- Se estimaría la magnitud y ritmo de las posibles recompensas, a lo largo del tiempo, para comparar las pérdidas potenciales.
- Se podría calcular la tasa de “depreciación” de los beneficios en función del tiempo.
- Se necesitaría comparar resultados de naturaleza distinta, llevándolos, de alguna manera, a una representación de valores comunes para que pudiera establecerse la comparación.

Si dicha estrategia es la única de la que disponemos, no funcionará.

En el mejor de los casos,

1. Nuestra decisión tomará un tiempo excesivamente largo, mucho más de lo que puede ser aceptable.
2. No usaremos correctamente la teoría de probabilidades y de la estadística en problemas de la vida cotidiana
3. No será fácil conservar en la memoria todos los análisis. La atención y la memoria funcional tienen una capacidad limitada.

En el peor de los casos, no llegaremos a tomar una decisión.

Lo que se extrae de la experiencia de Damasio con pacientes que tienen lesiones prefrontales, es que la fría estrategia lógica propuesta por la postura racionalista tiene más que ver con las decisiones que toman estos pacientes que tienen lesionada una parte del cerebro, que con las operaciones usuales en las personas normales. El modelo racionalista puro no es satisfactorio. Necesitamos una concepción alternativa.

La hipótesis del marcador somático

La concepción alternativa que buscamos puede proceder de la “hipótesis del marcador somático” [Damasio, 1994]. Consideremos de nuevo la situación que exigía decidir si seguimos adelante o no con un determinado acuerdo comercial. Se generan en nuestra conciencia rápidamente variadas imágenes mentales imaginarias sobre las posibles opciones de respuesta y de resultados relacionados. Veamos cómo explica Damasio lo que en ese momento sucede:

«antes de que apliquemos ningún tipo de análisis coste-beneficio a un problema, ocurre algo muy importante: cuando un resultado negativo que viene vinculado a una determinada opción de respuesta nos viene a la mente, por fugaz que sea, experimentamos una coronada desagradable. Como la sensación tiene que ver con el cuerpo, le doy al fenómeno el término de estado somático, [...] , y como nos indica una representación, lo he llamado indicador. [El] indicador somático [...] atrae la atención hacia el resultado negativo al que puede conducir una acción determinada y funciona como una señal de alarma automática que dice: “cuidado con el peligro que se avecina si eliges la opción que conduce a ese resultado”. La señal puede llevarte a rechazar inmediatamente la alternativa de acción negativa y puede hacer, así, que elijas otras opciones. Esa señal automática te protege de futuras pérdidas, sin mayor dilación, y te permite escoger entre menos alternativas. Es posible emplear un análisis de coste-beneficio y cierta capacidad deductiva, pero sólo después de que ese paso automático haya reducido sensiblemente el número de opciones. »

Básicamente, el modelo de Damasio viene a reconocer la existencia de indicadores somáticos positivos, que conducen a la elección de determinadas opciones, y negativos, que nos llevan a rechazar inmediatamente la alternativa de una acción considerada negativa.

¿Cuál es el origen de los marcadores somáticos en términos neurales? ¿Cómo hemos llegado a poseer estos dispositivos tan útiles? ¿Son innatos? Si no lo son, ¿Cómo surgieron?

Tal como hemos visto, nacimos con una maquinaria neural que nos permite generar estados somáticos de respuesta a estímulos (vasodilatación, hiperventilación, etc.) a determinadas clases de estímulos: es la *maquinaria de las emociones primarias*.

No obstante, los marcadores somáticos que empleamos para la toma de decisiones se crean, probablemente, tras un proceso de aprendizaje, al conectar clases específicas de estímulos con clases específicas de estados somáticos. Es decir, utilizan para su generación la *maquinaria de las emociones secundarias*.

«Los indicadores somáticos son [...] emociones y sentimientos [que] han sido ligados, mediante el aprendizaje, a los resultados futuros predichos de ciertos escenarios» [Damasio, 1994].

Dada la exposición anterior sobre la fisiología de las emociones, esperamos no sólo un mecanismo para los procesos de los marcadores somáticos, sino dos:

- Un mecanismo básico, tal que dado un estímulo cognitivo que se procesa en la corteza prefrontal y pasa a la amígdala, genere cambios y reacciones corporales que definan un perfil de estado somático, cuyo resultado sea registrado en la corteza somatosensorial, donde se le hace consciente en forma de sentimiento.
- Un mecanismo alternativo, donde se prescinde del cuerpo, y en el que la corteza prefrontal y la amígdala le dicen simplemente a la corteza somatosensorial que se organice en la pauta de actividad explícita que habría asumido si el cuerpo hubiera sido situado en un estado somático determinado. Para la corteza somatosensorial es como si estuviera recibiendo señales sobre un estado corporal real. La pauta de actividad “como si” no es la misma que si viniera generada por un estado corporal real, pero influye igualmente en la toma de decisiones.

Es probable que a medida que aprendíamos y nos ajustábamos “socialmente” a las situaciones y experiencias vividas, la mayor parte de nuestra toma de decisiones era moldeada por estados somáticos reales relacionados con castigo (*fracaso*) y recompensa (*éxito*). Pero a medida que categorizamos situaciones repetidas, se fue desarrollando otro nivel de automatización más económica. Las estrategias de toma de decisiones empezaron a depender *en parte de “símbolos” sobre los estados somáticos* y no sobre los estados mismos, y así seguimos fortaleciendo nuestro aprendizaje.

Además, podemos encontrar que los marcadores somáticos tienen más de una vía de acción sobre nuestras decisiones:

- Acción a través de la conciencia. Dado un estímulo, se activan las señales provocadas por estados corporales (o por los sustitutos “como si”) que se hacen conscientes y constituyen un sentimiento. Esta sensación influye conscientemente en nuestra decisión.
- Acción fuera de la conciencia. Dado un estímulo, se activan las señales emocionales, pero no pasan nunca a entrar en el foco de atención, sino que forman parte de una acción encubierta sobre los mecanismos que rigen, sin control voluntario, nuestras actitudes apetitivas (*de acercamiento*) o aversivas (*de rechazo*) hacia algunas situaciones.

Es obvio que en organismos cuyo cerebro no se encarga de la conciencia y el razonamiento, estos mecanismos encubiertos son la base del aparato de toma de decisiones. Son un medio para construir “predicciones” de resultados y predisponen a los dispositivos de acción del organismo para que se comporten de una determinada manera (lo que a un observador externo le puede parecer una elección).

Condiciones emocionales para la deliberación racional

La hipótesis de Damasio defiende que los marcadores somáticos permiten elegir entre menos alternativas. Este planteamiento no propone simplemente que ciertas situaciones lleven un peso (*en la forma de un marcador somático*) que nos permita tomar una buena elección, sino el hecho de que *existe algún mecanismo en nuestro cerebro que responde a esta evaluación como una señal de alarma para reducir nuestras opciones automáticamente*. En palabras del mismo Damasio [Damasio 1994]:

«*¿Qué es lo que los marcadores somáticos consiguen? Centran la atención en el resultado negativo que una acción dada puede alcanzar, y tienen funciones como señales de alarma automáticas que dicen: cuidado con el peligro si eliges la opción que nos lleva a este resultado. La señal puede llevar a rechazar, inmediatamente, el curso de acciones negativo y de este modo hacerle elegir entre menos alternativas».*

Neurofisiológicamente, para la activación de un marcador somático, el cortex prefrontal *elabora representaciones de las alternativas posibles*, y estas representaciones son rápidamente evaluadas con la ayuda de las estructuras límbicas, que generan un tipo de “sentimientos”, un remedio de las reacciones emocionales que despertarían en nosotros si fuesen reales. Por tanto, la evaluación emocional descarta con rapidez las alternativas que generan reacciones negativas y preselecciona como candidatas a la elección final aquellas que despiertan respuestas positivas.

¿Y qué ocurre si afrontamos un reto novedoso para el que no haya ninguna regla general disponible? Las emociones nos inducen, entonces, a tomar una decisión basada en algún rasgo

destacado de la situación. Damasio postula que ese rasgo tiene un gran valor predictivo a la hora de hacer una buena elección y que destaca algún aspecto de la situación para la que evoca una emoción concreta.

Por tanto, la hipótesis sobre la existencia de marcadores somáticos lo que propone es la existencia de mecanismos en nuestro cerebro que criban automáticamente entre varias opciones sobre las que tenemos que decidir, manifestándose como una especie de señal de alarma, que bloquea algunas de ellas.

A.2. De Sousa: Irracionalidad sin emociones

Ronald de Sousa lleva más allá la tesis de Damasio y sostiene que las emociones no son elementos perturbadores de nuestros procesos de reflexión racional. Al contrario, las emociones pueden fomentar la conducta racional en situaciones de indeterminación [De Sousa, 1987]. Sus argumentos se pueden resumir en las siguientes tesis:

1. Sujetos sin emociones tomarían decisiones irracionales.
2. En muchas situaciones los sujetos sin el empuje de las emociones no tomarían ninguna decisión o la retrasarían durante mucho tiempo. Esta idea, como se destaca en [Elster, 2002], se plantea que en algunos entornos la abstención o falta de inmediatez en la toma de decisiones puede convertirse en algo tan irracional como reaccionar sin criterio deliberativo.

Tal como están formuladas las tesis, no se afirma que las emociones tengan un papel causal para alcanzar una mejor decisión por el hecho de que se llegue antes que a través de un proceso de optimización basado en la deliberación, sin embargo la postura de De Sousa se encuentra muy cerca de una tesis más fuerte según la cual las emociones intervienen *causalmente* en la toma de decisiones.

En los casos citados, y siguiendo la perspectiva de De Sousa, la urgencia de las emociones nos ayuda a tomar decisiones óptimas. Dada las limitaciones de nuestra racionalidad, se acuden a otro tipo de reglas que indican la mejor acción.

Por tanto, la visión de las emociones, no es la de elementos al margen de la razón. La emoción no está haciendo lo que la razón no puede hacer, sino lo que la razón también podría hacer, pero de modo diferente. De Sousa y otros autores explican, con esta interpretación de las emociones como quedan en evidencia los defectos estructurales y formales de las teorías de elección racional que pretenden explicar el comportamiento de los agentes racionales en función del análisis de la probabilidad que conllevan todas las consecuencias posibles de cada una de sus acciones.

Según De Sousa, sin la participación de las emociones, los procesos de razonamiento práctico y la toma de decisiones sucumbirían a lo que en *inteligencia artificial* se denomina “el problema del marco” [McCarthy y Hayes, 1969]. En [De Sousa, 1987], nos dice:

las emociones serían el modo en que la naturaleza ha respondido al problema del marco en el caso de los seres humanos.

Apéndice B

Emociones e inteligencia artificial

B.1. Limitaciones de los mecanismos de racionalidad pura para explicar el comportamiento

En la inteligencia artificial, el agente inteligente se ha considerado clásicamente como un agente puramente racional, cuyo comportamiento busca optimizar la consecución de objetivos [Russell y Norvig, 2003] . Sin embargo, somos conscientes de que en nuestras acciones las emociones cumplen una función innegable en la construcción de la motivación (para la disposición a la acción) y en la interpretación de los sucesos [Ver apéndice A]. Si queremos entender el funcionamiento de las funciones mentales, no tenemos más remedio que discutir las extrañas relaciones de las emociones con la racionalidad. El modo tradicional que a lo largo de la historia ha considerado cuál era la relación entre racionalidad y emociones, *el Dualismo*, ha interpretado los fenómenos emocionales como *eventos perturbadores de la racionalidad*.

Si aceptamos que a la elección racional se le atribuye el propósito de tomar decisiones vía un criterio de optimización, éste debe darse:

1. En la acción escogida, que ha de ser óptima dados los deseos y las creencias del agente.
2. En las creencias, que han de ser óptimas dada la información que el agente tiene disponible.
3. En la cantidad de recursos dedicada a la adquisición de información, que ha de ser óptima dados los deseos del agente y sus creencias acerca de los costes y beneficios esperados de tal información.

Esta perspectiva de la racionalidad nos hace caer en el problema del marco [McCarthy y Hayes, 1969], un problema recurrente en inteligencia artificial y que consiste en la dificultad de operar lógicamente en un dominio en el que no se han definido explícitamente las condiciones del problema que no son afectadas por nuestras acciones, haciendo que los sistemas inteligentes tengan que procesar continuamente información que un ser humano simplemente ignoraría.

En esta descripción de la acción racional, las emociones podrán afectar a la racionalidad pudiendo perturbarla de dos maneras, incorporando en los procesos de decisión y razonamiento:

1. **Irracionalidad.** Según la perspectiva tradicional, las emociones son causa de irracionalesidad en la toma de decisiones, afectando a las acciones, a la formación de creencias y a la adquisición de información:
 - a) las emociones afectan a las “estimaciones de probabilidad” acerca de hechos que pueden suceder en un entorno incierto e indeterminado.
 - b) son la “causa” de que haya cierto nivel de creencia en la eficacia de acciones en la que uno no creería si estuviera en otras condiciones.
2. **Indeterminación.** En la información disponible para un sujeto o un sistema artificial, la indeterminación puede manifestarse a partir de las limitaciones de sus mecanismos para percibir el mundo, o por la falta de certeza que las evidencias del mundo puedan sugerirnos. Pueden adoptar, en nuestros procesos de decisión, dos formas:

- a) **Indiferencia.** Dado un conjunto de opciones posibles, y una relación *R de equivalencia* entre los elementos del conjunto indicado —que llamaremos relación de preferencia— diremos que dos opciones *a, b* son indiferentes, en términos de nuestra elección, si la relación de equivalencia es conmutativa, esto es, *si “a es al menos tan bueno como b”*.

$$\exists aRb, \exists bRa / aRb = bRa$$

- b) **Incommensurabilidad.** Dado un conjunto de opciones posibles, y una relación *R de equivalencia* entre los elementos del conjunto indicado —que llamaremos relación de preferencia— diremos que dos opciones *a, b* son incommensurables, *si “no están definidas las relaciones entre a y b”*.

$$\nexists aRb, \nexists bRa$$

La indiferencia es una propiedad excepcional, raramente observada en situaciones no triviales de la vida real. Sin embargo, la incommensurabilidad es una *fuente de indeterminación mucho más presente*. Podemos encontrar diferentes situaciones que nos llevan a condiciones de imposibilidad de elección, si nuestras decisiones sólo las basamos en criterios formales, por problemas de incommensurabilidad. Por ejemplo:

- Podemos pensar en casos en los que no podamos (por limitaciones técnicas o por la forma de presentarse la información en el medio) recoger suficientes evidencias como para hacer una elección justificada. En tales casos, los sujetos que actuasen siguiendo un criterio de optimización, no tendrían fundamentos para la actuación.
- Existen otras situaciones en las que, pudiendo ser posible recoger suficiente información para hacer una elección informada, los costes de oportunidad *podrían ser* (en lugar de *serían*) muy elevados. Cuando no se saben los costes que supondría recoger más información (y no se sabe tampoco si valdría la pena averiguarlo) nos encontramos con una forma radical de indeterminación.
- Existen otros casos en los que, pudiendo ser posible recoger suficiente información para hacer una elección informada, los costes de oportunidad *serían* muy elevados (conocemos *cómo de elevados serían*). Cuando se sabe con certeza que los costes esperados de recoger más información superarán a los beneficios, la elección sólo puede hacerse en términos de una intuición, o una corazonada, es decir, dejando el mando del control decisional a las emociones.

Observamos, tras el análisis expuesto, que **existen numerosos casos donde el criterio racional no es capaz de decir nada**. Además, a veces nos encontramos con otros casos donde surgen problemas de índole práctica, en los cuales aunque se *podría* determinar la opción más adecuada, dado un criterio de racionalidad, los costes de recogida de información *serían* (en lugar de *podrían ser*) muy elevados (conocemos con certeza *cómo de elevados serían*). Son contextos donde **los costes esperados para tomar una decisión racional superan a los beneficios**: ser *racional* en estos casos *no lo es tanto*.

Por tanto, aunque tradicionalmente las emociones se han asociado con comportamientos irracionales y con acciones en cuya planificación falta juicio —de ahí el tópico de la frialdad en la decisión como criterio de calidad para la actividad mental—, hemos mostrado cómo existen situaciones cotidianas, en las que la elección racional no puede llevarse a la práctica, sólo puede hacerse en términos de una intuición, o una corazonada, es decir, dejando el mando del control decisional a las emociones.

Las emociones, en estos casos, no parece que vengan a interrumpirnos, sino a proporcionarnos un mecanismo alternativo de decisión.

B.2. Inteligencia de los mecanismos emocionales

A partir de los resultados de los trabajos de neurólogos y neurofisiólogos [apéndice A] hemos intentado formalizar algunas cuestiones acerca de la naturaleza de las emociones, sobre cómo operan en el cerebro y por qué tienen tanta influencia. Algunos resultados están en sintonía

con propuestas clásicas, pero hemos obtenido otros resultados que revelan aspectos aspectos contraintuitivos y difíciles de integrar con la psicología popular.

Según los trabajos de la Neurofisiología moderna, debemos concluir que:

- Los términos del lenguaje con los que identificamos las emociones no representan nada que suceda en la mente o el cerebro, son tan sólo una etiqueta, una manera de referirse a aspectos, mecanismos y procesos del cerebro y la mente.
- Las emociones son un tipo de funciones evolutivas. Organizan neurológicamente ciertos mecanismos de conducta emocionales como los mecanismos subyacentes en la respuestas del miedo, de manera muy similar en todas las especies, para satisfacer condiciones de supervivencia. Estos mecanismos cerebrales que generan conductas emocionales, y su significado evolutivo, se han conservado intactos desde casi su origen.
- El carácter consciente de las emociones no es la regla sino la excepción. Las emociones conscientes, obviamente, ocurrirán en seres que tengan la capacidad de tener conciencia de sí mismos. Cuando esto no es así, el cerebro cumple su función sin tener conocimiento consciente. El modo que tenemos de identificar emociones se remite al sentimiento consciente que provoca, sin embargo esto no permite centrar adecuadamente el estudio de las emociones desde una perspectiva científica. Esta afirmación es contraintuitiva a priori, ya que tenemos convicciones arraigadas para creer que las emociones son sentimientos conscientes.

El planteamiento que seguiremos en este trabajo, no es el de explicar las emociones a partir de su registro consciente, ni tampoco a través de sus respuestas fisiológicas (temblores, palpitations o sudores). Lo que nos interesa es identificar y reproducir en el diseño de un artefacto inteligente, los mecanismos que detectan las situaciones que provocan las emociones. Las funciones emocionales nos permiten resolver problemas que nuestra capacidad deliberativa es incapaz de abarcar. Estamos interesados en estudiar las soluciones biológicas generadas por la naturaleza, las emociones, a la hora de resolver problemas a los que no se ha encontrado solución desde la inteligencia artificial. Esto no quiere decir que al abordar las emociones desde la inteligencia artificial pretendamos programar ordenadores que vivan experiencias emocionales. Significa, al contrario, que buscamos representar:

- **La inteligencia evolutiva de los mecanismos sobre los que se despliegan nuestros procesos cognitivos**, ya que ahí es donde se encuentran las emociones.
- **La funcionalidad de las emociones como complemento a la deliberación racional en procesos de toma de decisiones.**

En general, pretendemos extender el modelo de procesamiento de información de la mente para que explique algo más de nuestra complejidad que la que ofrecen los modelos clásicos sobre nuestros procesos mentales.

Apéndice C

Metodología: construir emociones

C.1. El enfoque sintético: “entender construyendo”

El enfoque propuesto, denominado constructivo o sintético, se presenta como una metodología a través de la cual las ciencias de lo artificial (inteligencia artificial, robótica y ciencias cognitivas) pueden contribuir a la investigación científica sobre la vida y la cognición. Cada vez más, estas ciencias manifiestan que son capaces de ir más allá de un mero enfoque de ingeniería, y proporcionar un enfoque puramente científico a la hora de tratar temas cruciales de las “ciencias naturales” (como la biología, la psicología y las ciencias cognitivas). La propuesta consiste en ofrecer a estas ciencias una nueva metodología capaz de encontrar los mecanismos ocultos en los sistemas naturales.

Esta metodología, cuyo objetivo podría resumirse en “entender construyendo”, da la vuelta al orden habitual entre el análisis del comportamiento y la construcción de modelos. Hace necesario que el investigador integre en primer lugar una hipótesis sobre la cognición en sistemas artificiales para después poder examinar los comportamientos que se producen. El objetivo de esto es proporcionar a la ciencia una nueva forma de realizar explicaciones científicas simples de sistemas naturales complejos [Damiano y Cañamero, 2010].

Esta propuesta surge en el marco del estudio del comportamiento de sistemas complejos, en los que se muestra que la aparente complejidad de este comportamiento surge como resultado de la interacción de elementos simples interaccionando con entornos complejos. Estos comportamiento de sistemas complejos cumplen las siguientes propiedades:

- Aunque se basan en la interacción de elementos simples, no pueden ser explicados tomando estos elementos por separado.
- Surgen por la interacción entre los elementos, el sistema global constituido por ellos y el ambiente con el que el sistema interacciona.
- Como resultado de esta interacción, tienden a sobrepasar las capacidades del observador de cálculo y previsión, incluso cuando es quien ha construido el sistema que los manifiesta.

Estas hipótesis nos permiten considerar a los modelos artificiales de sistemas cognitivos como generadores de conocimiento. Ya que, una vez que situamos nuestros modelos en un entorno, pueden expresar comportamientos inesperados para quienes los construyeron, permitiendo revisar las hipótesis sobre la cognición que se habían integrado en el modelo. En este sentido los modelos creados nos sirven como guías para el descubrimiento, ya que en una investigación científica nos permiten hacer predicciones sobre experimentos futuros, extender el alcance de la investigación a nuevos fenómenos, resolver problemas conceptuales en su planteamiento, proporcionar nuevos puntos de vista para descubrimientos empíricos, etc. [Barandiaran y Chemero, 2009].

La perspectiva presentada desobedece los viejos postulados mecanicistas acerca de que “el todo es la suma de sus partes” y que es posible conocer el comportamiento pasado, presente y futuro del objeto de estudio, así como presenta una forma de describir la naturaleza diferente del viejo ideal de construir un reflejo objetivo de ella. El “entender construyendo” propone una práctica científica como un acto premeditado de construcción. Exige al investigador poner en práctica el pensamiento científico como una forma de conocimiento que crea activamente, y no refleja pasivamente, el fenómeno explorado [Damiano y Cañamero, 2010].

C.2. Artificial Intelligence-Inspired Biology: Animats

Es evidente que la inteligencia artificial ha estado inspirada por la biología desde sus inicios hasta el presente. Sin embargo, esta inspiración también puede tener lugar en el sentido contrario. Bajo lo que se ha denominado Artificial Intelligence-Inspired Biology (AIIB) o Artificial Intelligence-Inspired Cognitive Science (AIICS), se propone que la inteligencia artificial puede servir como inspiración a la biología, ampliando nuestros conocimientos acerca de la cognición natural/biológica. Se plantea que los conceptos, problemas, teorías y técnicas de la IA a la hora de diseñar máquinas que imitan el comportamiento de humanos u otros animales pueden motivar nuevas investigaciones que nos ayuden a entender cuáles son los logros de la cognición, identificando el papel de aspectos relevantes del entorno, explicando cómo evolucionaron o se desarrollaron las capacidades de los individuos, y proponiendo nuevos mecanismos o arquitecturas de procesado de información, que expliquen mejor los tipos de cognición natural. [Sloman, 2010]:

- La forma más obvia en que la IA puede ser relevante para la biología es proporcionando modelos que funcionen como explicaciones potenciales de las capacidades cognitivas de los organismos. Esto incluye modelos sobre el funcionamiento de las diferentes formas de percepción, aprendizaje, generación y comprobación de hipótesis, etc.
- Pero además existe un rol menos evidente en la AIIB, que consiste en generar preguntas de investigación, y conceptos que pueden ser usados para formular esas preguntas y sus respuestas. Por ejemplo, estas preguntas pueden estar relacionadas con cuántos sistemas son necesarios para una tarea, qué tipo de cosas hacen, cómo son controlados, cómo se desarrollan, qué tipo de información adquieren y manipulan, cómo evolucionaron, etc.

Dentro del marco descrito anteriormente, según el cual estamos interesados en crear sistemas complejos que posean ciertas propiedades de emergencia que les permitan representar sistemas cognitivos biológicos, la aportación de la AIIB puede proporcionar es la siguiente: si identificamos un modelo formal o de inteligencia artificial que posea dichas propiedades de emergencia en las cuales estamos interesados, investigar dicho modelo puede sernos de utilidad a la hora de entender sistemas naturales que exhiban dichas propiedades. Si el enfoque sintético nos proporciona una metodología para explicar la cognición natural a partir de sistemas biológicos, la AIIB nos proporciona un punto de partida en forma de modelos en los que basarnos.

En [Barandiaran y Chemero, 2009] se clasifican los modelos de vida artificial (Figura C.1) según la posición que estos ocupan entre la teoría y los datos empíricos: (1) Los más cercanos al dominio empírico son los modelos mecanicistas, en los cuales existe una representación casi exacta entre las variables del modelo y las que se pueden observar en el sistema estudiado y su entorno; (2) los modelos funcionales, por su parte, buscan una correspondencia funcional o de comportamiento entre sus modelos y el sistema estudiado, en lugar de una correspondencia entre variables concretas; (3) los modelos genéricos, cubren un espectro amplio de fenómenos, con el objetivo de buscar principios generales de sistemas complejos; y finalmente (4), los modelos conceptuales no tienen como objetivo ningún sistema concreto, ni un amplio espectro de ellos. Por el contrario, los modelos conceptuales se construyen a partir de teorías, de las cuales ilustran conceptos, simulan principios teóricos, etc.

El trabajo presentado esta basado en este último modelo, siguiendo la recomendación hecha por Dennett¹ de crear modelos conceptuales de agentes cognitivos que no se corresponden con ningún animal real. Estos agentes cognitivos de animales inventados que pretenden proporcionar conocimiento sobre asuntos sobre biología o ciencias cognitivas son conocidos como “animats”. Los animats no son abstracciones para representar conjuntos específicos de datos biológicos, sino que representan teorías biológicas, y los modelos compuestos por animats no intentan capturar aspectos específicos del mundo, sino que constituyen mundos virtuales. Detrás de todo esto está la idea de que la biología como ciencia no sólo trata sobre animales reales concretos, sino que también (y quizás esto es más importante) sobre principios generales de organización, desarrollo y evolución biológica [Barandiaran y Chemero, 2009].

¹“one does not want to get bogged down with technical problems in modeling the cognitive eccentricities of turtles if the point of the exercise is to uncover very general, very abstract principles... So why not then make up a whole cognitive creature, a Martian three-wheeled iguana say, and an environmental niche to cope with” [Dennett, 1978].

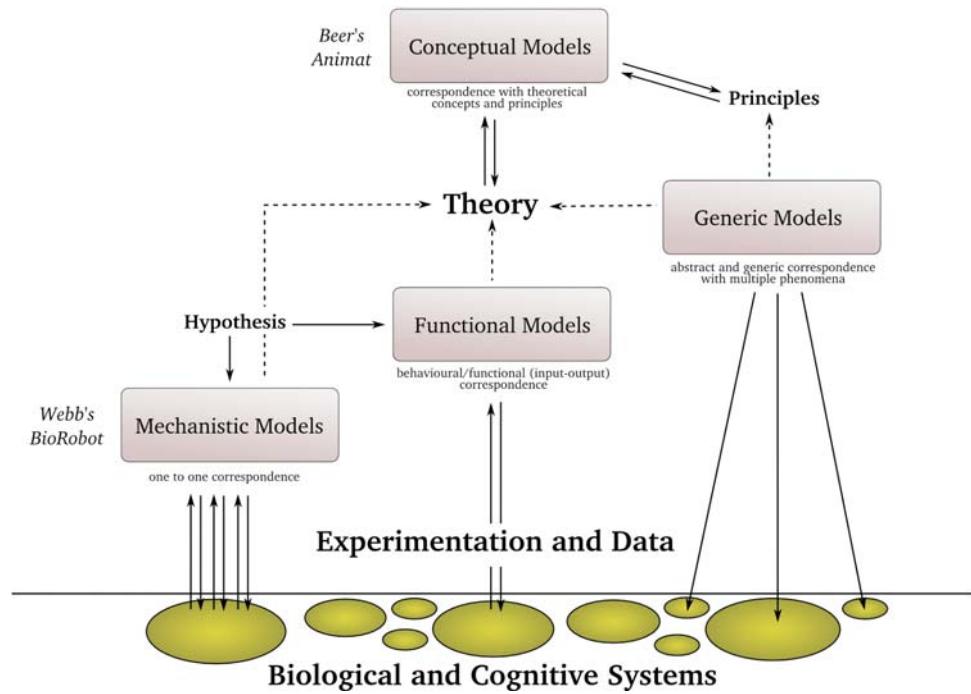


Figura C.1: Tipos de modelos. Copyright (c) 2009 Xabier E. Barandiaran under a Creative Commons Attribution Share-Alike licence, freedom to copy, modify and distribution provided that this notice is preserved.

Apéndice D

Herramientas: modelos basados en agentes

La utilización de modelos de simulación basados en agentes (ABMs) por parte de la comunidad científica está creciendo de manera acelerada. Las plataformas ABM más utilizadas, están definidas como un conjunto de funciones y librerías que posibilitan la creación, ejecución y seguimiento de las simulaciones. En este apéndice se hace una descripción de las diferentes criterios por los que se eligió NetLogo como herramienta para implementar los modelos de agentes cognitivos propuestos.

D.1. Comparación de diferentes plataformas de ABMs

Las plataformas analizadas son las siguientes:

Código abierto

- Swarm
- MASON
- Repast

Freeware y Shareware

- Starlogo
- NetLogo
- Obeus

Software propietario

- AgentSheets
- AnyLogic

A continuación se describe cada una de las plataformas:

Swarm Es un software de código abierto para modelado y simulación de sistemas, diseñado específicamente para el desarrollo de simulaciones Multiagente de sistemas adaptativos complejos [Swarm, 2006]. [Najlis et al, 2001] determina que la curva de aprendizaje de Swarm es un factor importante a considerar antes de elegir esta herramienta, porque se requieren conocimientos de programación en el proceso de modelado.

MASON Multi Agent Simulation Of Neighbourhood. Fue desarrollado por el Laboratorio de Computacion evolutiva de la Universidad de George Mason. No tiene incorporado la funcionalidad para realizar gráficas dinámicas (histogramas, gráficos) y el grupo de trabajo es reducido en comparación a otras herramientas.

REPAST Recursive Porous Agent Simulation Toolkit. Originalmente fue desarrollado por la Universidad de Chicago, y en la actualidad las versiones recientes han sido distribuidas por el Laboratorio Nacional de Argonne y administrado por la Organización Repast para la arquitectura y Desarrollo. Repast permite la implementación de los modelos en tres lenguajes de programación: Python (RepastPy); Java (RepastJ); and Microsoft.Net (Repast.Net). RepastPy posibilita el desarrollo de modelos básicos de una manera sencilla a través de una interfaz gráfica [Collier and North, 2005]; además RepastPy tiene habilitada la opción de exportar modelos a JAVA. Repast.Net and RepastJ pueden modelar sistemas de mayor complejidad con respecto a RepastPy [Vos, 2005].

STARLOGO Fue desarrollado por el Media Lab del Instituto de Massachusetts (MIT). Al igual que en NetLogo, los modelos son programados de forma procedimental, lo cual que no propicia el tipo de procesos de abstracción de la programación orientada a objetos. StarLogo carece de la posibilidad de complementar sus funcionalidades a través de desarrollos propios en contraposición a las ventajas de los sistemas de código abierto, sin embargo, en 2006 salió su versión de código abierto: OpenStarLogo. StarLogo es fácil de usar y no exige un conocimiento previo de programación a nivel de avanzado. Tiene incorporado elementos gráficos que pueden ser configurados con el resultado de los modelos; además tiene un gran número de modelos de ejemplo disponibles.

NETLOGO Es una variante de StarLogo, originalmente fue desarrollado por el Centro de Aprendizaje y Modelamiento en la Northwestern University de Chigaco. NetLogo y StarLogo tienen la opción de importar archivos gráficos que pueden utilizarse para definir los entornos en donde los agentes interactúan, y sociales, con una documentación extensa, tutoriales, modelos de prueba, todos estos recursos disponibles desde el sitio web.

OBEUS (Entorno Basado en Objetos para Simulaciones Urbanas) Fue desarrollado de Autómatas Geográficos [GAS; Benenson and Torrens, 2004]. Fue diseñado geoespacial. Su implementación se realizó utilizando el Framework de .NET. OBEUS tiene un entorno gráfico para desarrollar el modelo, el comportamiento del agente y las reglas de interacción deben ser programadas utilizando los lenguajes de la plataforma .NET (C, J, Visual Basic) lo que hace imprescindible tener conocimientos básicos de programación.

AgentSheets Es un sistema que permite desarrollar modelos basados en agentes sin necesidad de un conocimiento alto de programación. Los modelos son desarrollados a través de una interfaz gráfica [Repennig et al, 2000]. En el sitio web se pueden descargar algunas demostraciones de los modelos. El sistema es fácil de usar, brinda la posibilidad de realizar modelos sencillos, pero sin embargo los modelos son limitados (como la complejidad del comportamiento de los agentes y su interacción) [Carvalho, 2000]. El Software carece de funcionalidad para representar datos gráficos obtenidos de la simulación.

Anylogic Tiene como una de sus principales características la posibilidad de leer y escribir datos en una hoja de cálculo o base de datos durante el proceso de simulación. Programas externos pueden ser lanzados desde un modelo en AnyLogic, lo que favorece el intercambio dinámico de datos. Los modelos sólo pueden ser creados utilizando un sistema operativo Windows, y se pueden ejecutar en otros sistemas operativos una vez se hayan compilado y se tenga tutoriales y ejemplos no está disponible.

Algunos Criterios Adicionales para Elegir NetLogo como la plataforma de las simulaciones

- Comparado con otras herramientas, como Swarm, Mason, Repast u Obeus, NetLogo es fácil de aprender, con un lenguaje de programación simple y eficaz, interfaz gráfica y mucha documentación y ejemplos. Es el lenguaje de más alto nivel de los presentados, recomendable para modelos en los que las interacciones sean locales y a corto plazo, y que se puedan modelar en un entorno de tipo malla. [Railsback et al, 2006]
- Evaluando las características de OBEUS, AgentSheets y AnyLogic no se encuentran mejores ventajas de las ofrecidas por NetLogo, ni en la curva de aprendizaje ni en las posibilidades de modelado y desarrollo de sistemas complejos.

- NetLogo separa la capa de implementación de la capa de presentación.
- Modelar el movimiento con NetLogo es muy fácil, debido a que simplemente se actualizan las coordenadas X, Y en cada instante, mientras que en otras plataformas hay que recurrir a funciones de más bajo nivel para realizar primero el desplazamiento y después informar de la posición.

Características Adicionales de NETLOGO

NetLogo [Tisue y Wilensky, 2004] es un ambiente de modelado programable para agentes que sirve para simular fenómenos naturales y sociales. Este proyecto fue desarrollado en el Massachusetts Institute of Technology (M.I.T.) en 1999 y su autor fue Uri Wilensky; estando en la actualidad en continuo desarrollo.

Esta herramienta está adaptada para modelar sistemas complejos que se modifican constantemente, de forma que los usuarios pueden dar instrucciones a miles de agentes que actúan concurrentemente y se relacionan entre sí. NetLogo es una herramienta simple y fácil de utilizar para realizar simulaciones. También posee una documentación de todo tipo, así como ejemplos prácticos de simulaciones desarrolladas en diferentes áreas científicas que pueden servir de bases para nuestras simulaciones. Esta herramienta está escrita bajo el lenguaje JAVA, lo que hace que pueda ser utilizadas en diferentes plataformas como son Windows, Linux etc., y además los modelos individuales pueden ejecutarse como Java applets dentro de un web browser y se pueden ver en ambientes 2D o 3D.

La forma de programar en NetLogo está basada en los conceptos de *turtles* (agentes que participan en la simulación) y *patches* (diferentes posiciones en el terreno). Mediante una gran cantidad de funciones predefinidas, es posible hacer interaccionar a los diferentes turtles y patches de diferentes maneras, ya que estos serán capaces de reconocerse entre sí, conocer su posición, etc.

NetLogo tiene como base los conceptos de turtles, que vienen siendo los agentes dentro de la simulación y también tiene una segunda clase de agente llamados patches, estos son como los turtles pero con la diferencia que están siempre alrededor y no se mueven.

Una de las principales ventajas de NetLogo es que el proceso de aprendizaje es muy rápido, debido a la sencillez del lenguaje de programación, la gran cantidad de documentación y ejemplos y que posee una interfaz gráfica intuitiva.

Apéndice E

Resolución del sistema de Markov

Denominamos cadena de markov, y representambos como $\vec{V}^* = \{X_t\}$, a una secuencia de estados más un conjunto de valores que determina las probabilidades de transición, $\{P_{ij}\}_{ij=1,\dots,n}$, donde $P_{ij} = P(X_t = X_i | X_{t+1} = X_j)$ que podemos expresar matricialmente como:

$$\Pi = [P_{ij}]_{ij=1,\dots,n} = \begin{bmatrix} P_{11} & - & P_{1n} \\ | & & | \\ P_{n1} & - & P_{nn} \end{bmatrix}$$

Los elementos de la matriz represenran la probabilidad de que el agente, si se encuentra en (X_i, t) tome la decisión de actuar en la dirección de alcanzar $(X_j, t + 1)$. Suponemos que no hay incertidumbre sobre las acciones, esto es, al ejecutar $a_{ij} : x_i \rightarrow x_j$ se cumple que $P(X_t = X_i | X_{t+1} = X_j) = 1$ y que la incertidumbre se restringe a las deficiencias de nuestra capacidad de predicción para hacer inferencias sobre la idoneidad del estado que alcanzamos. En estas circunstancias, podemos identificar las probabilidades de transición $\{P_{ij}\}_{ij=1,\dots,n}$ del sistema de Markov con las capacidades predictivas definidas en la sección anterior.

Definición 9 Sea un sistema cognitivo X caracterizado por su capacidad predictiva P_X . Sea un estado x_i y sea $a_{ij} : x_i \rightarrow x_j$ la mejor acción en ese estado, es decir, aquella que cumple $P(X_t = X_i | X_{t+1} = X_j) = 1$ o en términos de P^* aquella que $P^*(x_i, a_{ij}) = 1$ o más brevemente $P^*_{ij} = 1$ donde i, j son los índices que identifican los estados x_i, x_j y la acción a_{ij} .

En este contexto la capacidad predictiva del sistema X , que denotamos por P_X , para el mismo estado y acción, $P_X(x_i, a_{ij})$. y que necesariamente cumple $P_X(x_i, a_{ij}) < P^*(x_i, a_{ij})$ representa el valor $v \in [0, 1]$ que el sistema X asocia a la transición $x_i \rightarrow x_j$. En nuestro modelo las $\{P_{ij}\}_{ij=1,\dots,n}$ de Markov son las $P_X(x_i, a_{ij})$ que asocia el sistema a la transición $x_i \rightarrow x_j$.

E.1. Modelo $(D+E_1+E_2)$ markoviano

Supongamos, por sencillez, que la estructura de un sistema $D + E_1$ viene representada por $\frac{2}{3}D + \frac{1}{3}E_1$ (donde un tercio de las decisiones que toma el sistema serán desconocidas y consideradas como peligrosas por E_1), entonces el sistema de Markov del sistema X puede reducirse al estudio del vector \vec{V}^*

$$\vec{V} \equiv \vec{V}^* \bmod 3, \quad \text{donde } \vec{V} = (V_E, V_D, V_D)$$

Conocer el comportamiento de este sistema significa resolver su ecuación de evolución. Si denotamos

$$P_{ij} = P(X_t = x_i | X_{t+1} = x_j) = P(i, t | j, t + 1)$$

la probabilidad total de encontrarse en el estado X_j viene representada por

$$P_j(- \mid X_{t+1} = x_j) = \sum_{i=1}^n P_{ij} \cdot P(i, t),$$

$$\text{o más brevemente } P_j(t+1) = \sum_{i=1}^n P_{ij} \cdot P_i(t)$$

En general $\forall j$, hay que resolver el sistema

$$\vec{V}(t+1) = \Pi \cdot \vec{V}(t)$$

donde

$$\Pi_{DE_1} = [P_{ij}]_{ij=1,2,3} = \begin{bmatrix} 0 & 1 - P_D & P_D \\ P_{E_1} & 0 & 1 - P_D \\ 1 - P_{E_1} & P_D & 0 \end{bmatrix}$$

Resolver, por tanto, resolver la ecuación de la dinámica del sistema de Markov –obtener $\vec{V}(t)$ estacionario denotado \vec{V}^{st} – consiste en resolver un problema clásico de autovalores y autovectores.

Definición 10 Decimos que \vec{V}^{st} es la solución del sistema si \vec{V}^{st} es autovector de la matriz de transiciones Π , es decir, si existe un vector $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ tal que $\Pi \cdot \vec{V}^{st} = \lambda \cdot \vec{V}^{st}$. Como es sabido, el vector $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ se obtiene resolviendo $|\Pi - \lambda \cdot Id| = 0$, y el vector $\vec{V}^{st} = \{V_1^{st}, V_2^{st}, \dots, V_n^{st}\}$ el sistema de ecuaciones $(\Pi - \lambda_i \cdot Id)V_i^{st} = 0$, $\forall i = 1, \dots, n$.

Example 11 Resolvemos, como ejemplo, el caso del sistema $D + E_1 = \frac{2}{3}D + \frac{1}{3}E_1$. a partir de un vector inicial $\vec{V} = (V_{E_1}, V_D, V_D)$ y la matriz Π_{DE_1} .

$$\vec{V}^{st} = \Pi_{DE_1} \cdot \vec{V}^{st} \Leftrightarrow \vec{V}^{st} = (V_{E_1}^{st}, V_D^{st}, V_D^{st}) = \frac{1}{\mu} \begin{bmatrix} 1 - P_D + P_D^2 \\ 1 - P_D + P_D \cdot P_{E_1} \\ 1 - P_{E_1} + P_D \cdot P_{E_1} \end{bmatrix}$$

donde $\mu = 3 - P_D - 2P_{E_1} + 2P_D \cdot P_{E_1} + P_{E_1}^2$ para garantizar que $\sum_{i=1}^n V_i^{st} = 1$.

Una vez conocidas las componentes del vector \vec{V}^{st} tenemos todo lo necesario para identificar el comportamiento del sistema. Si nos preguntamos cuál es la probabilidad de que el sistema recorra la secuencia $e = \{X_{t-1}, X_t, X_{t+1}\}$ a partir de su capacidad deliberativa y su estructura de emociones primarias, obtenemos que:

$$P_{\{X_{t-1}, X_t, X_{t+1}\}} = \sum_{i=1}^n P_i \cdot V_i^{st} = P_{E_1} \cdot V_{E_1}^{st} + P_D \cdot V_D^{st} + P_D \cdot V_D^{st}$$

Conclusion 12 Podemos ya presentar el primero de los atributos importantes de un sistema $D + E_1 = \frac{2}{3}D + \frac{1}{3}E_1$ tal como se ha formulado: el efecto de la contribución de E_1 sobre el sistema global es mayor que su contribución a la estructura (1/3 en el ejemplo). Veámoslo.

Si entendemos que la forma en que afecta E_1 a la capacidad predictiva P_D es lineal (algo que está implícito en la teoría de marcadores somáticos) la probabilidad de que el sistema recorra la secuencia $e = \{X_{t-1}, X_t, X_{t+1}\}$ se calcularía (donde $\vec{V}^{(T)}$ es el vector traspuesto de \vec{V}).

$$P_{\{X_{t-1}, X_t, X_{t+1}\}} = (P_{E_1}, P_D, P_D) \cdot \vec{V}^{(T)} = (P_{E_1}, P_D, P_D) \cdot \begin{pmatrix} V_{E_1} \\ V_D \\ V_D \end{pmatrix} = \frac{1}{3}P_{E_1} + \frac{2}{3}P_D$$

puesto que $\vec{V} = (V_{E_1}, V_D, V_D) = (1/3, 1/3, 1/3) \cdot \vec{u}$, donde \vec{u} es un vector unitario y se eligen coeficiente para que $\sum_{i=1}^3 V_i = 1$.

Pero aceptar que la estructura $D + E_1$ debe ser acoplada nos lleva a un resultado distinto. Es necesario calcular el vector \vec{V}^{st} para obtener:

$$P_{\{X_{t-1}, X_t, X_{t+1}\}} = (P_{E_1}, P_D, P_D) \cdot \vec{V}^{st(T)} = (P_{E_1}, P_D, P_D) \cdot \begin{pmatrix} V_{E_1}^{st} \\ V_D^{st} \\ V_D^{st} \end{pmatrix} =$$

$$= \frac{1}{3-P_{E_1}-2P_D+2P_D\cdot P_{E_1}+P_D^2} \cdot (P_{E_1}, P_D, P_D) \cdot \begin{bmatrix} 1-P_D+P_D^2 \\ 1-P_D+P_D\cdot P_{E_1} \\ 1-P_{E_1}+P_D\cdot P_{E_1} \end{bmatrix} =$$

$$\frac{P_{E_1}+2P_D-P_D^2-2P_D\cdot P_{E_1}+3P_D^2\cdot P_{E_1}}{3-P_{E_1}-2P_D+2P_D\cdot P_{E_1}+P_D^2}$$

Veamos un caso numérico. Si por ejemplo, $P_{E_1} = 0,1$ y $P_D = 0,8$ obtenemos que $\vec{V}^{st} = (0,4, 0,13, 0,47)$ que contrasta con los valores de $\vec{V} = (0,33, 0,33, 0,33)$.

Esto nos permite cuantificar el acoplamiento: el verdadero efecto de E_1 es que rompe el equilibrio entre las componentes del sistema total redistribuyendo las probabilidades. El acoplamiento hace que la capacidad deliberativa pase, en el ejemplo, de “pesar” un 0,66 a un 0,60.

El efecto refleja algo que es obvio: el bajo poder predictivo de E_1 que nos lleva a tomar la decisión de no avanzar en ciertos estados, realimenta el uso de los mecanismos de E_1 puesto que el sistema se conserva en el estado de partida para el que E_1 no actúa.

Sin embargo lo que nos interesa es modelar la inclusión de emociones secundarias E_2 y los efectos de acoplamiento que se generan. Para ello supondremos dos cadenas de Markov:(1) modelando un sistema $D+E_1$ y (2) modelando un sistema E_2 . Construiremos el efecto de incorporar la cadena E_2 a la inicial.

que denotaremos respectivamente $\vec{V}(t+1) = \Pi_{DE_1} \cdot \vec{V}(t)$, $\vec{W}(t+1) = \Pi_{E_2} \cdot \vec{W}(t)$, y el

sistema acoplado $\vec{U}(t+1) = \Pi_{DE_1E_2} \cdot \vec{U}(t)$, de modo que la dinámica del sistema DE_1E_2 venga establecida por la matriz $\Pi_{DE_1E_2} = \frac{1}{2}(\Pi_{DE_1} + \Pi_{E_2})$ que se obtiene fácilmente de:

$$\Pi_{DE_1} = \begin{bmatrix} 0 & 1-P_D & P_D \\ P_{E_1} & 0 & 1-P_D \\ 1-P_{E_1} & P_D & 0 \end{bmatrix}, \quad \Pi_{E_2} = \begin{bmatrix} 0 & 1-P_{E_2} & P_{E_2} \\ P_{E_2} & 0 & 1-P_{E_2} \\ 1-P_{E_2} & P_{E_2} & 0 \end{bmatrix}$$

$$\Pi_{DE_1E_2} = \begin{bmatrix} 0 & 1-\frac{1}{2}(P_D+P_{E_2}) & \frac{1}{2}(P_D+P_{E_2}) \\ \frac{1}{2}(P_{E_1}+P_{E_2}) & 0 & 1-\frac{1}{2}(P_D+P_{E_2}) \\ 1-\frac{1}{2}(P_{E_1}+P_{E_2}) & \frac{1}{2}(P_D+P_{E_2}) & 0 \end{bmatrix}$$

el resultado anterior nos permite conocer el \vec{V}^{st} en este caso simplemente cambiando $P_D \Rightarrow \frac{1}{2}(P_D+P_{E_2})$ y $P_{E_1} \Rightarrow \frac{1}{2}(P_{E_1}+P_{E_2})$. Las conclusiones son equivalentes: el efecto de los elementos en el sistema global no es lineal con respecto a la contribución que hacen a la estructura.

E.2. Efectos constraintuitivos: paradojas markovianas

Hemos mostrado cómo en los sistemas de Markov no se cumple la propiedad aritmética. Esto es, $P_{DE_1E_2} \neq P_D + P_{E_1} + P_{E_2}$. A continuación vamos a describir un efecto denominado “losing strategies cooperate to win” aplicado a nuestro sistema y a examinar sus consecuencias. Este efecto, mostrado en diferentes trabajos de economía [Boman et al, 2008], física [Buceta et al, 2002] y teoría de juegos [Harmer y Abbott, 1999], surge en condiciones especiales en sistemas donde se solapan fenómenos aleatorios. En el marco mostrado podemos demostrar que:

Proposición 13 *Existen condiciones en las que la incorporación de una estructura de emociones E_2 errónea (esto es, $P_{E_2} < 1/2$) a un sistema X con una capacidad predictiva P_{DE_1} puede hacer que mejore su capacidad predictiva global, esto es, $P_{DE_1E_2} > P_{DE_1}$.*

La importancia de este resultado es que constituye un contraejemplo al modelo de marcadores somáticos de Damasio: en ocasiones en las que esté dañada la estructura E_2 en un sistema cognitivo, su comportamiento podría mejorar. Si esto es cierto, la función de los marcadores somáticos no residiría, al menos completamente, en el valor predictivo de las experiencias registradas. Pasemos a describir las condiciones en las que este fenómeno no-lineal, producto del acoplamiento, se manifiesta.

Demostración: El fenómeno citado se produciría si existe solución al sistema siguiente.

$$\left\{ \begin{array}{l} P_D \cdot [V_1^{st} + V_2^{st}] + P_{E_1} \cdot [V_3^{st}] < 1/2 \\ P_{E_2} \cdot [V_1^{st} + V_2^{st} + V_3^{st}] < 1/2 \\ \frac{1}{2}(P_D + P_{E_2}) \cdot [V_1^{st} + V_2^{st}] + \frac{1}{2}(P_{E_1} + P_{E_2}) \cdot [V_3^{st}] > 1/2 \end{array} \right\}$$

donde la primera ecuación representa un sistema $D+E_1$ con baja capacidad de predicción, la segunda ecuación expresa que se registran emociones E_2 incorrectas (capacidad de inferencia menor que 0.5), y la última un sistema $D+E_1+E_2$ que mejora al incorporar registros E_2 equivocados.

El sistema tiene solución, esto es, existen ciertas combinaciones de parámetros P_D, P_{E_1}, P_{E_2} que hacen compatibles las tres ecuaciones. Estos vienen fijados por las siguientes ecuaciones:

$$\left\{ \begin{array}{l} (1 - P_{E_1}) \cdot (1 - P_D)^2 > P_{E_1} \cdot P_D^2 \\ P_{E_2} < 1 - P_{E_2} \\ [1 - \frac{1}{2}(P_{E_1} + P_{E_2})] \cdot [1 - \frac{1}{2}(P_D + P_{E_2})]^2 < \frac{1}{2}(P_{E_1} + P_{E_2}) \cdot [\frac{1}{2}(P_D + P_{E_2})]^2 \end{array} \right\}$$

Apéndice F

Caracterización de modelo de serendipia en cadenas de Markov en una y dos dimensiones

Una vez que hemos mostrado cómo tiene lugar el fenómeno de serendipia en cadenas de Markov, consideramos que es interesante llevar a cabo un análisis más detallado del comportamiento de este marco matemático. Así, en este apéndice, analizamos algunos aspectos de su comportamiento, y estudiamos si el modelo sigue siendo válido para un número mayor de dimensiones, en el que ya no tengamos cadenas lineales, sino mallas de estados.

Caracterización del modelo de Markov

Observando la función $P_{DE_1E_2}$ respecto a P_{E_2} (Figura F.1.a) con detalle, podemos observar que no es totalmente recta, sino que presenta cierta curvatura. Para poder analizar cuál es el comportamiento de la función, es preciso observar su derivada respecto a P_{E_2} , obteniendo la función representada en la Figura F.2.

De esta función deducimos que $P_{DE_1E_2}$ crece más rápidamente para valores de P_{E_2} mayores. Esto significa que cuando mejoramos la capacidad predictiva de las emociones secundarias, el efecto sobre el resultado obtenido es no es lineal, y que **a medida que mejoramos la “calidad” de los marcadores somáticos, la mejora que obtenemos sobre el resultado final es no lineal y cada vez de mayor magnitud**.

Para explorar la relación entre la calidad del sistema deliberativo y del sistema emocional secundario, hemos calculado, para un P_{E_1} fijo igual a 0,1 y distintos valores de P_D y P_{E_2} , el punto en el que el sistema $(D+E_1+E_2)$ comienza a ser mejor que el sistema $(D+E_1)$, esto es, los límites de la región de acoplamiento positivo (Figura F.3). Observamos de nuevo que el resultado es no lineal, que cuando la capacidad deliberativa es suficientemente alta ($P_D \gtrsim 0,8$), las mejoras en su capacidad predictiva van a tener muy poco efecto a la hora de mejorar con respecto a un sistema $(D+E_1+E_2)$ con una capacidad predictiva P_{E_2} en torno a 0,5. En otras palabras, **llega un momento en que un sistema $(D+E_1)$, por mucho que mejore su capacidad deliberativa, apenas va a mejorar respecto a un sistema $(D+E_1+E_2)$** .

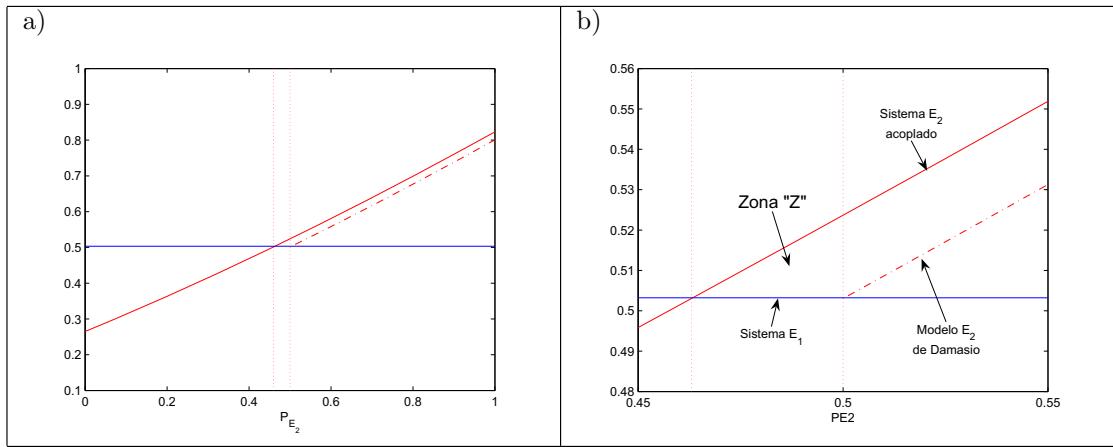


Figura F.1: (a) $P_{DE_1E_2}$: Probabilidad efectiva del sistema ($D+E_1+E_2$). (b) Zoom de la gráfica anterior en la que se muestran los resultados del modelo acoplado ($P_{DE_1E_2}$) frente a los del sistema ($D+E_1$) y los resultados que deberían darse según la el modelo de Damasio. Se observa también una zona "Z" de acoplamiento positivo, en la que introducir marcadores somáticos erróneos nos proporciona mejores resultados.

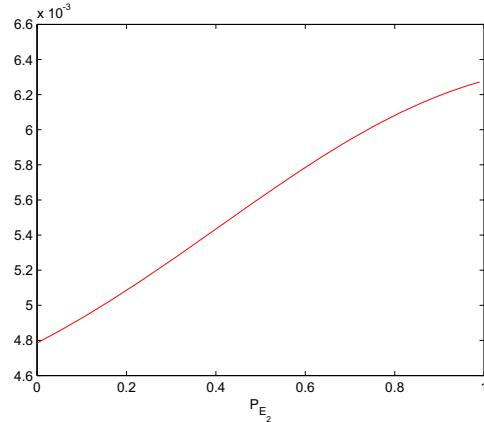


Figura F.2: Derivada de la función de probabilidad efectiva: $\frac{\partial(P_{DE_1E_2})}{\partial(P_{E_2})}$.

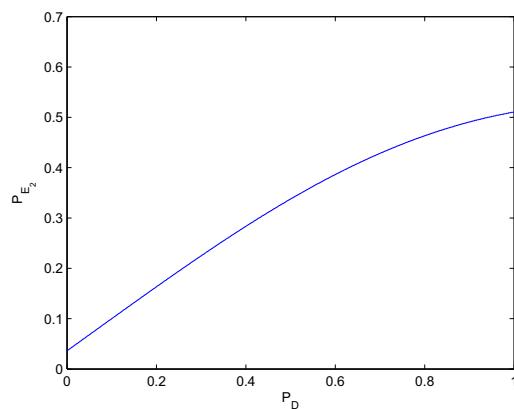


Figura F.3: Punto de acoplamiento positivo para un sistema con $P_{E_1} = 0, 1$

Extensión a dos dimensiones

Una vez que hemos caracterizado el efecto desatascador de las emociones secundarias en cadenas de Markov unidimensionales, es interesante comprobar si el efecto persiste cuando tenemos modelos de Markov con un número de dimensiones mayor. Para ello, no hemos resuelto el sistema teórico, sino que mediante simulaciones hemos comprobado que se obtienen resultados similares, encontrando de nuevo la existencia de una zona de acoplamiento en la que emociones secundarias erróneas hacen que el sistema mejore. Si promediando 50.000 simulaciones del sistema de Markov en dos dimensiones calculamos la probabilidad $P_{DE_1E_2}$ para $P_{E_1} = 0,1$ y $P_D = 0,8$, obtenemos el resultado de la Figura F.4, que es un resultado ligeramente mejor que el que veíamos en una dimensión, ya que la zona de acoplamiento positivo aparece antes.

Y si calculamos la zona el límite de la zona de acoplamiento positivo calculando realizando 10.000 simulaciones para cada valor de P_{E_2} y P_D , y calculando la tendencia de los resultados, obtenemos que (Figura F.5) cuando P_D es suficientemente alta, la zona en la que el acoplamiento del sistema E_2 más amplia que en una dimensión, y empieza para valores de P_{E_2} más pequeños (para valores de $P_D \gtrsim 0,6$ un agente emocional primario no obtiene ventajas significativas frente a un agente emocional secundario al mejorar su capacidad deliberativa). Según estos resultados podemos pensar que **aumentar la dimensionalidad no hace sino favorecer el efecto de acoplamiento positivo de emociones secundarias erróneas, permitiendo soportar mayor nivel de error.**

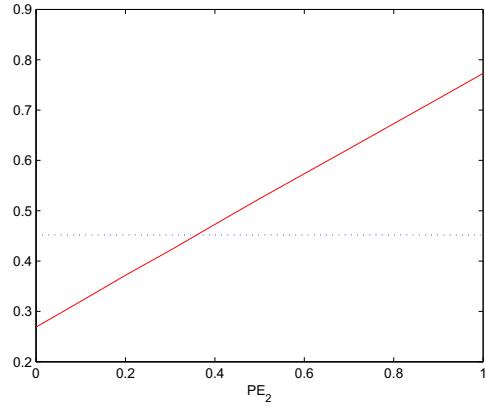


Figura F.4: $P_{DE_1E_2}$, Probabilidad efectiva del sistema ($D + E_1 + E_2$) de dos dimensiones. En linea discontinua el valor de P_{DE_1} en dos dimensiones. Se observa que es necesario un valor menor de P_{E_2} para obtener un acoplamiento positivo que en el resultado de una dimensión (Figura F.1).

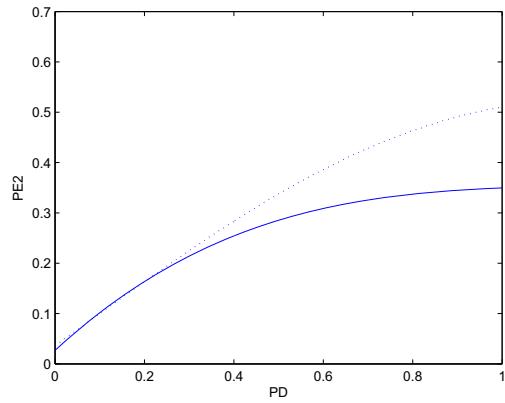


Figura F.5: Zona de acoplamiento positivo para un sistema de dos dimensiones con $P_{E_1} = 0.1$. En linea de puntos el resultado obtenido en el sistema de una dimensi n. Observamos que la zona de acoplamiento aparece para valores de P_{E_2} m s peque os.

Apéndice G

Resolución del algoritmo de Bellman

G.1. Modelo alternativo: sistema mínimo de toma de decisiones

Buscamos representar en un modelo mínimo las hipótesis descritas en el apartado anterior. Las asunciones de nuestro modelo serán muy básicas para que las conclusiones del modelo sean generales. Estos son los ingredientes:

G.1.1. Condiciones del problema

a. Soluciones ordenadas por ventanas temporales: Las soluciones que puede generar el sistema están clasificadas según el tamaño de la ventana temporal en la que son compatibles.

Si el conjunto de soluciones se denota $\{x_1, \dots, x_i, x_{i+1}, \dots, x_j, x_{j+1}, \dots, x_k, \dots\}$, se cumple que $\{x_1, \dots, x_i\}$ son compatibles con la ventana temporal Δt_1 , $\{x_1, \dots, x_i, x_{i+1}, \dots, x_j\}$ compatibles con Δt_2 , $\{x_1, \dots, x_i, x_{i+1}, \dots, x_j, x_{j+1}, \dots, x_k\}$ compatibles con $\Delta t_3, \dots$

b. Maximización del ajuste del sistema: Buscamos calcular el ajuste medio del sistema en un tiempo T . El ajuste medio será una medida de la capacidad media para encontrar soluciones a problemas del entorno. Llamaremos ajuste de la solución del sistema en un instante t , y denotaremos $a(t)$, a una medida de similitud entre la solución puesta en práctica, $x(t)$, y la ideal $x^*(t)$.

$$a(t) = \begin{cases} 1, & \text{si } x(t) = x^*(t) \\ 1 - \frac{|x(t) - x^*(t)|}{x(t)}, & \text{si } \frac{|x(t) - x^*(t)|}{x(t)} < 1 \\ 0, & \text{si } \frac{|x(t) - x^*(t)|}{x(t)} \geq 1 \end{cases}$$

c. Se conoce la relación funcional entre tiempo de generación de soluciones y ajuste. La función más general no es lineal (el esfuerzo en obtener soluciones muy específicas crece en términos relativos con el tiempo), que podemos suponer exponencial, $a(t) = S(1 - e^{-t/\tau})$.

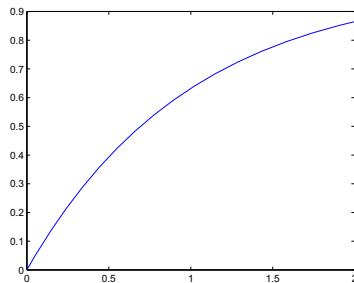


Figura G.1: Función generadora de solución.

d. Se conoce la relación funcional entre el tiempo de validez (ventana temporal) y la pérdida de ajuste de la solución: Suponemos igualmente una dependencia exponencial (podemos asumir gaussiana) para modelar la dependencia funcional entre el ajuste de la solución aplicada y el tiempo, esto es, $a(t) = S(e^{-t/\varepsilon})$, para una evolución exponencial.

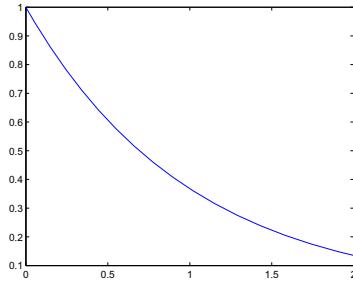


Figura G.2: Desajuste en la interacción con el mundo.

e. Comportamiento cualitativo del proceso de decisión. Con estos elementos, se modela el comportamiento más general de un sistema de decisión. Dado un entorno con una dinámica que permite generar soluciones en ventanas temporales de tamaño Δt , y cuyas soluciones en promedio duran Δt^* , la representación funcional del ajuste del sistema será de este modo.

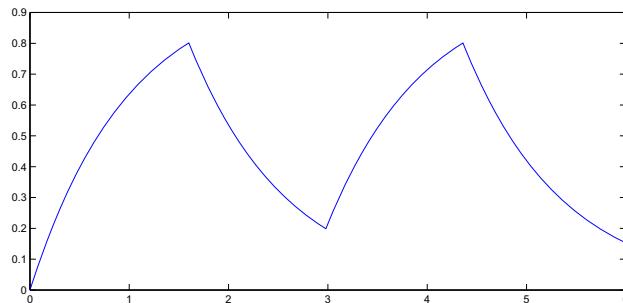


Figura G.3: Dependencia funcional inicial.

Llamaremos ajuste global de la solución de un sistema, en un periodo T , al promedio del ajuste,

$$a(T) = \frac{1}{T} \int_0^T a(t) dt$$

Denotamos como $\gamma(t) \in \{\gamma_0, \gamma_1\}$ la decisión binaria que, para cada instante de tiempo tiene el sistema (γ_0 : generar una solución; γ_1 : ejecutar una solución) y queremos que optimice su ajuste global. Resolvemos el problema a continuación.

Se supone que maximizar cada una de las etapas constituye el máximo del sistema. Para tener una medida que permita hacer comparaciones cuantitativas calculemos la ley de la dinámica del sistema y calculemos el valor de la solución ‘avariciosa’. Compararemos este resultado con el óptimo del sistema a partir del criterio de Bellman [Oviedo, 2005].

Solución estándar

Veamos qué ecuación resulta del modelo propuesto. Conocemos las soluciones y debemos obtener las ecuaciones de las que provienen:

- **Fase de generación:** $a(t) = a_M(1 - e^{-t/\tau})$

■ **Fase de ejecución:** $a(t) = a_M(e^{-t/\varepsilon})$

Para el caso $\gamma(t)=\gamma_0$, integrando en el intervalo $t \in \{t_0, t_1\}$ tenemos:

$$\frac{d}{dt}a(t) = \frac{1}{\tau}(a_M - a(t))$$

Para el caso $\gamma(t)=\gamma_1$, entre $t \in \{t_1, t_2\}$ se tiene:

$$\frac{d}{dt}a(t) = -\frac{1}{\varepsilon}a(t)$$

Ambas pueden combinarse, tomando como valores $\gamma_0 = 0$, $\gamma_1 = 1$. Sin pérdida de generalidad, se toma ($a_M = 1$), y se tiene la ecuación global del comportamiento:

$$\frac{d}{dt}a(t) = -\gamma(t) \cdot \frac{1}{\varepsilon}a(t) + (1 - \gamma(t)) \cdot \frac{1}{\tau}(1 - a(t))$$

Evaluamos el comportamiento en un intervalo $(0, T)$. Para calcular el valor promedio, simplemente integramos:

$$\bar{p}(T) = \frac{1}{T} \int_0^T \gamma(t) \cdot a(t) dt$$

G.1.2. Cálculo de política óptima

Una vez definido el problema, procedemos a calcular los valores que nos proporcionan un sistema que ofrezca un máximo valor de $\bar{p}(T)$. Si tenemos,

$$\begin{cases} \dot{a}(t) = \frac{1}{\tau}(1 - a(t)) - \gamma(t) \cdot (\frac{1}{\varepsilon} + a(t) \cdot (\frac{1}{\varepsilon} - \frac{1}{\tau})) \\ \dot{p}(t) = \gamma(t) \cdot a(t) \end{cases}$$

donde $\gamma(t) = \{0, 1\}$ y queremos encontrar el conjunto $\{\gamma_k(t_k)\}$ que maximice $p(t)$. Discretizando,

$$\begin{cases} a_{k+1} - a_k = -h(\frac{1}{\tau}(1 - a_k) - \gamma_k \cdot (\frac{1}{\varepsilon} + a_k \cdot (\frac{1}{\varepsilon} - \frac{1}{\tau}))) \\ p_{k+1} - p_k = h(\gamma_k a_k) \end{cases}$$

donde h es un paso temporal, $k = 0, 1, 2, \dots, N$, de forma que $a(0) = a_0$, $p(T) = p_N$, puesto que $T = \{t_1, t_2, \dots, t_N\}$. En la versión discretizada el problema se puede reformular así (sabiendo que h es constante):

"Encontrar el conjunto de decisiones $\{\gamma_k(t_k)\}$ tal que maximice $\sum_{k=0}^N \gamma_k a_k$ "

Esto es, debemos calcular las $\{\gamma_k(t_k)\}$ tal que

$$p_N = \max_{\gamma_0, \gamma_1, \dots, \gamma_N} \sum_{k=0}^N \gamma_k a_k,$$

que como se inicia en a_0 , denotamos $p_N^{MAX}(a_0)$.

Para resolverlo aplicamos la Regla de Bellman (*An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.* [Bellman, 1957]). El criterio de Bellman calcula la secuencia completa $(\gamma_0, \gamma_1, \dots, \gamma_N)$ de forma recursiva y hacia atrás, por tanto,

$$p_N^{MAX}(a_0) = \max_{\gamma_0} [\gamma_0 a_0 + \max_{\gamma_1, \dots, \gamma_N} \sum_{k=1}^N \gamma_k a_k], \text{ donde}$$

$$p_{N-1}^{MAX}(a_1) = \max_{\gamma_1, \dots, \gamma_N} \sum_{k=1}^N \gamma_k a_k,$$

iterando de nuevo. Obtenemos la secuencia,

$$p_N^{MAX}(a_0) = \max_{\gamma_0} [\gamma_0 a_0 + \max_{\gamma_1} [\gamma_1 a_1 + \max_{\gamma_2} [\gamma_2 a_2 + \dots + \dots + \max_{\gamma_N} [\gamma_N a_N]] \dots]$$

Para resolver el sistema, empezamos a resolverlo desde la última decisión a la primera. Como la última no afecta al futuro, la maximización es local. En nuestro caso, consiste en:

$$\gamma_N(t_N) = \begin{cases} 1, & \text{si } a_N \in (a_M, 0), \dot{a}_N < 0 \\ 0, & \text{si } a_N \in (0, a_M), \dot{a}_N > 0 \end{cases}$$

por tanto,

$$p_0^{MAX}(a_N) = \begin{cases} a_N, & \text{si } a_N \in (a_M, 0), \dot{a}_N < 0 \\ 0, & \text{si } a_N \in (0, a_M), \dot{a}_N > 0 \end{cases}$$

esto es, según el sistema se encuentre en fase de generación de soluciones o de ejecución. Una vez que sabemos cuál es la decisión óptima en $\gamma_N(t_N)$ calculamos para el instante anterior $\gamma_{N-1}(t_{N-1})$ aplicando la ecuación:

$$p_1^{MAX}(a_{N-1}) = \max_{\gamma_{N-1}} [\gamma_{N-1} a_{N-1} + p_0^{MAX}(a_N)]$$

sabemos que

$$a_N = a_{N-1} - \frac{h}{\tau}((1 - a_{N-1}) - \gamma_{N-1} \cdot (1 + a_{N-1} \cdot (\frac{\tau}{\varepsilon} - 1))), \text{ por tanto}$$

$$p_1^{MAX}(a_{N-1}) = \max_{\gamma_{N-1}} [\gamma_{N-1} a_{N-1} + p_0^{MAX}(a_{N-1} - \frac{h}{\tau}((1 - a_{N-1}) - \gamma_{N-1} \cdot (1 + a_{N-1} \cdot (\frac{\tau}{\varepsilon} - 1))))]$$

Dado que $\gamma_{N-1} = \{0, 1\}$, sólo tendremos que calcular cuál de los dos casos es mayor:

$$a_{N-1} + p_0^{MAX}[a_{N-1} - \frac{h}{\tau}((1 - a_{N-1}) - (1 + a_{N-1} \cdot (\frac{\tau}{\varepsilon} - 1)))] \geq p_0^{MAX}[a_{N-1} - \frac{h}{\tau}((1 - a_{N-1}) - \gamma_{N-1} \cdot (1 + a_{N-1} \cdot (\frac{\tau}{\varepsilon} - 1)))]$$

que, simplificando, queda como:

$$a_{N-1} + p_0^{MAX}[(1 - \frac{h}{\varepsilon}) \cdot a_{N-1}] \geq p_0^{MAX}[(1 - \frac{h}{\tau}) \cdot a_{N-1} + \frac{h}{\tau}]$$

La condición de equilibrio se cumple para un valor de a_{N-1} crítico, que denotamos, a_{N-1}^* , con el que se puede reescribir la ecuación del siguiente modo:

$$p_1^{MAX}(a_{N-1}) = \begin{cases} a_{N-1}^* + p_0^{MAX}[(1 - \frac{h}{\varepsilon}) \cdot a_{N-1}], & \text{si } a_{N-1} > a_{N-1}^* \\ p_0^{MAX}[(1 - \frac{h}{\tau}) \cdot a_{N-1} + \frac{h}{\tau}], & \text{si } a_{N-1} \leq a_{N-1}^* \end{cases}$$

El procedimiento se puede repetir para ($k = 2, \dots, N$) obteniéndose los valores $\{a_0^*, a_1^*, \dots, a_{N-1}^*, a_N^*\}$, que se obtienen resolviendo iterativamente la ecuación:

$$p_{N-k}^{MAX}(a_{N-k}) = \max_{\gamma_{N-k}} [\gamma_{N-k} a_{N-k} + p_{k-1}^{MAX}(a_{N-k} - \frac{h}{\tau}((1 - a_{N-k}) - \gamma_{N-k} \cdot (1 + a_{N-k} \cdot (\frac{\tau}{\varepsilon} - 1))))],$$

donde $k = 1, \dots, N$.

La Figura G.4 muestra el aspecto de la estrategia óptima para diferentes valores de τ y ε :

Con esta estrategia se puede medir el valor de la p y de la P promedio. El resultado es mayor que en el caso del modelo clásico de toma de decisiones.

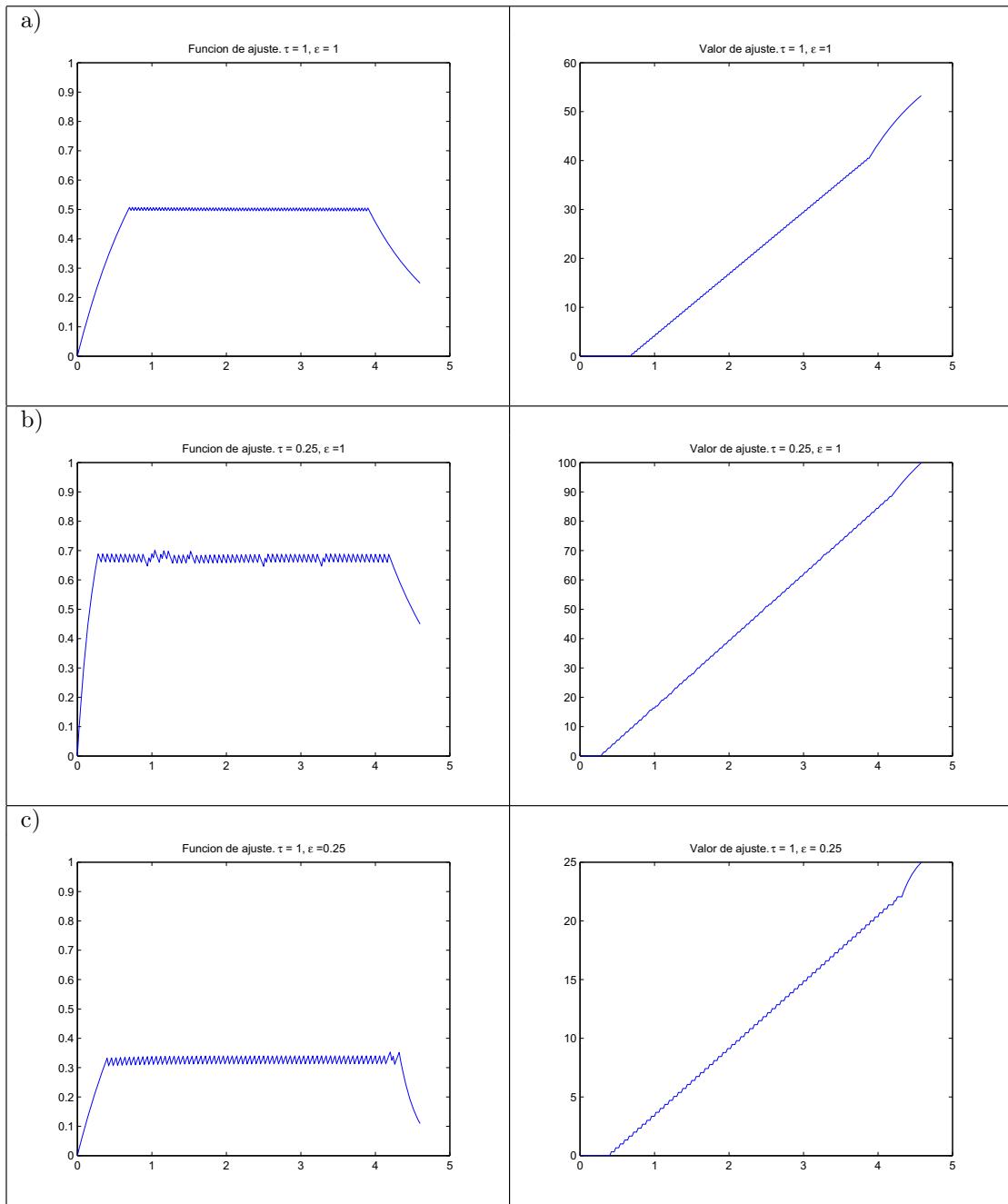


Figura G.4: Representación de diferentes funciones y valores de ajustes para diferentes valores de τ y ε . Vemos que la calidad de la solución escogida en cada caso varía dependiendo de los parámetros.

Apéndice H

Escala de las redes de mundo pequeño

Según sus propiedades estadísticas, las redes de mundo pequeño encontradas en el mundo real se puede dividir en tres tipos de redes [Amaral et al, 2000]: (a) redes sin escala, que poseen una distribución de conexiones entre nodos que decae según una ley de potencia; (b) redes de escala amplia, caracterizadas por una distribución de conectividad que tiene una distribución en ley de potencia seguido de una caída brusca y (c) redes de escala simple, caracterizadas por distribuciones de conectividad con caídas rápidas, como pueden ser las distribuciones exponenciales o gaussianas. Las redes sin escala emergen en contextos en los que en la creación de la red los nuevos nodos se conectan preferencialmente a aquellos nodos que poseen más conexiones. Estas redes tienen la propiedad de ser muy resistentes a la eliminación aleatoria de nodos, aunque la eliminación de alguno de los nodos que poseen un gran número de conexiones puede suponer el colapso de la red. Las redes de escala amplia y escala única emergen en situaciones similares a las de la creación de redes sin escala, con la diferencia de que existen restricciones que limitan de alguna forma la creación de nuevas conexiones en los nodos. Cuando las redes no son sin escala, se evita que casi todo el tráfico dependa de unos pocos nodos, de forma que la eliminación de estos no supondrá un coste tan grave, si bien serán más sensibles a la eliminación aleatoria de grandes cantidades de nodos.

Para analizar la escala de las redes que surgen del algoritmo evolutivo, realizamos 4 simulaciones diferentes y una vez que las redes han evolucionado, tomamos 10.000 conexiones de cada red en instantes aleatorios. Finalmente, calculamos la distribución de probabilidad del número de conexiones por nodo, e intentamos ajustar el resultado a diferentes tipos de distribuciones: de ley de potencia, exponenciales y gaussianas. Obtenemos como resultado que el ajuste gaussiano casi no tiene error, siendo mucho mejor que los otros (Figura H.1). Esto significa que la distribución de las conexiones tiene una caída rápida, impidiendo la existencia de nodos que tengan una gran parte de las conexiones de la red, como pasa en las redes sin escala. Podemos decir por lo tanto que las redes de mundo pequeño que se forman son redes de escala simple.

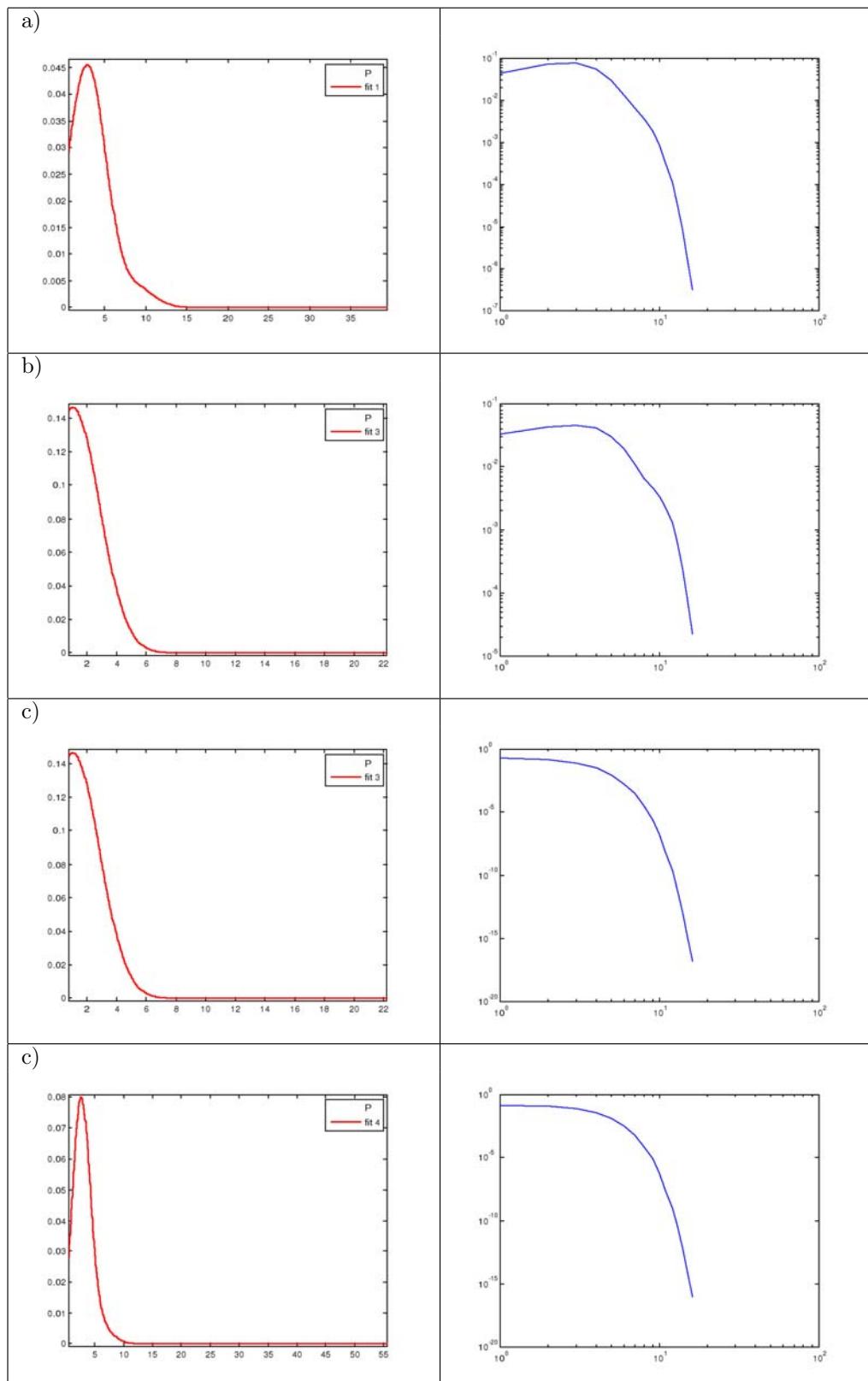


Figura H.1: Resultados al calcular la escala de cuatro de las redes surgidas del algoritmo evolutivo. En la columna de la izquierda está el ajuste de los datos de 10.000 conexiones utilizadas mediante una función gaussiana. A la derecha está la representación log-log de dicha función. Para que las redes fueran sin escala, la representación log-log debería ser lineal, y sin embargo muestra una caída rápida, lo que significa que lo que tenemos son redes de escala única.

Parte III

Bibliografía

- Amaral, L. A. N. Scala, A. Barthélémy, M. Stanley, H. E. (2000). Classes of small-world networks. *PNAS, September 26, 2000*.
- Barandiaran, X. E. & Chemero, A. (2009). Animats in the Modeling Ecosystem. *Adaptive Behavior, Volume 17, Issue 4 (August 2009)*.
- Bechara, A. y Damasio, A.R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, Volume 52 (August 2005) pp. 336-372.*
- Benenson, I. y Torrens, P. (2004), *Geosimulation: Automata-Based Modelling of Urban Phenomena*. John Wiley Sons, London.
- Boman, M., Johansson S.J. y Lybäck, D. (2008). Parrondo strategies for artificial traders. *Intelligent Agent Technology Research and Development, pp 150-159.*
- Buceta, J., Lindenberg, K. y Parrondo, J.M.R. (2002) Stationary and oscillatory spatial patterns induced by global periodic switching, *Physical Review Letters, Vol. 88, No. 2, 024103, 2002.*
- Busemeyer,T., Ryan K. Jessup, Joseph G. Johnson, James T. Townsend (2006). Building bridges between neural models and complex decision making behaviour. *Neural Netw. 2006, 19(8):1047-58.*
- Carvalho, J. (2000), Using AgentSheets to Teach Simulation. *Journal of Artificial Societies and Social Simulation, 3(3).*
- Clark, A. (1997). *Being There: putting, body and world together again*. MIT, Cambridge, MA.
- Collier, N. and North, M.J. (2005), Repast for Python Scripting, *Annual Conference of the North American Association for Computational Social and Organizational Science (NAACSOS), Notre Dame, Indiana, USA.*
- Colombetti, G. (2003). Complexity as a new framework for emotion theories. *LPS - Logic and Philosophy of Science. Vol. 1 · No. 1 · 2003.*
- Damasio, A. R. (1994). *Descartes' Error. Emotion, Reason, and the human Brain*. Grosset/Putnam, New York.
- Damasio, A.R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt Brace & Co., New York.
- Damasio, A.R. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Harcourt Brace & Co., New York
- Damiano, L. y Cañamero, L. (2010). Constructing Emotions. Epistemological groundings and applications in robotics for a synthetic approach to emotions. *AIIB Symposium 2010.*
- De Sousa, R. (1987). *The Rationality of Emotion*. Cambridge, Mass., MIT Press.
- Dennett, Daniel C. (1978). Why not the whole iguana?. *Behavioral and Brain Sciences 1, pp. 103-104.*
- Elster, J. (2002). *Alquimias de la mente*, Paidós.
- Fuster, J.M., (1996). *The Prefrontal Cortex. Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. Raven Press, New York.
- Goodwin, B. and Sole, R. (2000). *Signs of Life: how complexity pervades biology*. Basic Books. 2000

- Harmer, G. P. y Abbott, D. (1999). Game theory: Losing strategies can win by Parrondo's paradox. *Nature* 402, 864 (23 December 1999).
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Kleinberg, J. (1999). The Small-World Phenomenon: An Algorithmic Perspective. *Proceedings of the 32nd ACM Symposium on Theory of Computing*.
- Krichmar, J. and Edelman, G. (2002). Machine Psychology: Autonomous Behavior, Perceptual Categorization, and Conditioning in a Brain-Based Device. *Cerebral Cortex* 12:818-830
- Macia, J. y Solé, R.V. (2008). Distributed robustness in cellular networks: insights from synthetic evolved circuits. *J R Soc Interface*. 2009 Apr 6;6(33):393-400. Epub 2008 Sep 16.
- Martel, C. Nguyen, V. (2004). Analyzing Kleinberg's (and other) small-world Models. *Proc. of ACM Symp. on Princ. of Dist. Comp. (PODC)*.
- MASON (2006), Multi Agent Simulation Of Neighbourhood, Disponible en:
<http://cs.gmu.edu/~eclab/projects/mason>
- McCarthy, J. y Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, 4:463-502.
- Megill, J. y Cogburn, J.. (2005). Easy's Getting Harder all the Time: The Computational Theory and Affective States. *Ratio (New Series)* XVIII 3 (September 2005): 306-316.
- Milgram, S. (1967). The Small-World Problem. *Psychology Today*, vol. 1, no. 1, May 1967, pp. 61-67.
- Najlis, R., Janssen, M.A. and Parker, D.C. (2001), Software Tools and Communication Issues. *Parker, D.C., Berger, T. and Manson, S.M. (eds.), Meeting the Challenge of Complexity: Proceedings of a Special Workshop on Land-Use/Land-Cover Change, Irvine, California*.
- Nichelli, P., (2002). The processing of temporal information in the frontal lobe. In: Grafman, J. (Ed.), *Handbook of Neuropsychology: Frontal Lobes*. Elsevier, Amsterdam, pp. 175-193.
- Oviedo, J. M. (2005). *Programación Dinámica. La Ecuación de Bellman y el Teorema de la Envolvente*. Universidad Nacional de Córdoba - Argentina
- Pope, C. (2007). *Somatic Computationalism: Damasio's Clever Error*. Louisiana State University.
- Port, R. and van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. MIT Press.
- Repenning, A., Ioannidou, A. and Zola, J. (2000), AgentSheets: End-User Programmable Simulations. *Journal of Artificial Societies and Social Simulation*, 3 (3).
- Russell, S. y Norvig, P. (2003). *Inteligencia Artificial Un Enfoque Moderno*. Segunda Edición. Stuart Russell (Prentice Hall)
- Sloman, A. (2010). The Design-Based Approach to the Study of Mind (in humans, other animals, and machines), Including the Study of Behaviour Involving Mental Processes. <http://www.cs.bham.ac.uk/~axs/>
- Spivey, J. (2007). *The continuity of mind*. New York: Oxford University Press.
- Swarm (2006), Swarm: A Platform for Agent-Based Models. Disponible en:
<http://www.swarm.org/>

Vos, J.R. (2005), Repast .NET: The Repast Framework Implemented in the .NET. *Annual Conference of the North American Association for Computational Social and Organizational Science (NAACSOS), Notre Dame, Indiana, USA.*

Watts, D.J. y Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature, vol, June 1998, pp. 440-442.*