

Autopoiesis y una Biología de la Intencionalidad¹

versión 0.8 de la traducción

Francisco J. Varela

Traducción: *Xabier Barandiaran*
xabier@barandiaran.net

13 de junio de 2003

¹Parte de este texto ha sido publicado en Varela (1991)

Notas Previas

Referencia original

Varela, Francisco J. (1992) Autopoiesis and a Biology of intentionality. McMullin, B. and Murphy, N. (eds.) *Autopoiesis & Perception*. pp.1–14. Proceedings of a workshop held in Dublin City University, August 25th & 26th 1992. School of Electronic Engineering Technical Report, Dublin, 1994.

Este artículo, así como el resto de los presentados al congreso pueden encontrarse en:

<ftp://ftp.eeng.dcu.ie/pub/alife/bmcm9401/>
<http://www.eeng.dcu.ie/alife/bmcm9401/>

Copyright del original

Copyright © Francisco Varela 1994. Permission is granted to private individuals to copy and distribute the files in this archive, for purposes of private study only, provided that the files are unmodified, that they are accompanied by a copy of this notice, and that **no** charges are made. Commercial use or distribution of any sort is specifically prohibited.

All other rights reserved.

Copyright de la traducción

Copyright © Xabier Barandiaran 2003. Licencia aún por determinar.

Notas del traductor

La presente versión está pensada para revisar la traducción y no para ser publicada. Es por ello que se incluyen comentarios y referencias a términos originales en inglés totalmente fuera de lugar para un versión publicable.

- He traducido sistemáticamente:
 - *living* por vivo (en lugar de viviente)
 - *entity* por entidad (en lugar de ente)
 - *becomes* por deviene
 - *address* por tratar
 - *network* por red
 - *surplus* por excedente

- *argument* por discusión
- *disucussion* por argument
- *environment* por entorno (fenoménico)
- *surrounding* por alrededores (no fenoménico)¹
- *to emerge through* por emerger a través de²

- no encuentro fácil traducción para:
 - *Distinct*, ni distinguible, ni distinto, quizás distinguido (pero queda mal su uso impersonal)
 - *Bring forth* a veces por hacer surgir, crear... pero no acaba de satisfacerme.
- Entre corchetes [] se encuentran tres tipos de añadidos/comentarios:
 - EN:** es una expresión en inglés en cuya traducción he dudado
 - otros** añadidos posibles al texto que no son traducciones
- Todas las notas a pie de página son notas del traductor, el original no incluye ninguna nota.

¹Varela no lo pone fácil porque hasta el párrafo 13 del texto utiliza esta dicotomía entre *environment/surrounding*, entorno/alrededores y luego más tarde la sustituye por *environmet/world*, entorno/mundo.

²el uso más común es 'emerge de', pero también lo es en inglés 'to emerge from', así que he decidido mantener la expresión "heterodoxa" de Varela

1. Introducción

Como ustedes saben³, autopoiesis es un neologismo, introducido en 1971 por H. Maturana y por mí para designar la organización de un sistema vivo mínimo. El término se hizo representativo de una perspectiva de la relación entre el organismo y su entorno en la que los aspectos de su autoconstitución y autonomía son considerados elementos fundamentales. Desde 1971 hasta ahora han sucedido muchas cosas que vienen a reforzar esta idea. Algunos de los avances tienen que ver con la misma noción de autopoiesis en lo referido a la organización celular y el origen de la vida. Muchos otros tienen que ver con la autonomía y propiedades auto-organizativas del organismo en lo referido a su actividad cognitiva. Por tanto, frente a la postura cognitivista de procesamiento de símbolos dominante en los años 70, hoy en día somos testigos, en las ciencias cognitivas, de un renacimiento de la importancia de la imbricación [*EN:embeddedness*] del agente cognitivo, sea este natural o artificial. Esto viene sucediendo bajo diferentes denominaciones [*EN:labels*] como *nouvelle-AI* (Brooks, 1991), el problema del enraizamiento del símbolo (*symbol grounding problem*) (Harnad, 1991), agentes autónomos en vida artificial (Varela and Bourgine, 1992), o funcionalidad situada (*situated functionality*) (Agree, 1988), por citar algunas de las denominaciones auto-explicativas más recientes.

Cualquiera de estos desarrollos merecería una conferencia completa; evidentemente no puedo hacer eso aquí. Mi intención, sin embargo, valiéndome de la posición de inaugurar este encuentro, es la de intentar señalar una serie de temas fundamentales y fundacionales de la relación entre la autopoiesis y la percepción. De ahí el título de esta conferencia: una biología de la intencionalidad. Desde que la crisis de las ciencias cognitivas clásicas ha dejado abierta la cuestión de la intencionalidad, en mi opinión la autopoiesis ofrece un punto de partida natural hacia una perspectiva [*EN:view*] de la intencionalidad que es fundamental para responder a los grandes obstáculos que han sido señalados recientemente. Volveré sobre este punto al final. Permítanme comenzar por el principio.

³Este artículo está basado en una conferencia para psicólogos que tubo lugar en Dublin City University durante los días 25 y 26 de Agosto de 1992 bajo el título *Autopoiesis and Perception*

2. La Cognición y los Sistemas Vivos Mínimos

2.1. La autopoiesis como esqueleto bio-lógico

La célula bacteriana es el más simple de los sistemas vivos porque posee la capacidad de producir, a través de una red (*network*) de procesos químicos, todos los componentes que conducen a la constitución de una unidad distinta [*EN:distinct*] y limitada [*EN:bounded*]. Para evitar ser trivial el atributo ‘vivo’ (*living*) en la descripción anterior debe señalar el *proceso* que permite tal constitución, no las materialidades [*EN:materialities*] que la constituyen, ni una enumeración de propiedades. ¿Pero qué es este proceso básico? Su descripción debe situarse a un nivel muy específico: debe ser lo suficientemente universal como para permitirnos reconocer los sistemas vivos como una clase, sin una referencia esencial a los componentes materiales. Pero al mismo tiempo no debe ser demasiado abstracta, esto es, debe ser lo suficientemente explícita como para permitirnos ver tales patrones dinámicos en acción en los mismos [*EN:actual*] sistemas vivos que conocemos en la tierra, en aquellos que pudieran ser encontrados en otros sistemas solares, e incluso [*EN:eventually*] en aquellos creados artificialmente por el hombre. En palabras del organizador de un congreso de vida artificial: “sólo cuando seamos capaces de ver la *vida-tal-como-la-conocemos* (*live-as-we-know-it*) en el contexto más amplio de la *vida-tal-como-podría-ser* (*life-as-it-could-be*) podremos realmente comprender la naturaleza de la bestia” (Langton, 1989a, p. 2)

Hace ya algunos años que la biología celular contemporánea ha hecho posible desplegar [*EN:to put forth*] la caracterización de esta organización básica de lo vivo —una bio-lógica— como la de un sistema *autopoiético* (del griego: auto-producción (Maturana and Varela, 1980; Varela et al., 1974)). Un sistema autopoiético —la organización mínima de lo vivo— es aquél que continuamente produce los componentes que lo especifican, al tiempo que lo realizan (al sistema) como unidad concreta en el espacio y en el tiempo, haciendo posible la propia red de producción de componentes. Definido con mayor precisión: Un sistema autopoiético está organizado (definido como unidad) como una red de

procesos de producción (síntesis y destrucción) de componentes, de tal manera que estos componentes:

- (i) continuamente regeneran y realizan la red [EN:network] que los produce, y
- (ii) constituyen el sistema como una unidad distinguible en el dominio en que existen

Por lo tanto, la autopoiesis pretende capturar los mecanismos o procesos que generan la *identidad* de lo vivo, y servir, así, como distinción categorial [EN:as a categorical distinction] de lo vivo frente a lo no-vivo. Esta identidad equivale a coherencia auto-producida: el mecanismo autopoietico se mantendrá a sí mismo como unidad distinta mientras su concatenación básica de procesos se mantenga intacta en presencia de perturbaciones, y desaparecerá cuando se enfrente a perturbaciones que superen cierto umbral de viabilidad que depende del sistema bajo consideración. Obviamente, todas las trayectorias [EN:pathway] bioquímicas y formación de membranas en las células pueden ser inmediatamente recogidas bajo esta definición de autopoiesis.

Un ejercicio diferente —que no persigo aquí en modo alguno— es el de ver cómo esta organización autopoietica básica, presente en los orígenes de la vida terrestre (Fleischaker, 1988) se convierte progresivamente más compleja a través de mecanismos reproductivos, compartimentalización, dimorfismo sexual, formas de nutrición, simbiosis, etc. dando lugar a la vida pro- y eucariota en la tierra (Margulis, 1981; Fleischaker, 1988). Concretamente, en este punto asumo la perspectiva de que la reproducción *no* es intrínseca a la lógica mínima de lo vivo. La reproducción debe ser considerada como una añadida complejificación superpuesta sobre una identidad más básica, la de la unidad autopoietica, una complejificación que es necesaria dadas las constricciones impuestas por las condiciones primarias/iniciales/tempranas [EN:early] en un planeta turbulento. La reproducción es esencial para la viabilidad de lo vivo, pero sólo cuando hay una identidad puede una unidad reproducirse. En este sentido, la identidad tiene una prioridad lógica y ontológica sobre la reproducción, aunque no precedencia histórica.

No perseguimos aquí estas complejificaciones históricas, ni tampoco persigo otra, igualmente pertinente, pregunta empírica: ¿puede una estructura

molecular, más simple que la ya intrincada célula bacteriana, satisfacer el criterio de organización autopoietica? Esta pregunta puede ser contestada a través de dos enfoques diferentes: (1) simulaciones y (2) síntesis de sistemas autopoieticos mínimos. Hay avances en ambos frentes. En relación al primero [EN:as to the first], hay algunos resultados nuevos en el estallido de investigaciones en vida artificial, en parte extendiendo nuestro temprano [EN:early] trabajo en autómatas de tesela [EN:tesselation automata⁴] (Varela et al., 1974). El segundo frente toma la forma de un acercamiento al origen de la vida centrado en la célula y que busca corporalizaciones [EN:embodiments] químicas de sistemas autopoieticos mínimos. De hecho, la encapsulación [EN:encapsulation] de macromoléculas por vesículas lípidas [EN:lipide vesicles] ha sido activamente investigada como una prometedora candidata a célula primigenia [EN:early cell] (Deamer and Barchfeld, 1982; Lazcano, 1986; Baeza et al., 1987); véase (Deamer, 1986). Luisi y Varela (1989) afirman que un sistema micelar inverso [EN:a reverse micellar system] puede acercarse mucho al límite de ser considerado un sistema autopoietico mínimo. En concreto, discuten el caso de un sistema micelar inverso [EN:a reverse micellar system] que acoge en su interior acuoso una reacción que conduce a la producción de surfactantes [EN:surfactants], que constituyen una membrana [EN:boundary] para la reacción micelar inversa [EN:reverse micellar reaction]. El interés de este caso radica en que se sabe mucho de estos sistemas químicos, lo que hace posible la puesta en marcha de un sistema autopoietico mínimo. Pero debo abandonar estos fascinantes resultados [EN:issues] para volver sobre el tema que he escogido para esta conferencia.

2.2. Identidad de lo vivo y su entorno

La autopoiesis trata el tema del organismo como sistema vivo mínimo caracterizando su *modo básico de identidad*. Esto es, propiamente hablando, tratar el tema a un nivel ontológico: el acento recae sobre la forma en que un sistema se convierte [o deviene EN:becomes] en una entidad distinguible, y no sobre su composición molecular específica y configuraciones históricas contingentes. Ya que,

⁴N.T. Explicar lo que es un autómata de tesela

en tanto que existe, la organización autopoietica se mantiene invariante. En otras palabras, un modo de aclarar la especificidad de la autopoiesis es pensar en su auto-referencialidad como aquella organización que mantiene la propia organización como invariante. Toda la constitución físico-química está en flujo constante; el patrón se mantiene, y solamente a través de su invarianza puede determinarse [EN:*ascertained*] el flujo de los componentes que lo realizan.

Aquí sólo he señalado la organización mínima que da lugar a semejante autonomía viva. Tal y como he dicho, mi propósito es el de subrayar la bio-lógica básica que sirva de fundamento desde el que poder considerar la diversidad presente en los organismos actuales: sólo cuando hay una identidad pueden sus ramificaciones [EN:*elaborations*⁵] considerarse variaciones familiares de una clase común de unidades vivas. Toda clase de entidad tiene una identidad que la caracteriza; la singularidad de lo vivo reside en el tipo de organización que posee.

Ahora bien [EN:*now*], la historia de la biología está, por supuesto, enturbiada [EN:*marred*] por la tradicional oposición entre mecanicistas/reduccionistas por un lado y holistas/vitalistas por otro, herencia del espacio-problemático [EN:*problem-space*] de la biología del siglo XIX. Una de las contribuciones específicas del estudio de los mecanismos auto-organizativos —de los cuales la autopoiesis es un ejemplo particular [EN:*specific*]— es que la tradicional distinción entre elementos componentes y propiedades globales desaparece. En el sencillo ejemplo del autómata celular explicado anteriormente [sic]⁶ lo que queda en evidencia es precisamente la *reciprocidad causal* de las reglas locales de interacción (e.d. las reglas de los componentes, que son análogas a las interacciones químicas) y las propiedades globales de la entidad (su demarcación topológica que afecta a la difusión y creación de condiciones locales para la reacción). Me parece que esta causalidad recíproca hace mucho para evacuar [EN:*does much to evacuate*] la oposición mecanicismo/vitalismo, y nos permite movernos a una fase más productiva de identificación de diferentes *modos* de auto-organización en los que lo local y lo global se

encuentran explícitamente entrelazados gracias a esta causalidad recíproca. La autopoiesis es un inmejorable ejemplo de este tipo de dialéctica entre el nivel de componentes locales y el todo global, unidos en relación recíproca a través de los requerimientos constitutivos de una entidad que se auto-separa de su entorno. En este sentido la autopoiesis, como caracterización de lo vivo, no cae en los extremos tradicionales tanto del vitalismo como del reduccionismo.

Una segunda y complementaria dimensión de la bio-lógica básica que resulta central para enfocar nuestra argumentación es la naturaleza de la relación entre unidades autopoieticas autónomas y su entorno. Es evidente *ex-hypothesis* que un sistema autopoietico depende de su entorno físico-químico para su conservación como unidad separada, de lo contrario volvería a disolverse en él. De aquí la intrigante paradoja propia de una identidad autónoma: el sistema vivo debe diferenciarse de su entorno, mientras mantiene *al mismo tiempo* su acoplamiento [con él]; este engarzamiento (*linkage*) es inseparable ya que es precisamente en contraste con ese entorno que el organismo surge [EN:*comes forth*][como unidad diferenciada]. Ahora, en este acoplamiento dialógico entre la unidad viva y su entorno físico-químico la balanza se encuentra ligeramente inclinada del lado de lo vivo ya que posee el rol activo en el acoplamiento recíproco. Al tiempo que se define a sí mismo como unidad define qué queda fuera de él, esto es, su entorno circundante [EN:*surrounding environment*]. Un examen más detenido también hace evidente que esta exteriorización sólo puede ser entendida, por decirlo de algún modo, desde “dentro”; la unidad autopoietica *crea una perspectiva* desde la que el exterior es uno, que no puede ser confundido con sus alrededores físicos [EN:*physical surroundings*] tal y como se nos aparece como observadores, exentos de tal perspectivismo, i.e. el terreno [EN:*the land*] de las leyes físicas y químicas *simpliciter*.

En nuestra práctica como biólogos alternamos constantemente entre estos dos dominios. Usamos y manipulamos principios y propiedades físico-químicas, mientras cambiamos alternativamente al uso de la *interpretación* y la significatividad vista desde la perspectiva del sistema vivo. Así pues [EN:*thus*], una bacteria nadando en un gradiente de azúcar queda convenientemente analizada en términos de los efectos locales del azúcar en la permeabilidad

⁵ELABORATE = complicar.

⁶N.T.: Esto es un error del original compuesto en parte por fragmentos de otro artículo (Varela, 1991) FALTA, HABRÍA QUE VER EL ORIGINAL

de la membrana, la viscosidad del medio, la hidromecánica del movimiento del flagelo, etc. Pero por otro lado el gradiente de azúcar y el movimiento del flagelo resultan de interés para el análisis sólo porque la bacteria en su totalidad señala esos elementos como relevantes: su significatividad específica [*EN:specific significance*] como componentes de la conducta alimenticia sólo es posible a través de la presencia y la perspectiva de la bacteria como una totalidad. Quítese a la bacteria como unidad y todas las correlaciones entre gradiente y propiedades hidrodinámicas se convierten en leyes químicas del entorno, evidentes para nosotros como observadores pero carentes de significatividad especial alguna.

Me he metido en esta larga arena [*EN:I have gone into this lengthy harangue*] porque creo que esta relación verdaderamente dialéctica es un punto clave. De hecho, puede parecer tan obvio que no apreciemos sus ramificaciones más profundas. Me refiero a las distinción entre el entorno del sistema vivo tal y como se le aparece a un observador y sin referencia a la unidad autónoma —que denominaremos de aquí en adelante simplemente *entorno* [*EN:environment*]— y el entorno *para* el sistema que es definido en el mismo momento que dio lugar a su identidad y que sólo existe en esa definición mutua —de aquí en adelante el *mundo* del sistema.

La diferencia entre entorno y mundo es el excedente [*EN:surplus*] de *significación* que acoge la comprensión de lo vivo y la cognición, y que se encuentra a la raíz [*EN:at the root*] de cómo un yo deviene en uno [*EN:of how a self becomes one*]. En otras palabras, este excedente es la madre de la intencionalidad. En la práctica resulta bastante difícil mantener a la vista [*EN:to keep in view*] la dialéctica de esta definición mutua: ni aislamiento rígido, ni simple continuidad con la química física [*physical chemistry*]. Por el contrario [*EN:in contrast*], es fácil confundir [*EN:conflate*] el mundo de la unidad con su entorno ya que es *tan* obvio que estamos estudiando ésta o aquella interacción molecular en el *contexto* de una unidad celular autónoma que resulta fácil fallar completamente a ver [*EN:miss completely*] el excedente *añadido* por la perspectiva del organismo. No existe significatividad alimenticia en la sacarosa [*EN:sucrose*] excepto cuando una bacteria nada gradiente arriba y su metabolismo hace uso de esa molécula como un modo de permitir que continúe su identidad. Este excedente no es, evidentemente, indiferente a las regularidades y textura

que operan en el entorno (i.e. las “leyes”), que la sacarosa pueda crear un gradiente y atravesar una membrana celular, etc. Por el contrario [*EN:on the contrary*], el mundo del sistema está construido sobre estas regularidades, que es lo que asegura que pueda mantener su acoplamiento en todo momento.

Lo que hace el sistema autopoietico —gracias a su mismísimo modo de identidad [*EN:due to its very mode of identity*]— es enfrentarse constantemente a los encuentros (perturbaciones, choques [*EN:shocks*], acoplamientos) con su entorno y tratarlos desde una perspectiva que no es intrínseca a los encuentros mismos. Seguramente a las rocas o abalorios [*EN:beads*⁷] de cristal no les llame la atención el gradiente de azúcar, de entre las infinitas posibilidades de interacciones físico-químicas, como especialmente significativa —para que esto suceda es esencial una perspectiva *desde* una identidad activamente constituida. Es tentador, en este punto, deslizarse dentro de las vaporosas nubes [*EN:vaporous clouds*] acerca del “significado”, reminiscente del peor tipo de vitalismo del pasado o de la jerga informacional del presente. Lo que enfatizo aquí es que aquello que es significativo para un organismo es precisamente dado por su constitución como un proceso distribuido, con una unión [*EN:link*] indisoluble entre procesos locales en los que ocurre una interacción (i.e. fuerzas físico-químicas actuando en la célula), y la entidad coordinada que es la unidad autopoietica, dando lugar a manejarse con su entorno sin necesidad de acudir a un agente central que gire la manilla desde el exterior —al modo de *élan vital*— o una orden preexistente en una localización particular —al modo de un programa genético esperando ser expresado.

Me gustaría reconstruir esta idea básica dándole la vuelta. La constante creación [*EN:bringing forth*] de significación es lo que podemos describir como una ausencia permanente en lo vivo: está constantemente creando [*EN:bringing forth*] una significación que falta, no pre-dada o pre-existente. La relevancia debe ser provista [*EN:provided*] *ex nihilo*: distinguir entre especies moleculares relevantes e irrelevantes, seguir un gradiente en contra y no a favor, aumentar la permeabilidad a este ion y no a este otro, etc. Hay un inevitable contratiempo [o revés *EN:contretemps*⁸] entre un sistema autónomo

⁷viene así traducido en el diccionario —para cristales

⁸ambos contratiempo y revés aparecen como traducciones

y su entorno: siempre hay algo que el sistema debe proporcionar desde su perspectiva como todo funcional. De hecho, un encuentro molecular adquiere una significación en el contexto de sistema operativo *completo* y de muchas interacciones simultáneas.

La fuente de esta creación-de-mundos es siempre producida por discontinuidades en la autopoiesis, sean menores, como cambios en la concentración de algún metabolito, o mayores, como la disrupción de la membrana. Debido a la naturaleza misma de la autopoiesis —ejemplificada en la reparación de la membrana del ejemplo anterior de simulación mínima [sic⁹]— toda discontinuidad puede ser vista como el comienzo de una acción sobre lo que falta en la parte del sistema para que la identidad pueda ser mantenida. Repito: no hay teleología alguna supuesta [EN:*implied*] en este “para que”: esto es lo que la lógica auto-referencial de la autopoiesis implica en primer lugar. La acción realizada podrá ser vista como un intento de modificar su mundo —cambio de lugar de diferentes nutrientes, aumento del flujo de un metabolito para síntesis metabólica, etc.

En resumen, esta permanente e implacable acción sobre lo que falta [(sobre lo ausente)] se convierte, desde el lado del observador, en la continua actividad *cognitiva* del sistema, que está a la base de la inconmensurable diferencia entre el entorno en el que el sistema es observado, y el mundo en el que opera el sistema. En su propio origen [EN:*at its very root*] esta actividad cognitiva resulta paradójica. Por un lado la acción que da lugar [EN:*brings forth*] a un mundo es un intento de restablecer un acoplamiento con un entorno que desafía la coherencia interna a través de encuentros y perturbaciones. Pero estas acciones, al mismo tiempo, demarcan y separan al sistema de su entorno, dando lugar a un mundo diferenciado [EN:*distinct*].

Puede que el lector rechace mi uso del término cognitivo para sistemas celulares, así como mi atrevido deslizamiento en el terreno de la intencionalidad. Como ya he dicho, uno de mis principales argumentos [EN:*main points*] aquí es que salimos ganando al tener en cuenta la *continuidad* entre este nivel fundamental del yo [EN:*self*] y otro yoes regionales [o particulares EN:*regionals*], incluyen-

de contreptem. A mí simplemente no me parece un término apropiado ni en castellano ni en el original inglés

⁹Una vez más conviene mirar el original Varela (1991) para ver a qué ejemplo se refiere.

do los neuronales y lingüísticos con los que no vacilaríamos al utilizar la palabra cognitivo. Supongo que otros, en cambio, preferirían introducir la palabra “información”. Pues bien [EN:*well*], hay razones por las que creo que esto sería incluso más problemático. A pesar de que está claro que describimos un *X* que perturba desde la exterioridad del organismo, *X* no es información. De hecho, para el organismo sólo es un *esto*, un *algo*, una *cosa* [EN:*basic stuff*] básica que in-formar desde su propia perspectiva. En términos físicos hay cosas [EN:*there is some stuff*], pero no son para nadie. En cuanto hay alguien (un cuerpo¹⁰)—incluso en esta forma mínima— se convierte en in-formada para un yo [EN:*self*], a través de la dialéctica recíproca que acabo de explicar. Tal in-formación no es jamás una significación fantasmal o un bit de información esperando a ser recolectado [EN:*harvested*] por el sistema. Es una presentación, una ocasión para el acoplamiento y es en este *entre-deux* que surge el significado (Varela, 1979; Castoriadis, 1987).

Por lo tanto el término cognitivo posee dos dimensiones constitutivas: en primer lugar su dimensión de *acoplamiento*, esto es, una unión [EN:*link*] con su entorno que permite su continuidad como entidad individual; en segundo lugar —forzando el lenguaje, lo reconozco— su dimensión *imaginaria*, es decir, el excedente de significatividad que una interacción adquiere desde la perspectiva de la acción global del organismo.

3. Percepción-acción y neurológica básica

3.1. Clausura operacional del sistema nervioso

En la sección anterior, he presentado la imbricación [EN:*interlock*] fundamental entre identidad y cognición tal como se muestra en un organismo

¹⁰Se da entre estas dos frases un juego de palabras intraducible y que refleja además la importancia radical del cuerpo y la corporalización en la filosofía de la biología y la cognición de Francisco Varela. Estas dos últimas frases en el original dicen así: “In physical terms there is stuff, but it is for nobody. Once there is body—even in this minimal form—it becomes in-formed for a self, in the reciprocal deialectics I have just explicated”. La palabra ‘nadie’ (*nobody*) significa lo mismo que ‘sin-cuerpo’ o ‘ningún-cuerpo’ (*no-body*).

mínimo. En esta sección quiero mostrar cómo el nivel tradicionalmente considerado el propiamente cognitivo, el que gira en torno al cerebro de animales multicelulares, es en cierto modo la continuación del mismo proceso.

El cambio de celularidad mínima [*EN: minimal cellularity*]) a organismos con sistema nervioso es repentino y esconde [*EN: skips*] la complejidad de las diversas maneras en las que los organismos multicelulares surgen y evolucionan (Margulis and Schwartz, 1988; Buss, 1987; Bonner, 1988). Ésta es una transición de unidades de selección, y una que implica el equilibrio somático [*EN: somatic balance*] de poblaciones celulares diferenciadas en un organismo adulto, así como ingeniosos recorridos de desarrollo que permiten establecer una estructura corporal. Tal como Buss afirmaba recientemente: “La evolución del desarrollo [somático] es la generación de una ‘ecología somática’ que media en los conflictos potenciales entre las células y el individuo, mientras el organismo interactúa efectivamente con el entorno extrasomático” (Buss, 1987).

Para la mayoría de los vertebrados, esta “ecología somática” está formada [*EN: bound together*] por la red de linfocitos que constituyen el centro del sistema inmune. Una vez más la discusión entorno a un yo [*EN: self*] inmunológico no es mi propósito aquí. No puedo resistir a la tentación, sin embargo, de señalar [*EN: for completeness sake*] que en otro lugar he presentado *in extenso* un tratamiento de red del sistema nervioso y su papel en el establecimiento de un yo celular/molecular flexible durante el desarrollo ontogénico [*EN: ontogeny*] de los mamíferos, véase (Varela et al., 1988; Varela and Coutinho, 1991). En mi opinión [*EN: in my view*] esta identidad no es, tal como se considera tradicionalmente, una demarcación del yo como defensa contra el *no-yo* de antígenos invasores. Es la afirmación positiva auto-referencial de una unidad coherente — una “ecología somática” — mediada por inmunoglobulinas y marcadores celulares libres en intercambio dinámico. Las reacciones inmunológicas contra infecciones, aunque claramente importantes, están mediadas por un sistema inmune “periférico”, una sub-población de diferentes linfocitos movilizada no a través de mecanismos de expansión de la red sino a través de de mecanismos de expansión clónicos, al modo de una reactividad refleja [*EN: reflex reactivity*] adquirida a través de la evolución. Pero limitemos nuestro excursus [*EN: But enough of this*

excursus]. Para mis objetivos aquí daré por dada la identidad de un organismo multicelular, distinta, en su modo de identidad, a una entidad autopoiética mínima, pero similar en tanto que demarca una entidad autónoma de su entorno.

Ahora bien, ¿cuál es el lugar específico que ocupa el sistema nervioso en las operaciones corporales [*EN: in the bodily operations*] de un multicelular? En cuanto el *movimiento* sea parte integral del estilo de vida de un multicelular, hay un consiguiente desarrollo del sistema nervioso que une efectores (músculos, secreciones) y superficies sensoras (órganos sensitivos, terminaciones nerviosas). *La lógica fundamental del sistema nervioso es la de acoplar movimientos con una corriente de modulaciones sensoriales, de una forma circular.* El resultado neto son correlaciones de percepción-acción [o sensomotoras] surgidas de y moduladas por un conjunto de neuronas entrelazadas, la red *interneuronal*. Igualmente, las neuronas son únicas entre las células de un organismo multicelular en sus ramificaciones axonales y dentríticas que permiten múltiples contactos, que se extienden a largas distancias (relativas al tamaño somático de las células) constituyendo el medio esencial para estas correlaciones sesomotoras intra-organísmicas.

Frente a lo que suele ser habitual [*EN: contrary to current habit*], me gustaría enfatizar desde el comienzo el *situacionismo* [*EN: situatedness*] de esta neuro-lógica: el estado de activación de los sensores viene dado, la *mayoría* de las veces [*EN: most typically*], por los movimientos del organismo. En gran medida, la conducta es la regulación de la percepción. Esto no excluye, por supuesto, perturbaciones independientes que vienen del entorno. Pero lo que típicamente se describe como “estímulo” en el laboratorio, una perturbación que es deliberadamente independiente de la actividad presente del organismo, es menos pertinente [*EN: less pertinent*] (fuera del laboratorio) para entender la biología de la cognición.

Las coherencias percepto-motoras, que describimos externamente como conducta, disfrazan el surgimiento, dentro de la red interneuronal, de un gran subconjunto — una agrupación [*EN: ensemble*] como suele decirse — de neuronas fugazmente [*EN: transiently*] correlacionadas. Estas agrupaciones son tanto la fuente como el resultado de la actividad de las superficies sensoras y motoras. Lo que cambia es la cantidad de interneuronas me-

diadoras, y la arquitectura específica del sistema nervioso concreto, que contiene varias regiones corticales, capas y núcleos. En los seres humanos alrededor de 10^{11} interneuronas interconectan alrededor de 10^6 neuronas motoras en relación a 10^7 neuronas sensoras distribuidas por todo el cuerpo en superficies receptoras. Ésta es una proporción de 10:100.000:1 de interneuronas mediando el acoplamiento entre superficies sensoras y motoras. La ascensión y el declive de la auto-organización neuronal, digamos, en la modesta retirada gradual de la *Aplysia* [*EN:modest Aplysia siphon withdrawal*] (Zecevic et al., 1989), es incluso más válida en cerebros de mayor tamaño [*EN:is all the more valid in larger brains*]. Así, por ejemplo, un estudio en el gato (John et al., 1986) muestra que de 5 a 100 millones de neuronas están activas a lo largo del cerebro durante la simple tarea visuo-motora de apretar una palanca. Semejantes conjuntos neuronales surgen a través de un mosaico de áreas regionales, mostrando el enorme paralelismo distribuido de los cerebros de vertebrados.

Las dinámicas neuronales que subyacen a una tarea percepto-motora, es un asunto de redes [*EN:is a network affair*], un sistema bidireccional, altamente cooperativo, y no un una secuencia paso a paso de abstracción de información. La densa interconexión entre sus sub-redes implica que toda neurona activa opere como parte de un largo y distribuido conjunto de neuronas en el cerebro, incluyendo regiones locales y distantes. Por ejemplo, a pesar de que las neuronas en el cortex visual tienen respuestas diferentes para “características” específicas del estímulo visual (posición, dirección, contraste, etc.), estas respuestas ocurren sólo en una animal anestesiado con un entorno (interno y externo) enormemente simplificado. En cuanto se permiten condiciones sensoras más normales se ha puesto crecientemente [*EN:increasingly*] de manifiesto que las respuestas neuronales estereotipadas a esas “características” son altamente dependientes [*EN:labile*] y sensitivas al contexto. Esto ha sido mostrado, por ejemplo, para el efecto de la inclinación corporal en el estimulación auditiva. Más aún, las características de respuesta de la mayoría de las neuronas en el cortex visual dependen directamente de otras neuronas localizadas lejos de sus campos receptivos (ver, por ejemplo, Allman et al. (1985)); incluso un cambio en la postura, al tiempo que se mantiene el mismo e idéntica estímulo sensorial, altera las respuestas

neuronales, demostrando que incluso el supuesto final del flujo neuronal, el *motorium*, entra en resonancia con el *sensorium* (Abeles, 1984).

Si me permiten seguir usando la visión como ejemplo, puedo llevar la discusión previa a un nivel superior de generalización, para señalar [*EN:to note*] que en los últimos años la investigación se ha convertido en el estudio, no de una “reconstrucción” centralizada de una escena visual para el beneficio de un ulterior homúnculo [*EN:homunculus*], sino el de un *mosaico* de modalidades visuales que incluyen, al menos, la forma (figura [*EN:shape*], tamaño, rigidez), propiedades de la superficie (color, textura, reflectancia [*EN:specular reflectance*]¹¹, transparencia), relaciones espaciales tridimensionales (posiciones relativas, orientación tridimensional en el espacio, distancia) y movimiento tridimensional (trayectoria, rotación). Se ha hecho evidente que estos aspectos diferentes de la visión son propiedades emergentes de sub-redes concurrentes, que tienen cierto grado de independencia e incluso separabilidad anatómica, se correlacionan entrecruzadamente [*EN:cross-correlate*] y trabajan conjuntamente de tal manera que un percept visual es esta coherencia.

Este tipo de arquitectura es fuertemente reminiscente de una “sociedad” de agentes, por usar la metáfora de Minsky (1987). Esta multiplicidad multidireccional es contraintuitiva pero típica de los sistemas complejos. Son contraintuitivas porque estamos acostumbrados al modo causal tradicional de direccionalidad tipo input-procesamiento-output. No hay nada en la subsiguiente [*EN:foregoing*] descripción que sugiera que el cerebro opera como una computadora digital, con procesamiento de información de estado-a-estado; esta descripción tan popular para este tipo de sistemas simplemente no da en el clavo [*EN:simply goes against the grain*]¹².

Por el contrario, a la red y a la arquitectura paralela le corresponde un tipo diferente de operación: hay un tiempo de “relajación” de idas y venidas de señales hasta que todas se estabilizan [*EN:settle down*] en una actividad coherente. Por lo que el ejercicio cooperativo conjunto requiere cierto tiempo para culminar [o finalizar *EN:culimante*], y esto es evidente en tanto que

¹¹no vienen en el diccionario

¹²to go against the grain está traducido como ir cuesta arriba —en el diccionario

conductualmente todos los animales muestran un lapso temporal natural. En el cerebro humano esta ráfaga de cooperación suele durar unos 200-500 mseg, la “ahoridad” [EN: “nowness”] de una unidad percepto-motora. Contrariamente a lo que pueda parecer a primera vista, tanto etológicamente como en nuestra propia introspección, la vida cognitiva no es un flujo constante, sino que aparece puntuado [EN: *punctuated*] por patrones conductuales que surgen y amainan [EN: *subside*¹³] en segmentos temporales [EN: *chunks of time*]. Esta revelación [EN: *insight*] de la neurociencia reciente —y de hecho de las ciencias cognitivas en general— es fundamental ya que nos alivia de la tiranía de la búsqueda de una cualidad homuncular centralizada de la conducta normal de un agente cognitivo.

Permítanme retroceder [EN: *backtrack*] un momento y recontextualizar [EN: *reframe*] nuestra discusión sobre el yo cognitivo junto con aquella en torno a un yo [EN: *self*] molecular mínimo. Estoy defendiendo que las neurociencias contemporáneas —como la biología celular en el caso de la organización de lo vivo— proveen elementos suficientes para concebir la organización básica de un yo cognitivo en términos de la *clausura operacional* (¡que no interactiva!) del sistema nervioso (Maturana and Varela, 1980; Varela, 1979). Hablo de “clausura” para subrayar la cualidad autorreferencial de la red interneuronal y de las superficies percepto-motoras cuyas correlaciones sostiene. El calificativo de “operacional” enfatiza que el término de clausura es utilizado en su sentido matemático de recursividad y *no* en su sentido de clausuramiento [EN: *closedness*] o aislamiento frente a interacciones, lo cual sería, por supuesto, absurdo. Más concretamente, el sistema nervioso está organizado por la clausura operacional de una red de sub-redes modulares recíprocamente relacionadas que dan lugar a conjuntos de actividad coherente de tal forma que:

- (i) median continuamente patrones invariantes de correlaciones sesomotoras de las superficies sensoras y efectoras;
- (ii) dan lugar a una conducta para el organismo en su conjunto [EN: *for the total organism*] como unidad móvil en el espacio.

La clausura operacional del sistema nervioso

¹³SUBSIDE: bajar, hundirse, amainar, hacerse menos violento, disminuir, alejarse, calmarse

hace surgir así un *modo* específico de coherencia, que se encuentra imbricado [EN: *embedded*] en el organismo. Esta coherencia es un *yo cognitivo* [EN: *cognitive self*]: una unidad de percepción/movimiento en el espacio, invariantes sensoromotoras mediadas a través de la red interneuronal. El tránsito [EN: *passage*] a la cognición sucede al nivel de la entidad conductual, y no, como en el caso del yo celular básico, como una entidad espacialmente limitada [EN: *not as a spatially bounded entity*]. La clave [EN: *the key*] en este proceso cognitivo es el sistema nervioso a través de su neuro-lógica. En otras palabras el yo cognitivo es la manera en la que un organismo, a través de su actividad auto-producida, se convierte en unidad distinguida en el espacio, pero siempre acoplada a su correspondiente entorno del que sin embargo se mantiene distinguido [EN: *distinct*]. Un yo coherente distinguido que, a través del mismo proceso de constitución de sí mismo, configura un mundo externo de percepción y acción.

3.2. Yo cognitivo y mundo perceptivo

La naturaleza de la *identidad* del yo cognitivo que acabamos de discutir es, igual que la del yo celular básico, una naturaleza *emergente* a través de un proceso distribuido. Las propiedades emergentes de una red interneuronal son, sin embargo, bastante diferentes en sus propiedades y posiblemente sean muchísimo más ricas en posibilidades. Lo que me gustaría subrayar aquí son los descubrimientos recientes sobre la facilidad con la que muchos agentes simples con propiedades simples pueden ser agrupados, incluso en de una forma azarosa, para *dar lugar* a lo que se le aparece a un observador externo como un todo integrado y con propósitos [EN: *purposes*], *sin* la necesidad de una supervisión central. Ya hemos tocado este punto al discutir la naturaleza de un proceso autopoiético y el modelo de autómeta celular [sic!], y más tarde al discutir el surgimiento constante y relajación [EN: *subside*] del complejo [EN: *ensemble*] neuronal que sostiene [EN: *underlying*] la conducta. Este tema de la propiedades emergentes es crucial para el conjunto de mi argumentación aquí, si bien sostengo mis conclusiones en estudios contemporáneos de varios sistemas complejos de inspiración biológica (Farmer et al., 1986; Langton, 1989b).

Lo que resulta particularmente importante es que podamos admitir que (i) un sistema pueda estar formado por componentes locales separados que (ii) sin embargo no tienen un yo central o localizado, y que aún así el todo se comporte como una unidad y que se le aparezca al observador como si hubiera un “agente” virtual coordinándolo todo desde el centro. Esto es lo que quiero decir al referirme a un yo *sin centro* —también podríamos postular un yo virtual: un patrón global coherente que emerge a través de componentes locales simples, dando la apariencia de tener una localización [EN:location] central cuando tal localización no puede ser encontrada, y aún así esencial como nivel de interacción para la conducta de la unidad como un todo [EN:whole unity].

La importación [EN:import] de tales modelos, formalismos y casos de estudio de sistemas complejos (i.e. propiedades emergentes a través de elementos simples coordinados) es, desde mi punto de vista [EN:in my eyes], bastante profunda [EN:profound] para nuestra comprensión de las propiedades cognitivas. Introduce en el estudio de las propiedades cognitivas una alternativa explícita a la tradición dominante computacional/cognitivista para la que la idea central es la de una sintaxis independiente de la materialidad, que puede sostener [EN:support] una semántica para un entorno. Esto se está convirtiendo cada vez más evidente para los investigadores de sistemas cognitivos artificiales, tal como las actuales escuelas conexionistas han dejado ya claro. Lo que encontramos en los cerebros es un promiscuo remiendo [EN:promiscuous tinkering] de redes y sub-redes que no da lugar a ninguna evidencia en favor de una estructura de descomposición de arriba a abajo [EN:from top to bottom] típica de los algoritmos informáticos [EN:computer algorithms¹⁴]. Del mismo modo, uno de los primeros mensajes que surgen del estudio de redes neuronales en términos conexionistas modernos es la ausencia de una distinción primaria [EN:principled¹⁵] entre software y hardware, o, para ser más precisos entre símbolos y no-símbolos. De hecho, *todo* lo que encontramos en las máquinas modernas de redes neuronales artificiales son actividades relativas entre conjuntos [o complejos EN:ensembles] que sostienen [EN:underlying] las regularidades que deno-

minamos su conducta o acción [EN:performance]. Puede que observemos que algunos de estos conjuntos [o complejos EN:ensemble] se repiten con suficiente regularidad como para describirlos como siendo programas, pero esto es otro tema. A pesar de estar artificialmente contruidos,

estos complejos [EN:ensembles] emergentes no se pueden llamar “computaciones” en el sentido en que sus dinámicas no pueden ser formalmente especificables como implementaciones de algún algoritmo de alto nivel. Las redes neuronales *no* pueden ser interpretadas como lenguaje máquina nisqueiera en los detalles más finos, ya que simplemente no hay transición entre semejantes átomos operacionales elementales con semántica y el nivel superior emergente

en el que ocurre la conducta. Si lo fueran, el clásico juicio computacionalista [o informático EN:computer wisdom] se aplicaría inmediatamente: ignorar el hardware ya que no añade nada significativo a la propia computación (aparte de constricciones de tiempo y espacio). Por el contrario [EN:in contrast] en los modelos de red, distribuidos, son precisamente estos “detalles” los que hacen posible un efecto global, y los que marcan una afilada ruptura con la tradición en IA. Naturalmente esto refuerza la conclusión paralela que es aplicable a las redes neuronales naturales

en el cerebro, tal como discutimos anteriormente.

He sacado este tema [EN:I have raised this point] para prevenir al lector frente a la fuerza de muchos años de dominio del computacionalismo, y la consiguiente tendencia a identificar el yo cognitivo con algún programa informático [EN:computer program] o descripciones computacionales de alto nivel. Esto no funcionará. El yo cognitivo *es* su misma implementación: su historia y su actuar son la misma pieza [EN:its history and its action are of one piece]. Ahora bien, esto exige que clarifiquemos ahora el segundo aspecto del yo que nos queda por tratar: su forma de relacionarse con el entorno.

3.3. Intencionalidad y neuro-lógica

La vida cotidiana [EN:ordinary] es necesariamente la de agentes *situados*, continuamente habiéndoselas con qué hacer frente a continuas actividades paralelas en marcha en varios de sus sistemas percepto-motores. Esta redefinición constante del qué hacer no es, en absoluto, como un plan,

¹⁴No es lo mismo computational algorithm que computer algorithm, pero no sé cómo traducirlo más apropiadamente

¹⁵no viene en el diccionario

almacenado en un repertorio de alternativas potenciales, sino que es enormemente dependiente de contingencias e improvisación, y más flexible que una planificación. Situacionismo [*EN:situatedness*] significa que una entidad cognitiva tiene —por definición— una perspectiva. Esto quiere decir que no está relacionada con su entorno “objetivamente”, es decir independientemente de la localización del sistema, encaramiento [*EN:heading*¹⁶], actitudes [*EN:attitudes*] e historia. Por el contrario, se refiere a él [(al entorno)] en relación a la perspectiva establecida por la emergencia constante de propiedades del agente mismo y en términos del rol que semejante redefinición en marcha [*EN:running*] juega en la coherencia completa [*EN:entire*] del organismo.

Una vez más, tal como hicimos para el yo celular mínimo, tenemos que diferenciar con precisión [*EN:sharply*] entre entorno y mundo. Y, una vez más, el modo de acoplamiento es doble. Por un lado, semejante cuerpo-en-el-espacio sucede evidentemente a través de las interacciones con el entorno del que depende. La naturaleza de estas interacciones es la de encuentros macrofísicos —transducción sensora [*EN:sensory transduction*], fuerza y efectividad [*EN:performance*] muscular, luz y radiaciones, etc. — nada hay de sorprendente en ello. Sin embargo este acoplamiento es posible sólo si los encuentros son abarcados [*EN:embraced*] desde la perspectiva del sistema mismo. Esto equivale, en concreto [*EN:quite specifically*], a la elaboración de un excedente de significación [*EN:a surplus signification*] relativo a tal perspectiva. Cualquier cosa que sea encontrada [*EN:Whatever is encountered*] debe ser valorada de una manera o de otra —gusto, disgusto, indiferencia— y se debe actuar sobre ella de un modo u otro —atracción, rechazo, neutralidad. Esta evaluación básica es inseparable del modo en que el evento de acoplamiento es un encuentro [*EN:encounters*] con una unidad percepto-motora en funcionamiento,¹⁷ y da lugar a una intención (estoy tentado a decir “deseo”), esa cualidad única de la cognición viva (Dennett, 1987).

¹⁶supongo que Varela quería decir encaramiento. La palabra en inglés (heading) significa encabezamiento, membrete, sección, apartado... pero nada que encaje con el contexto presente.

¹⁷En inglés en el original: *This basic assessment is inseparable from the way in which the coupling event encounters a functioning perceptuo-motor unit*

Reformulado en otros términos, la naturaleza del entorno para un yo cognitivo adquiere un status curioso: es aquello que *se presta*

(*es lehnt sich an...*) a un excedente de significatividad [*EN:significance*]. Como la improvisación de jazz, el entorno provee una “excusa” para la “música” neuronal desde la perspectiva del sistema cognitivo implicado [*EN:involved*]. Al mismo tiempo, el organismo no puede vivir sin este acoplamiento constante y sin las regularidades que emergen constantemente; sin la posibilidad de una actividad acoplada el sistema se convertiría en un mero fantasma solipsista.

Sin ir más lejos [*EN:for instance*], la luz y la reflectancia (entre otros muchos parámetros macrofísicos como las texturas y los contornos, pero simplifiquemos por el bien de la argumentación [*EN:for the arguments sake*]), se prestan a una amplia variedad de espacios de color, dependiendo del sistema nervioso implicado en el encuentro. Durante sus respectivos recorridos evolutivos, peces teleost [*EN:teleost*¹⁸ *fishes*], pájaros, mamíferos e insectos han dado lugar [*EN:brought forth*] a varios espacios de color diferentes, no sólo con distinta significatividad conductual sino con dimensionalidad diferente, de tal manera que no es un asunto [*EN:matter*] de más o menos resolución de color (Thompson et al., 1992). Se puede demostrar que el color no es una propiedad que deba ser “recuperada” de la “información” del entorno de alguna forma única. El color es una dimensión que se muestra [*EN:that shows up*] sólo en el diálogo filogenético entre un entorno y la historia de un yo [*EN:self*] autónomo activo que define parcialmente lo que cuenta como entorno. La luz y la reflectancia proveen [*EN:provide*] un modo de acoplamiento, una perturbación que gatilla [*EN:triggers*], que ofrece una ocasión para la enorme capacidad informativa de las redes neuronales para constituir correlaciones senso-motoras y por tanto para poner en acción su capacidad de imaginar y presentar [*EN:to put into action their capacity for imagining an presenting*]. Es sólo después de que todo esto haya sucedido, después de que el acoplamiento se convierta en regular y repetitivo, como el caso de los colores en nuestro mundo —y el de otros—, que nosotros, observadores, por facilidad del lenguaje, decimos que el color corresponde a o representa un aspecto del mundo.

¹⁸No aparece en mi diccionario

Un reciente ejemplo dramático de este excedente de significatividad y la titubeante actuación [*EN:dazzling performance*] del cerebro como generador de “narrativas” neuronales lo dan las tecnologías de las así llamadas “realidades virtuales”. Los movimientos y la percepción visual dan lugar, así, a regularidades que son propias de esta nueva forma de acoplamiento senso-motor. Lo que más significativo me resulta en este punto es la *veracidad* [*EN:veracity*] del mundo que brota súbitamente [*EN:which rapidly springs forth*]: *habitamos* un cuerpo dentro de este nuevo mundo después de un corto periodo de tiempo probando esta nueva situación (i.e. unos 15 minutos), y la experiencia es verdaderamente la de volar a través de paredes, o la de cavar en universos fractales. Y esto sucede a pesar de la pobre calidad de la imagen, la baja resolución de los sensores y la cantidad limitada de interconexión [*EN:interlinking*] entre superficies sensoras y superficies de imágenes generadas por un programa que corre en un ordenador personal. A través de su clausura, el sistema nervioso es un sintetizador de regularidades de tal virtuosidad que cualquier material básico basta como entorno para hacer surgir [*EN:bring forth*] un mundo convincente [*EN:compelling*¹⁹].

Esta misma estrategia del situacionismo [*EN:situatedness*] de un agente que se encuentra progresivamente dotado con cada vez más sofisticados [*EN:richer*] módulos internos auto-organizativos, se está convirtiendo en un programa de investigación productivo incluso para el tan pragmáticamente orientado campo de la inteligencia artificial. Citando, con cierta extensión a Rodney Brooks, uno de los mayores exponentes de esta tendencia:

... defendiendo [*EN:argue for*] otra forma de crear inteligencia artificial:

- Tenemos que ir construyendo las capacidades de los sistemas inteligentes progresivamente a cada paso del camino y asegurar por tanto que las piezas y sus interfaces son válidas.
- A cada paso debemos construir sistemas inteligentes completos que podamos dejar sueltos [*EN:let loose*] en el mundo real sintiendo y actuando de

forma real. Algo menos se convierte en candidato con el que engañarnos a nosotros mismos.

Hemos estado siguiendo este enfoque y hemos construido una serie de robots autónomos mínimos. Hemos llegado a una conclusión inesperada (**C**) y tenemos una hipótesis (**H**) bastante [*EN:rather*] radical:

C : Al examinar la inteligencia de un nivel muy simple encontramos que representaciones explícitas y modelos del mundo están de sobra [*EN:simply get in the way*²⁰]. Resulta que es mejor utilizar el mundo como su propio modelo [o un modelo de sí mismo *EN: as its own model*].

H : La representación es la unidad de abstracción errónea a la ahora de construir el grueso de un sistema inteligente.

La representación ha sido el tema central del trabajo en Inteligencia Artificial durante los últimos 15 años sólo porque proporcionaba un interfaz entre, lo que de otro modo, no serían más que asignaturas y artículos de conferencias aislados.

Brooks (1987, p. 1)

Cuando la síntesis de conducta inteligente es abordada de esta forma incremental, con estricta observancia [*EN:adherence*] de la viabilidad senso-motora de un agente, la noción de que el mundo es una fuente de información que debe ser representada simplemente desaparece. El foco de la cuestión se centra en la autonomía del yo cognitivo. Por lo tanto en la propuesta de Brooks sus criaturas minimalistas [*EN:minimal creatures*] unen [*EN:join together*] varias actividades a través de una regla de convivencia [*EN:cohabitation*] entre ellas. Esto es homólogo a un recorrido [*EN:pathway*] evolutivo a través del cual sub-redes modulares se entrelazaron entre sí en el cerebro. El resultado esperado son aparatos inteligentes autónomos más auténticamente inteligentes [*EN:more truly intelligent*], en

²⁰lo más parecido que he encontrado en el diccionario es GET INTO THE WAY: adquirir la costumbre. Pero no encaja nada bien.

¹⁹COMPELLING: irresistible, apremiante, convincente.

lugar de frágiles procesadores de información que dependen de un entorno pre-asignado o un plan óptimo.

Es interesante hacer notar [EN:to note] que en este artículo Brooks también traza el origen de lo que describe como “decepción de la IA” hasta la tendencia en IA (y en el resto de las ciencias cognitivas) a la abstracción, i.e., a la desfactorización [EN:for factoring out²¹] de habilidades motoras y perceptivas situadas. Tal y como he discutido aquí (y Brooks discute por sus propias razones), semejante abstracción pierde la esencia de la inteligencia cognitiva, *que reside sólo en su corporalización*. Es como si uno pudiera separar los problemas cognitivos en dos partes: aquellos que pueden resolverse por abstracción y los que no. La segunda parte la constituyen típicamente la percepción-acción y habilidades motoras de los agentes en entornos inespecíficos. Al ser tratado desde esta perspectiva auto-situada [EN:When approached from this self-situated perspective] no hay destino en el que la percepción pudiera entregar una representación del mundo en el sentido tradicional. El mundo se muestra a través de la enacción²² de regularidades percepto-motoras. “Del mismo modo que no hay representación central tampoco hay sistema central. Cada capa de actividad [EN:activity layer] conecta la percepción a la acción directamente. Es solamente el observador de la Criatura quien introduce [EN:imputes] una representación o control central. La criatura en sí misma no tiene ninguno: es una colección de conductas en competición. Del caos local de sus interacciones emerge, a ojos del observador, un patrón coherente de conducta” (Brooks, 1986, p. 11).

Para concluir, las dos ideas claves [EN:main points] que he estado intentando subrayar en esta sección dedicada al yo cognitivo son las siguientes²³. Primero, he intentado explicar de forma sencilla [EN:spell out] la naturaleza de su *identidad* como cuerpo en movimiento-y-espacio [EN:motion-and-space] a través de la clausura operacional de la red interneuronal. Esta actividad puede ser ob-

servada como múltiples sub-redes, actuando en paralelo y entrelazadas en complejos *bricolages*, dando lugar una y otra vez a patrones coherentes que se manifiestan como conductas. Segundo, he intentado clarificar cómo esta dinámica emergente, paralela y distribuida es inseparable de la *constitución de un mundo*, que no es otro que el excedente de significado [EN:meaning] e intenciones que lleva aparejada la conducta situada [EN:intentions carried by situated behaviour]. Si las conexiones [EN:links] al entorno físico son inevitables, la singularidad [EN:uniqueness] del yo cognitivo es esta constante génesis de significado. O, por invertir la descripción una vez más, la singularidad del yo cognitivo es la constitutiva *falta* de significación [EN:signification] que debe ser suplida frente a las permanentes perturbaciones y rupturas [EN:breakdowns] de la continua vida percepto-motora. La cognición es acción sobre lo que *falta*, rellenando la ausencia [EN:the fault] desde la perspectiva de un yo cognitivo.

Esta perspectiva equivale a una biología de la intencionalidad. De hecho, responde sin ambigüedad a dos preguntas clave: los problemas de enraizamiento del símbolo (Harnad, 1991) y de la sintaxis (Searle, 1990) (*the symbol and syntax grounding problems*). El primero se refiere al misterio del origen del significado de los símbolos naturales, ya que en la opción clásica cognitivista hay una necesidad intrínseca de una asignación semántica arbitraria. La respuesta que provee nuestra perspectiva [EN:The answer provided by this approach] es que el significado surge en la emergencia de un punto de vista propio de la constitución autónoma del organismo en todos sus niveles, empezando por el nivel autopoietico. El problema del enraizamiento de la sintaxis dice que todas las operaciones sintácticas en un sistema simbólico son dependientes del observador. Nuestra respuesta es, precisamente, que la constitución de una unidad autónoma provee [EN:provides] los medios para la aparición de regularidades que son la base de la composicionalidad. Esto puede manifestarse al nivel celular al modo del famoso [EN:celebrated] código genético para la síntesis de proteínas, o al nivel del cerebro en las propiedades composicionales de los complejos neuronales [EN:ensemble]. No hay nada misterioso en la emergencia de semejantes regularidades suje-

²¹Creo que la expresión es utilizada como metáfora matemática: factoring out puede que sea la extracción del máximo común divisor (*highest common factor*)

²²Explicar neologismo vareliano

²³La frase en el original no tiene mucho que ver pero es horrible: “To conclude, the two main points that I have been trying to bring into full view in this Section devoted to the cognitive self are as follows

tas a composición [*EN:composable*²⁴]. Por tanto al contrario que en la mayor parte del debate filosófico actual (sea este el de Searle, Harnad o Dennett) no necesitamos tener una asignación arbitraria y dependiente del observador, ni del significado y ni de la composicionalidad. La clave está en la propiedades identitarias [*EN:identity properties*] generadas por la auto-constitución del organismo.

4. La doble dialéctica del organismo

El organismo, por tanto, es la clave central para las ciencias cognitivas, y no puede ser abordado en un solo proceso. Estamos forzados a descubrir “regiones” que se entretujan de formas complejas, y, en el caso de los humanos, se extienden, más allá de los estrictos confines del cuerpo, en el registro socio-lingüístico.

Más aún, lo que vengo argumentando es que detrás de este desorden de varios yoes [*EN:selves*] que llevamos encima [*EN:carry around*], todos estos yoes comparten una lógica común y fundamental al tiempo que difieren en su especificidad. Esto es un caso de lo que Wittgenstein hubiera llamado “aires de familia”: en lugar de que todas las características sean comunes a todas las instancias [*EN:instances*]; tratamos con un aglomerado [*EN:cluster*] de características superpuestas. También podemos hablar de este aglomerado [*EN:cluster*] de características comunes como de una *dialéctica* compartida, ya que estamos tratando aquí con un proceso de dos caras, en el que la co-definición está en el núcleo del asunto. De hecho, propongo que la dialéctica organizacional del yo es un acontecimiento de dos niveles: Tenemos por un lado la dialéctica de identidad del yo; por otra lado la dialéctica por la que esta identidad, una vez establecida, hace surgir [*EN:brings forth*] un mundo de un entorno. La identidad y el conocimiento se encuentran en relación mutua al modo de las dos caras de un mismo proceso: esto forma el núcleo de la dialéctica de todos los yoes.

En primer lugar una *dialéctica de la identidad* establece una agente autónoma, un para-sí (*pour soi*). Esta identidad se establece a través de un entrelazamiento [*EN:bootstrapping*] de dos términos:

- (i) un término *dinámico* que se refiere a un conjunto [o complejo, ensamblaje *EN:ensemble*] de componentes en una red de interacciones capaces de tener propiedades emergentes: redes metabólicas, complejos [*EN:assemblies*] neuronales, redes clónicas de anticuerpos [*EN:clonal antibody networks*], recursividad lingüística;
- (ii) un término *global* que se refiere a las propiedades emergentes, una totalidad que condiciona (hacia abajo [*EN:downwardly*]) los componentes de la red: membranas celulares, el cuerpo senso-motor en el espacio, discriminación yo/no-yo, ‘yo’ personal [*EN:personal ‘I’*].

Estos dos términos están realmente en relación de codefinición. Por un lado el nivel global no puede existir sin el nivel de red ya que surge de él [*EN:it comes forth through it*]. Por otro lado el nivel dinámico no puede no existir [*EN:cannot not exist* [sic!]] y operar como tal sin ser contenido y alojado [*EN:lodged*] por una unidad abarcadora [*EN:encompassing*] que la hace posible.

En segundo lugar, una *dialéctica del conocimiento* establece un mundo de significatividad cognitiva *para* esta identidad. Esto sólo puede surgir desde la perspectiva proporcionada por esta identidad, que añade un excedente de significatividad [*EN:significance*] a las interacciones con el ambiente propias de las partes constitutivas.

La idea clave es, entonces, que el organismo hace aparecer [*EN:brings forth*] y especifica su propio dominio de problemas y acciones que deben ser “solucionados”; este dominio cognitivo no existe “ahí fuera” en un entorno que actúa como campo de aterrizaje para un organismo que de alguna manera salta o es lanzado en paracaídas en el mundo. En su lugar, los seres vivos y sus mundos de significado se encuentran en relación de *especificación mutua* o *co-determinación*. Por tanto lo que describimos como regularidades significativas del entorno no son características [*EN:features*] externas que han sido internalizadas, tal como asume la tradición representacionista dominante en ciencias cognitivas —y la adaptacionista en biología evolutiva. Las regularidades del entorno son el resultado de una historia conjunta [*EN:conjoin*], una congruencia que se despliega [*EN:unfolds*] desde una larga historia de co-determinación. En palabras de Lewontin (1983), el organismo es tanto el sujeto como el objeto de la

²⁴No viene en el diccionario pero quizás hubiera que traducir por composicionables

evolución.

Éste segundo nivel de la dialéctica del organismo es así también establecido a través del entrelazamiento [*EN:bootstrapping*] de dos términos:

- (i) un término de *significatividad* que se refiere a la emergencia necesaria de un excedente de significado propio de la perspectiva del yo constitutivo: semántica celular, percepción y acción conductual, yo/no-yo como afirmación [*EN:assertion*] somática, identidad personal.
- (ii) un término de *acoplamiento* que se refiere a la dependencia e imbricación necesaria y permanente del yo sobre su entorno, ya que sólo a través de ese acoplamiento puede surgir su mundo [*EN:can its world be brought forth*]: leyes físico-químicas para el mundo celular, propiedades físicas macroscópicas para la conducta cognitiva, interacción molecular para el yo inmunológico [*EN:immune self*], intercambios socio-lingüísticos para nuestros yoes subjetivos [*EN:subjective selves*].

Doble dialéctica: la naturaleza de una identidad y la naturaleza de una relación con un mundo. Doble paralogismo [*EN:paradoxicality*]: Auto-producción por contención dependiente [*EN:dependent containment*]; autonomía de conocimiento por acoplamiento con el entorno. Ambas dialécticas dan lugar a la alternante naturaleza del organismo, ineluctablemente formándose a sí mismo e informando donde está, e igualmente ineluctablemente implicado en el paisaje [*EN:background*] del que brota [*EN:springs forth*]. Organismos, esas fascinantes mezclas de yoes sin centro [*EN:selfless selves*], ni más ni menos que existencias circulares multinivel, abiertas al cambio [*EN:open-ended*], siempre conducidas por la falta de significatividad que generan al afirmar su presencia.

Agradecimientos

Reconozco agradecidamente el apoyo económico de la CNRS, la Fondation de France (Chaire Scientifique), y el Prince Trust Fund.

Referencias

- Abeles, M. (1984). *Local Cortical Circuits*. Springer Verlag, Berlin.
- Agree, P. (1988). The dynamic structures of everyday life. Report Number AI-TR 1085, MIT Artificial Intelligence Lab., Cambridge.
- Allman, J., Meizen, F., and McGuinness, E. (1985). Non-classical receptive field properties. *Annual Review of Neuroscience*, 8:407–430.
- Baeza, I., Ibañez, M., Lazcano, A., Santiago, C., Arguello, C., Wong, C., and Oró, J. (1987). Liposomes with polyribonucleotides as models of precellular systems. *Origins of Life*, (17):187–199.
- Bonner, J. T. (1988). *The Evolution of Complexity*. Princeton University Press, Princeton.
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, (47):139–161.
- Brooks, R. A. (1986). Achieving artificial intelligence through building robots. A.I. Memo 899, Massachusetts Institute of Technology, Artificial Intelligence Laboratory.
- Brooks, R. A. (1987). Intelligence without representation. Mit artificial intelligence report, Cambridge, MA.
- Buss, L. (1987). *The Evolution of Individuality*. Princeton University Press, Princeton.
- Castoriadis, C. (1987). L'état du sujet aujourd'hui. *Topique*, 38:7–39.
- Deamer, D. (1986). Role of amphiphilic compounds in the evolution of membrane structure on the early earth. *Origins of Life*, (17):3–25.
- Deamer, D. and Barchfeld, G. (1982). Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. *Journal of Molecular Evolution*, (18):203–206.
- Dennett, D. C. (1987). *The Intentional Stance*. The MIT Press, Cambridge, Massachusetts.
- Farmer, J., Lapedes, A., Packard, N., and Wenderoff, B., editors (1986). *Evolution, Games and Learning*. North-Holland, Amsterdam.

- Fleischaker, G. (1988). *Autopoiesis: System logic and the origin of life*. PhD thesis, Boston University, Boston, MA.
- Harnad, S. (1991). The symbol grounding problem. In Forrest, S., editor, *Emergent Computation*. MIT Press.
- John, E., Tang, Y., Brill, A., Young, R., and Ono, K. (1986). Double-labeled metabolic maps of memory. *Science*, 233:1167–1175.
- Langton, C. (1989a). Artificial life. In Langton, C., editor, *Artificial Life*, volume VI of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 1–47, Redwood City, California. Addison-Wesley Publishing Company, Inc.
- Langton, C. G., editor (1989b). *Artificial Life*, volume VI of *Series: Santa Fe Institute Studies in the Sciences of Complexity*. Addison-Wesley Publishing Company, Inc., Redwood City, California. Proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems held September, 1987, in Los Alamos, New Mexico.
- Lazcano, A. (1986). Prebiotic evolution and the origin of cells. *Treballs Societat Catal. Biol.*, (47):185–250.
- Lewontin, R. (1983). The organism as the subject and object of evolution. *Scientia*, 118:63–82.
- Luisi, L. and Varela, F. (1989). Self replicating micelles: A minimal version of a chemical autopoietic system. *Origins of Life*, (19).
- Margulis, L. (1981). *Symbiosis in Cell Evolution*. W.H. Freeman, San Francisco.
- Margulis, L. and Schwartz, K. (1988). *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth*. W. H. Freeman, New York.
- Maturana, H. and Varela, F. (1980). *Autopoiesis and Cognition: The Realization of the Living*, volume 42 of *Boston Studies in the Philosophy of Science*. D. Reidel Publishing Company, Dordrecht, Holland.
- Minsky, M. (1987). *The Society of Mind*. Simon and Schuster, New York.
- Searle, J. (1990). Is the brain a digital computer? *Proceedings of the American Philosophical Association*, 64:21–37.
- Thompson, E., Palacios, A., and Varela, F. (1992). Ways of coloring: Comparative color vision as a case study in cognitive science. *Behavioral and Brain Sciences*, 15:1–75.
- Varela, F. (1991). Organism: A meshwork of selfless selves. In Tauber, editor, *Organism and the Origin of Self*, pages 79–107. Kluwer Assoc., Dordrecht.
- Varela, F. and Bourgine, P., editors (1992). *Towards a Practice of Autonomous Systems: Proceedings of the First European conference on Artificial Life*. MIT Press.
- Varela, F., Maturana, H., and Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, (5):187–196.
- Varela, F. J. (1979). *Principles of Biological Autonomy*. North-Holland, New York.
- Varela, F. J. and Coutinho, A. (1991). Second generation immune networks. *Immunology Today*, 12:159–167.
- Varela, F. J., Dupire, B., and Coutinho, A. (1988). Cognitive networks: Immune, neural and otherwise. In Perelson, A., editor, *Theoretical Immunology*, volume 2 of *SFI Series on Complexity*, pages 359–375. Addison Wesley, New Jersey.
- Zecevic, D., Wu, J., Cohen, L., London, J., Höpp, H., and Falk, C. (1989). Hundreds of neurons in the aplysia abdominal ganglion are active during the gill-withdrawal reflex. *Journal of Neuroscience*, 9:3681–3689.