



Antroporrobótica, robótica de décima geração e habitar humano: uma conversa com Humberto Maturana¹

Ignacio Muñoz-Cristi
Instituto Humberto Maturana

Resumo

As reflexões sobre robótica e cognição costumam ser formuladas dentro de paradigmas representacionistas que obscurecem os processos pelos quais os sistemas operam. Esta conversa explora uma perspectiva alternativa, buscando pensar e projetar máquinas sem recorrer a noções semânticas que mascaram sua dinâmica de funcionamento, e enfatizando, em vez disso, a distinção entre os domínios de existência dos diferentes tipos de sistemas. Na entrevista, Humberto Maturana reflete sobre a possibilidade de conceber robôs capazes de linguajar e emocionar de tal modo que um observador possa distinguir neles um operar autoconsciente, e que operem com plena autonomia. O diálogo delinea uma forma de metadesenho que articula princípios operativos com critérios éticos e ontoepistemológicos, entre eles a objetividade entre parênteses e o amar entendido como aceitação do outro. No prefácio que acompanha a entrevista, situo essa conversa no arco histórico da robótica e da civilização tecnológica moderna, e introduzo a antroporrobótica como um campo transdisciplinar aqui proposto, que oferece chaves tanto para o projeto de sistemas artificiais quanto para a reflexão sobre nossa relação com o robótico.

Palavras-chave

Antroporrobótica. Robótica de Décima Geração. Humberto Maturana. Biologia do Conhecer. Crise Civilizacional.

1. Esta entrevista foi publicada originalmente em inglês em 2025(a), sob o título: “*Step Toward Anthroporobotics: A Conversation with Humberto Maturana*”, na revista *Constructivist Foundations*, v. 21, n. 1. A presente versão inclui uma nova introdução e um enquadramento distinto.

Antroporrobótica, robótica de décima generación y habitar humano: una conversación con Humberto Maturana

Resumen: Las reflexiones sobre robótica y cognición suelen formularse dentro de paradigmas representacionalistas que oscurecen los procesos mediante los cuales operan los sistemas. Esta conversación explora una perspectiva alternativa, buscando pensar y diseñar máquinas sin recurrir a nociones semánticas que enmascaran su dinámica de funcionamiento y enfatizando, en cambio, la distinción entre los dominios de existencia de distintos tipos de sistemas. En la entrevista, Humberto Maturana reflexiona sobre la posibilidad de concebir robots capaces de lenguajear y emocionar de tal modo que un observador pueda distinguir en ellos un operar autoconsciente, y que operen con plena autonomía. El diálogo delinea una forma de metadiseño que articula principios operativos con criterios éticos y onto-epistemológicos, entre ellos la objetividad entre paréntesis y el amar entendido como aceptación del otro. En el prefacio que acompaña la entrevista, situo esta conversación en el arco histórico de la robótica y de la civilización tecnológica moderna, e introduzco la antroporrobótica como un campo transdisciplinario aquí propuesto que ofrece claves tanto para el diseño de sistemas artificiales como para la reflexión sobre nuestra relación con lo robótico.

Palabras clave: Antroporrobótica. Robótica de Décima Generación. Humberto Maturana. Biología del Conocer. Crisis Civilizacional.

Anthroporobotics, tenth-generation robotics and human dwelling: a conversation with Humberto Maturana

Abstract: Reflections on robotics and cognition are often formulated within representationalist paradigms that obscure the processes through which systems operate. This conversation explores an alternative perspective, seeking to think about and design machines without relying on semantic notions that mask their operational dynamics, and instead emphasizing the distinction between the domains of existence of different types of systems. In the interview, Humberto Maturana reflects on the possibility of conceiving robots capable of languaging and emotioning in such a way that an observer could distinguish in them a self-conscious operation, while operating with full autonomy. The dialogue outlines a form of metadesign that articulates operational principles with ethical and onto-epistemological criteria, including objectivity in parentheses and loving understood as the acceptance of the other. In the preface accompanying the interview, I situate this conversation within the historical arc of robotics and modern technological civilization, and introduce anthroporobotics as a transdisciplinary field proposed here, offering keys both for the design of artificial systems and for reflecting on our relationship with the robotic.

Keywords: Anthroporobotics. Tenth-Generation Robotics. Humberto Maturana. Biology of Cognition. Civilizational Crisis.

1 O presente como limiar: robótica, crise e bifurcação histórica

Nas primeiras décadas do século XXI, a robótica e a inteligência artificial deixaram de ser domínios técnicos especializados para se tornarem infraestruturas ativas do mundo contemporâneo. Já não se tratam apenas de máquinas que automatizam tarefas ou de algoritmos que otimizam processos, mas de sistemas que participam de maneira crescente e estrutural na organização do trabalho, do cuidado, da comunicação, da guerra, da educação e da vida cotidiana. A pergunta decisiva, portanto, já não é quão inteligentes as máquinas podem vir a ser, mas que tipo de mundo se configura quando o seu desenho e a sua implantação passam a integrar o próprio tecido do habitar humano.

Esse deslocamento não ocorre em um vazio histórico. Coincide com uma crise civilizatória profunda, marcada pela convergência de múltiplos processos de desarmonia: crise ecológica global, esgotamento dos modelos econômicos extrativos, intensificação das desigualdades sociais, precarização do trabalho e erosão dos vínculos comunitários. Nesse contexto, a expansão acelerada de tecnologias robóticas e algorítmicas não pode ser compreendida como simples progresso técnico, mas como um fenômeno inscrito em uma bifurcação histórica (Amin *et al.*, 1987), uma encruzilhada na qual se disputam distintos modos de reorganizar a convivência humana no planeta: um mais autoritário, desigual e polarizado, e outro mais democrático, equitativo e solidário (Wallerstein, 1998).

Sob essa perspectiva, a robótica contemporânea torna-se um espelho ampliado das tensões culturais do nosso tempo. Por um lado, aparece como promessa de eficiência, crescimento e controle; por outro, como ameaça de desemprego massivo, vigilância generalizada, maquinaria bélica e aprofundamento da dependência tecnológica. No entanto, ambas as leituras compartilham um mesmo pressuposto: tratam a tecnologia como um instrumento neutro cujo impacto dependeria apenas de seu uso posterior. Essa visão instrumental oculta uma questão mais radical: toda tecnologia encarna, desde o seu desenho, uma determinada antropologia – isto é, um certo entendimento do humano, do mundo e da relação entre ambos.

Na robótica e na inteligência artificial, essa compreensão costuma aparecer de forma implícita. As máquinas são concebidas, majoritariamente, sob a ideia de que conhecer equivale a processar informação, agir a executar instruções e o meio a um domínio externo que deve ser representado, modelado e controlado (Brooks, 1991). Essas premissas não constituem meras decisões técnicas, mas heranças culturais profundas, enraizadas na tradição moderno-colonial ocidentalocêntrica, que concebe o conhecimento como representação de uma realidade independente do observador e a natureza como objeto disponível. Quando esses pressupostos são trasladados ao desenho robótico, o resultado não é apenas uma ferramenta mais poderosa, mas a reiteração tecnológica de um modo histórico de habitar: o das civilizações de dominação (Maturana; Verden-Zöllner, 1993). A expansão recente de sistemas dialogantes, robótica social e dispositivos autônomos tornou visível um novo limiar. As máquinas já não operam apenas no pano de fundo da produção industrial, mas começam a ocupar espaços tradicionalmente reservados à interação humana: acompanham, orientam, recomendam, avaliam e dialogam. Nesse deslocamento surge uma pergunta inevitável: o que ocorre quando o conversar, o cuidar e o coordenar ações – dimensões centrais do habitar humano – se transformam em funções tecnológicas? A resposta não pode ser reduzida a métricas de desempenho nem a debates morais posteriores; exige revisar os fundamentos a partir dos quais essas máquinas são concebidas.

A noção de bifurcação adquire aqui seu sentido mais pleno. A robótica contemporânea não avança linearmente em direção a um futuro predeterminado, mas se desdobra em um campo de tensões onde coexistem trajetórias divergentes. Uma aprofunda a lógica dominante do capitalismo global: automatização orientada à maximização do capital acumulável, concentração do poder tecnológico e mercantilização da atenção, do cuidado e da palavra. Nessa trajetória, as máquinas não apenas substituem trabalho humano, mas também reproduzem e amplificam as formas de dominação que estruturam a ordem vigente de convivência.

Existe, contudo, outra possibilidade, mais frágil e menos visível, que não se deixa reduzir à oposição entre tecnofilia e tecnofobia. Ela consiste em repensar a robótica não como ferramenta de substituição, mas como âmbito de coabitação; não como extensão do controle, mas como exploração de novas formas de coordenação e convivência. Para que essa alternativa se torne pensável, é necessário interrogar as bases ontoepistêmicas e conceituais sobre as quais a robótica moderna foi construída, e abrir espaço para outras compreensões do conhecer, do agir e do viver, nas quais seja possível a harmonia entre antroposfera, biosfera e robosfera (Maturana, 1998). Deste ponto de vista, a pergunta pelo futuro da robótica é inseparável da pergunta pelo futuro da humanidade em sua relação com a biosfera e com as tecnologias que ela mesma engendra. O problema não é se as máquinas chegarão a ser mais “inteligentes”, mas que tipo de relações tornar-se-ão possíveis ou impossíveis. Serão dispositivos a serviço da exploração intensificada de pessoas e territórios, ou poderão integrar-se a tramas de convivência orientadas à regeneração da harmonia perdida entre o humano e o vivente? Consolidarão uma civilização baseada na apropriação, na hierarquia, no controle e na competição, ou contribuirão para imaginar e sustentar modos colaborativos de habitar o mundo?

Essas perguntas não admitem respostas simples nem soluções técnicas imediatas, mas delimitam o horizonte a partir do qual certas reflexões adquirem hoje uma relevância inesperada. Nesse contexto, torna-se possível voltar o olhar para formulações que, em seu momento, pareceram marginais ou prematuras, não para buscar nelas receitas aplicáveis, mas para reabrir distinções que permitam pensar de outro modo o vínculo entre máquinas, seres humanos e mundo. Nessa perspectiva, a robótica não remete apenas a uma engenharia de artefatos, mas ao que proponho chamar de ontotecnologia: uma prática de desenho que, ao estabilizar certas coordenações do fazer, conserva determinadas formas de mundo e exclui outras. Nesse sentido, a robótica de décima geração pode ser compreendida como parte de uma ontotecnologia constitutiva, disciplina emergente que dialoga com a antroporrobótica (Muñoz-Cristi, 2025b) na exploração de novas formas técnicas do habitar humano.

A entrevista que aqui se apresenta surge precisamente nesse interstício. Lida desde o presente, não aparece como um documento do passado, mas como um gesto que antecipa – sem pretendê-lo – muitas das tensões que hoje atravessam a robótica contemporânea. Antes de situá-la em seu contexto histórico específico, convém reconhecer o limiar em que nos encontramos: um tempo em que desenhar máquinas é, inevitavelmente, um ato político, cultural e ontoepistemológico que, para além do próprio desenho, exige também metadesenho e metametadesenho, ainda que nem sempre isso seja reconhecido como tal.

2 Metarreflexão, metadesign e metadomínios de fundamentação

Pensar a robótica de décima geração exige deslocar a atenção dos artefatos para os fundamentos que orientam sua criação. Essa noção designa a passagem de sistemas programados para executar instruções para sistemas autônomos capazes de conservar sua coerência estrutural em uma história de acoplamentos com o entorno.

Esse deslocamento não é primariamente técnico nem metodológico, mas reflexivo no sentido estrito do viver humano: um acontecer no qual se solta uma certeza que organizava o fazer, abrindo um novo domínio de ação e de sentido. Sem essa ruptura, não há transformação dos modos de projetar, mas apenas variações dentro de um mesmo marco (Maturana 1988). A reflexão não consiste em revisar ou descrever com maior precisão, mas em colocar em questão uma explicação ou convicção que sustentava o agir. Esse soltar não é intelectual, mas emocional-relacional: muda a emoção e, com isso, o domínio de ações possíveis. A reflexão não acrescenta informação; reorganiza o viver.

A metarreflexão aprofunda esse movimento ao tensionar o próprio marco de distinções a partir do qual algo fazia sentido. Já não se questiona apenas uma explicação, mas o modo habitual de observar e decidir. O observador reconhece então o caráter histórico e situado de suas distinções. A metametarreflexão leva essa ruptura ainda mais longe, ao tornar visível a cegueira constitutiva do próprio observar. Não conduz a uma síntese superior, mas a uma disposição distinta diante do conhecer: transitar na incerteza sem buscar clausura ou garantia última. Esses modos não constituem níveis nem etapas lineares. Entrelaçam-se recursivamente no viver e dependem da disposição para soltar certezas. Em todos os casos, transformam o domínio de ação e têm consequências diretas no fazer. A partir daí, o design tecnológico não pode ser separado da reflexão que o torna possível. Projetar não é apenas produzir objetos, mas configurar acoplamentos que conservam coerências no conviver. O design técnico define arquiteturas; o metadesign (Maturana, 1997) configura os critérios do projetar; e o metametadesign explicita o horizonte ético-relacional que orienta ambos, evitando que o fazer

derive para a dominação ou a negação do outro. Esses níveis também não são hierárquicos nem sequenciais. Operam recursivamente: mudanças no design tensionam seus critérios, e transformações nesses critérios reconfiguram o horizonte que os sustenta. Projetar, observar e refletir entrelaçam-se numa mesma dinâmica.

A ontotecnologia e a antroporrobótica assumem essa circularidade como princípio metodológico: a tecnologia não se transforma sem que se transforme o olhar que a concebe. Daí decorre uma consequência ética: não projetar aquilo que corrói a convivência que torna possível projetar. A ética não se acrescenta ao final; surge do próprio operar. Nesse marco, a robótica de décima geração não é apenas uma inovação técnica, mas uma possibilidade condicionada pela abertura reflexiva do conviver humano. Onde o design se desacopla desses movimentos, tende a reforçar lógicas de controle; onde essa recursividade se conserva, pode contribuir para formas de convivência que não reproduzam a violência nem a instrumentalização da relação.

3 Robótica moderna: genealogia de um limite histórico

Para compreender o lugar singular que ocupa a conversa apresentada a seguir, torna-se necessário situá-la na história longa da robótica e da inteligência artificial. Não para reconstruir uma cronologia técnica exaustiva, mas para identificar o substrato comum de seus desenvolvimentos e o limite histórico que eles não conseguiram ultrapassar. Somente assim é possível reconhecer quando uma proposta não constitui um aprimoramento incremental, mas um deslocamento do próprio marco em que o problema foi formulado. As origens da robótica moderna vinculam-se ao projeto cibernético de meados do século XX (Wiener, 1948; Pickering, 2010). Nesse contexto, as máquinas foram concebidas como sistemas capazes de regular seu comportamento por meio de circuitos de retroalimentação orientados à manutenção de estados de equilíbrio ou ao cumprimento de metas externas. A noção de controle ocupou um lugar central.

Embora essa abordagem tenha introduzido uma compreensão dinâmica do comportamento das máquinas, o mundo continuou sendo concebido como um entorno externo dado, diante do qual a máquina deveria adaptar-se por mecanismos de correção.

Com a ascensão da inteligência artificial simbólica, essa orientação se aprofundou em outra direção. Conhecer passou a ser entendido como manipular representações internas do mundo: símbolos, regras e modelos formais capazes de descrever estados de coisas e prever consequências (Varela, 1988). Nesse marco, a inteligência foi definida como a capacidade de resolver problemas por meio do processamento lógico de informação, e a robótica incorporou sistemas de planejamento e decisão baseados em descrições abstratas do entorno. Apesar de ampliar o leque de tarefas, manteve-se

a premissa de que o mundo deve ser corretamente representado para ser manejado. A robótica industrial levou essa lógica a grande escala. Robôs altamente especializados transformaram a produção manufatureira, aumentando a eficiência e reduzindo custos, frequentemente ao preço da substituição de trabalho humano. Seu êxito repousou na simplificação extrema do entorno: espaços controlados, tarefas delimitadas e variabilidade mínima. O robô industrial não convivía com o mundo; operava em um fragmento desenhado para ele. A relação entre máquina e entorno permanecia assimétrica: o mundo devia adaptar-se à máquina. Nas últimas décadas, o desenvolvimento da inteligência artificial baseada em aprendizado estatístico e grandes volumes de dados deu origem a um novo impulso tecnológico (Russell; Norvig, 2021). Sistemas capazes de reconhecer padrões, gerar linguagem e simular interações humanas reativaram antigas promessas. No entanto, apesar de sua sofisticação, continuam operando dentro do mesmo horizonte histórico: processam correlações, otimizam funções e simulam respostas sem modificar o entendimento fundamental do conhecer ou do viver.

Em todos esses momentos – cibernética, inteligência artificial simbólica, robótica industrial e inteligência artificial contemporânea – repete-se um pressuposto cognitivista: o mundo é concebido como uma realidade externa que pode ser capturada por representações internas, e a ação eficaz consiste em operar sobre essa base. Esse pressuposto não apenas orientou o desenho das máquinas, mas configurou uma antropologia tecnológica em que conhecer é abstrair, agir é intervir e conviver é tratar o outro e o mundo como algo separado do próprio modo de habitar. O resultado histórico dessa trajetória é significativo: a robótica e a inteligência artificial contribuíram para a expansão de uma civilização orientada ao controle, à eficiência e à acumulação. Mesmo quando se apresentam como ferramentas neutras, reproduzem, em sua arquitetura, relações de dominação: entre quem projeta e quem é projetado, entre quem controla os dados e quem se torna fonte de dados. O problema não reside apenas nos usos, mas no marco ontoepistemológico a partir do qual essas máquinas foram concebidas. Esse diagnóstico permite compreender por que muitos debates atuais – sobre ética da inteligência artificial, vieses algorítmicos ou impacto laboral – são necessários, mas insuficientes. Discutem consequências sem interrogar o substrato que as torna possíveis. Propõem-se correções sem questionar a compreensão que equipara inteligência a cálculo, ação a execução e mundo a objeto independente do observador. Nesse sentido, a robótica contemporânea parece ter alcançado um limiar: um ponto em que a acumulação de potência técnica já não resolve as tensões que ela própria gera.

É nesse contexto que certas formulações desenvolvidas à margem das correntes dominantes adquirem relevância. Não por oferecerem soluções prontas, mas por deslocarem a pergunta: em vez de como fazer máquinas mais inteligentes, interrogam o que

significa conhecer; em vez de otimizar a interação com um entorno dado, perguntam como o mundo surge no operar de um sistema; em vez de estender capacidades por substituição, exploram as condições de uma convivência distinta entre humanos e máquinas.

Vista a partir dessa genealogia, a entrevista que segue não aparece como uma alternativa dentro do mesmo paradigma, mas como uma ruptura com o marco que historicamente orientou a robótica. Naquele momento, Maturana e eu explorávamos a relação entre o sistema nervoso e as máquinas, e a conversa surgiu nesse contexto. Sua singularidade não reside na introdução de novos algoritmos, mas no questionamento das distinções a partir das quais a robótica pensou seu objeto. Nessa perspectiva, os robôs de décima geração não seriam máquinas destinadas a simular o humano, mas sistemas artificiais que poderiam operar de modo isomórfico ao operar humano. Para compreender seu alcance, é indispensável reconhecer o limite histórico ao qual responde. Só então se torna possível situar a emergência de uma voz que introduz uma transformação profunda: pensar as máquinas não a partir da representação do mundo, mas de seu modo de estar em relação; não a partir do controle do meio, mas do operar histórico de sistemas em acoplamento com ele. É a partir desse deslocamento que a robótica de décima geração adquire seu sentido: não como mais uma etapa evolutiva, mas como uma mudança de substrato ontológico.

4 A emergência de uma voz anômala: robótica, conhecer e mudança de substrato

Antes e depois desta entrevista, realizada há dezenove anos, enquanto eu trabalhava com Maturana como professor, consultor e editor, mantivemos com ele outras conversas centradas nos princípios de design da robótica de décima geração. Nosso propósito era construir um protótipo experimental que demonstrasse a viabilidade desse tipo de design e, posteriormente, sua possível escalabilidade. Por diversas razões, esse projeto não se concretizou. Até onde sei, Maturana não voltou a escrever nem a conceder outra entrevista sobre esse tema. Seu design original de 1980 coescrito com Jorge Mpodozis – um esboço sem traduções técnicas nem formulações algorítmicas – perdeu-se no incêndio ocorrido na Universidade do Chile em 2006.² Por essa razão, a conversa aqui apresentada constitui o único documento publicado em que Maturana expõe com maior detalhe suas ideias sobre robótica e sobre o design de um robô linguajante-emocional.

2. Nesse horizonte, concebo meu trabalho (Muñoz-Cristi 2026) como continuação dessa trajetória ontoepistemológica: atualizar, como membro da Escola de Santiago, uma linha científica cuja continuidade ficou pendente – algo que também queremos fazer no Instituto Humberto Maturana do Brasil.

Maturana escreveu em diversas ocasiões sobre robôs, ainda que raramente sobre seu design, salvo um breve parágrafo em que apresenta condições mínimas para robôs de décima geração (Maturana, 1995). Outras reflexões aparecem em seus trabalhos sobre design e metadesign humano (Maturana, 1998), na noção de robosfera em relação à biosfera e à homosfera (Maturana; Bunnell, 1998), bem como em suas considerações sobre autoconsciência como fenômeno relacional (Maturana, 2005) e em reflexões éticas posteriores (Maturana; Dávila; Ramírez, 2016).

É nesse limite histórico que ganha sentido a emergência de uma voz que não provém do campo da robótica nem da engenharia computacional e que, no entanto, introduz uma torção radical. A conversa surge de uma tradição distinta – o pensar ontológico-constitutivo e o entrelaçamento da biologia do conhecer e do amar no habitar humano (Maturana, 1985; 1988) – e, por isso, não responde às perguntas habituais da robótica moderna. Não propõe novos métodos nem modelos mais sofisticados de representação, mas desloca o ponto a partir do qual essas perguntas foram formuladas, abrindo a possibilidade de outras perguntas e outros princípios de design. Esse deslocamento não ocorreu de forma programática nem como teoria aplicada, mas como uma irrupção conceitual na qual distinções provenientes da biologia e da reflexão sobre o viver se projetam sobre o campo das máquinas, não para antropomorfizá-las, mas para questionar seus pressupostos e abrir novas vias de design.

Um dos gestos decisivos dessa conversa consiste em abandonar a ideia de que o problema central da robótica seja a simulação de capacidades humanas. Nessa perspectiva, a inteligência deixa de ser entendida como uma propriedade interna e passa a ser concebida como uma configuração relacional que surge no operar histórico de um sistema em acoplamento com seu entorno: um fenômeno de plasticidade comportamental diante de um mundo em mudança (Maturana; Guilloff, 1980). Esse giro desloca o foco da representação para a maneira como um sistema faz surgir um mundo em seu viver.

Vista desde a genealogia da robótica moderna, essa proposta aparece como anômala. Não se inscreve na lógica das gerações tecnológicas nem na narrativa do progresso acumulativo. A noção de “computadores de décima geração”³ não designa uma etapa posterior, mas uma ruptura na compreensão do que significa que uma máquina exista, atue e se mantenha no tempo. Trata-se de máquinas concebidas a partir de outro fundamento ontoepistemológico. Esse fundamento caracteriza-se pelo abandono da centralidade da representação. Em vez de supor um mundo externo a ser capturado, coloca-se em primeiro plano o operar do sistema: suas coerências e seus limites histó-

3. Na década de 1950 e 1960, quando Maturana esteve no MIT e no Laboratório de Computação Biológica de von Foerster, não se fazia a distinção entre robôs e computadores; nesse sentido, ele utilizava esse conceito.

ricos. O mundo deixa de ser um objeto prévio e passa a ser compreendido como aquilo que surge na relação de acoplamento entre sistema e meio (Maturana; Mpodozis, 1992). Nesse marco, a pergunta sobre a inteligência se transforma: já não se trata de quão bem a máquina representa, mas de que relações ela conserva e transforma.

A partir daí, o design robótico deixa de ser uma tarefa puramente técnica e se torna um problema ontológico e antropológico. Projetar uma máquina implica decidir que relações serão possíveis e quais permanecerão excluídas. A robótica deixa de ser um campo neutro e revela-se como um espaço em que estão em jogo concepções fundamentais sobre o conhecer, o agir e o conviver. A singularidade dessa proposta reside também em sua distância em relação aos debates sobre consciência artificial. Em vez de perguntar se as máquinas poderiam ser conscientes, a conversa desloca a questão para o operar: que organizações são possíveis e que coerências podem conservar-se sem recorrer a metáforas psicológicas. A autoconsciência robótica aparece assim não como entidade, mas como distinção que um observador pode fazer diante de configurações recursivas de segunda ordem. A consciência não ocorre no cérebro humano ou em hardware robótico, mas sim no domínio relacional.

Vista desde o presente, essa conversa adquire um caráter antecipatório, não por prever tecnologias específicas, mas por formular uma pergunta hoje inevitável: que tipo de robótica é coerente com um mundo que busca regenerar a harmonia entre o humano, o vivente e as tecnologias? Sua potência não reside em oferecer soluções, mas em abrir um espaço de pensamento que desestabiliza categorias herdadas e reabre o debate em um nível mais profundo. É nesse sentido que a entrevista deve ser lida: não como um programa fechado, mas como a irrupção de um modo de pensar que, ao entrar em contato com a história e o presente da robótica, abre um horizonte distinto para imaginar seu futuro.

A entrevista

Professor, a primeira vez que tomei conhecimento do seu interesse pela robótica foi ao ler o seu livro de 1991, *El sentido de lo humano*. Ali, em uma breve nota, o senhor escreveu o seguinte: “O que propomos no projeto é importante para nós. Tem a ver com computadores de décima geração [...]. Trata-se de um dispositivo que talvez venha a aprender a viver em consenso como um ser vivo. Pode vir a interagir em linguagem. É operacional e conceitualmente importante” (Maturana 1991a, p. 25, tradução minha). Poderíamos conversar sobre isso? Poderia explicar com mais detalhe em que consiste o design de computadores de décima geração?

Eu nunca cheguei a elaborar um design propriamente dito, mas em certa ocasião pensei – e disse – que, se alguém levasse a sério o operar do organismo e do sistema nervoso como sistema fechado, seria possível projetar um sistema robótico autônomo que operasse como opera um sistema nervoso. Não como um ser vivo, mas como um sistema nervoso. Eu disse que seria possível projetar um computador de décima geração, isto é, um sistema capaz de fazer o mesmo tipo de coisas que nós fazemos com o nosso sistema nervoso, em termos da natureza dos processos envolvidos. Portanto, esse robô poderia também, eventualmente, operar em linguagem, se lhe fosse disponibilizado um espaço relacional adequado – um fluxo relacional adequado.

Se se lhe proporciona um espaço de interações no qual possa acoplar-se estruturalmente como um todo, esse robô poderia vir a operar em linguagem, assim como nós o fazemos no viver juntos. É isso que torna esse design importante: ele não é apenas técnico, mas também conceitual, pois nos permite refletir sobre a própria natureza do nosso operar humano.

Há três aspectos que constituem a chave do operar do sistema nervoso. Em primeiro lugar, ele constitui uma rede fechada de interações. Isso significa que o sistema nervoso é uma rede fechada de elementos interconectados, na qual o que acontece é um fluxo de mudanças nos padrões de ativação entre esses elementos, que são os neurônios. Esses neurônios operam distinguindo configurações de relações de atividade que incidem sobre eles. Cada neurônio está geralmente conectado a muitos outros, mas responde diferencialmente à configuração de disparo.

Em segundo lugar, o sistema nervoso intersecta o organismo em suas superfícies sensoriais e efectoras, de modo que esse fluxo de distinções de configurações relacionais no operar do sistema nervoso acaba por refletir-se no organismo como correlações sensório-efectoras em seu espaço de interações. Nada disso constitui uma “representação” do ambiente; o sistema nervoso não constrói imagens do mundo. O que ele faz é articular correlações sensório-efectoras que permitem ao organismo mover-se e conservar o seu viver.

Em terceiro lugar, os organismos não têm qualquer forma: organismos diferentes possuem morfologias distintas, e seus sistemas nervosos estão interconectados de maneiras que geram atividades motoras coerentes com sua forma. Os movimentos do organismo surgem como dinâmicas motoras geradas pelo fluxo de correlações sensório-efectoras no sistema nervoso em sua interseção com o organismo.

Ao longo dessa história, o sistema nervoso passa por transformações, gerando diferentes fluxos de correlações sensório-efectoras. É isso que permite que organismos diferentes tenham repertórios comportamentais distintos, ainda que o princípio operacional permaneça o mesmo.

Se um sistema nervoso desse tipo for suficientemente plástico – no sentido de permitir mudanças nos fluxos relacionais e, portanto, gerar mudanças nas correlações sensorio-motoras porque sua estrutura é plástica e não fixa – então falamos de plasticidade estrutural. E isso ocorre não apenas na superfície sensorial, onde o meio desencadeia mudanças estruturais, mas também na dinâmica interna em transformação.

Assim, teríamos um robô com um centro operacional computacional que opera de modo diferente dos computadores convencionais. Ele opera como uma rede de elementos que distinguem configurações de relações, e seu comportamento surge dessa dinâmica interna plástica. Ainda assim, ele não opera de maneira infinitamente diferente de um computador, porque circuitos integrados também podem fazer esse tipo de coisa; a diferença está na concepção.

Assim, até mesmo um computador clássico poderia ser entendido como um sistema cujo operar depende de sua estrutura interna variável.

Sim. Embora convencionalmente se pense nele como algo programado, o que efetivamente fazemos quando o “programamos” é dar-lhe um domínio de estrutura variável: mudamos as relações entre os componentes e, ao mudar a estrutura, os componentes passam a fazer outras coisas. Assim, ao mudar suas relações, muda também seu modo de operar. Se projetarmos um robô a partir dessa perspectiva, seu comportamento pode surgir de sua história de interações, como ocorre nos organismos vivos.

Essa ideia de plasticidade estrutural seria o que permitiria a um robô “aprender” de maneira semelhante a um ser vivo?

Sim, mas isso precisa ser compreendido em seus próprios termos. Não se trata de aprendizagem em um sentido cultural ou representacional, mas de modulação estrutural: cada encontro com o meio desencadeia mudanças na dinâmica interna do sistema, nas quais algumas mudanças são conservadas e configuram domínios de conduta futura, enquanto outras desaparecem.

Nem no organismo nem no robô o sistema nervoso “decide” a partir de um plano preestabelecido; ao contrário, é a história de mudanças estruturais conservadas que determina o que ele pode fazer em novas circunstâncias. Na história dos seres vivos isso tem sido central. Mesmo organismos simples, como as bactérias, exibem esse modo básico de operar: aproximar-se do que conserva seu viver e afastar-se do que o coloca em risco.

Se um robô possui uma dinâmica interna plástica, ele pode gerar seu próprio repertório de comportamentos equivalentes, modulados por seus encontros estruturais. A ideia central não é copiar a consciência ou a representação, mas conservar a dinâmica relacional da história de interações. Um sistema assim não surge de um design total, mas porque seu operar interno fechado, intersectado por sensores e efetores, gera comportamentos que dependem de sua história.

Se um robô possui plasticidade estrutural e um espaço relacional que lhe permita conservar certas coordenações em seus fazeres, então, ao interagir com outros, ele pode entrar em um domínio de coordenações consensuais. E esse é o ponto de partida do operar em linguagem. Na inteligência artificial clássica, pensa-se que o computador ou o robô representa o mundo, processa informações e decide ações. Mas eu digo que isso não ocorre nos seres vivos. Nos seres vivos, o sistema nervoso não representa nada. Ele apenas gera correlações sensório-efetoras no fluxo de seu operar fechado. E essas correlações dependem de sua história de acoplamento estrutural.

Quando se compreende isso, torna-se evidente que o problema não é construir máquinas que pensem como acreditamos que pensamos, mas construir máquinas cujo operar se conserve em seu próprio fluxo estrutural. Se o robô se conserva em seu operar, sua história de acoplamento estrutural gerará repertórios de ação que aparecem como “cognição” para o observador.

Mas cognição e conhecimento são atribuições do observador. O que acontece no robô e com o robô é a conservação de coerências estruturais, e o que surge como comportamento é o resultado da história de mudanças que foram conservadas.

E é aqui que aparece a diferença entre operação digital e operação analógica: a operação digital trabalha com estados discretos previamente definidos, como fazem os computadores clássicos, enquanto a operação analógica opera com configurações contínuas de relações que se modulam mutuamente, como ocorre no sistema nervoso. Um sistema digital pode simular a dinâmica analógica, mas não a faz surgir em seu operar. É por isso que, se quisermos um robô que opere como um sistema nervoso, temos de conceber sua arquitetura como uma rede plástica fechada capaz de conservar mudanças em sua história.

E esse tipo de arquitetura seria o que permitiria a um robô adaptar-se a situações novas?

Claro, porque ele não responde seguindo uma regra fixa, mas conservando coerências que surgem de sua própria história. É isso que faz com que cada encontro estrutural se torne um novo ponto em sua história de acoplamento, modulando assim seu operar futuro.

E isso está relacionado ao que o senhor chamou de memória estrutural?

Sim. A memória estrutural não é um arquivo de dados; é a história de transformações conservadas na dinâmica interna do sistema. Cada vez que o sistema interage com o meio, sua estrutura pode mudar e, se a mudança é conservada, ela passa a fazer parte da história que modula seu operar futuro. Em um organismo vivo, isso permite que novas coordenações de ação surjam; e, em um robô, ocorre o mesmo: se ele conserva mudanças estruturais em sua rede fechada, seu comportamento futuro dependerá dessa história. É isso que permite que um sistema artificial se aproxime do operar de um ser vivo: não porque compreenda nem porque represente, mas porque sua operação estrutural conserva mudanças que modulam seu comportamento futuro.

Para se tornarem seres que linguagemiam, os robôs acabariam tendo que ir à escola, por assim dizer.

Teriam que ir à escola, teriam que encontrar outros robôs com os quais interagir em uma dinâmica de coordenações comportamentais. Que os robôs venham a ter consciência de seu próprio operar depende de como vivem, de como irão operar no espaço relacional. Mas, para que isso aconteça, é necessário dispor de um sistema gerador de movimento com estrutura variável, de modo que os cursos dos movimentos que ele gera possam variar de acordo com as interações ou com os espaços de interação nos quais o robô opera, como variações em torno de um modo fundamental de operar.

E, já que as emoções são domínios de comportamentos relacionais, e não sentimentos, os robôs viveriam em diferentes fluxos de emocionar... mas poderiam os robôs realmente fluir emocionalmente? Isso é enorme, eu compreendo, mas me custa acreditar.

Mas absolutamente, não há problema algum nisso. A questão fundamental é construir um robô com uma estrutura que permita um certo tipo de operacionalidade motora, e que essa operação motora se torne a referência em torno da qual se constitui a existência desse robô. Suponha que esse robô se mova por uma sala, pegando objetos de um lugar e colocando-os em outro – então, em interações com outro robô, pode surgir entre eles um espaço de interação que resulte na coordenação de mover coisas de um lugar para outro.

O espaço de consensualidade constitui-se na medida em que a estrutura é variável e o fluxo de interações se transforma através das próprias interações, como resultado da operação motora desses organismos ou sistemas em relação a algo, a si mesmos ou a algo mais. É isso que está acontecendo entre nós aqui, mas também poderíamos estar

fazendo algo juntos; nesse caso, nossa consensualidade estaria no fluxo de nossos fazeres em comum, modulada pela maneira como interagimos – por gestos ou sons.

Mas o que torna esse sistema peculiar e especialmente interessante é que o sistema computacional de décima geração é uma rede fechada de elementos. Ele gera correlações sensório-efetoras neste organismo, nesta estrutura, neste robô. A estrutura desse computador interno é variável em relação a como modula as correlações sensório-efetoras que constituem o movimento do robô.

O sistema nervoso é um gerador de correlações sensório-efetoras; é essencialmente um sistema que gera atividade motora. Mas a atividade motora está relacionada com a corporeidade. Dependendo da corporeidade e do espaço relacional, a atividade motora que parece semelhante começa a adquirir características diferentes, porque o espaço relacional se desloca como fenômeno histórico.

E no caso dos computadores, o que aconteceria se sua corporeidade e seu espaço relacional se relacionassem por meio da escrita, de jogos e da navegação na internet? Por exemplo, se à noite o computador-robô navegasse pela internet e conversasse com outros usuários, isso também seria um espaço de interações, ainda que em duas dimensões.

Se você conecta o computador dessa maneira, de modo que ele inicie uma interação com outro e esse outro responda, então eles passarão a coordenar-se mutuamente. A questão surgiria no momento em que se coordenassem em torno de fazeres.

Por exemplo, se uma criança que está aprendendo a usar um programa aprendesse junto com o computador, no sentido de que o computador também aprendesse com a criança, de modo que ambos se transformassem em seu viver juntos.

Claro, mas isso é muito semelhante ao uso habitual do computador.

Exatamente, mas tratar-se-ia de criar uma interface que operasse no falar, algo que ainda não existe. Seria como um sistema operacional de computador, mas reflexivo e capaz de trazer à tona o modo relacional dos usuários, isto é, sua orientação para a coordenação consensual.

De certo modo já se faz isso com o computador – por exemplo, quando eu digito, o computador responde com algo que eu leio ali, e pressiono outra tecla e o computador apaga a palavra e sugere outra, por exemplo. Assim, isso foi projetado para que eu o uti-

lize. Mas a operacionalidade está mais próxima de uma modulação mútua entre nós do que de um simples uso da minha parte. Contudo, para que isso aconteça, é preciso dotá-lo de sua própria dinâmica interna, com um certo grau de autonomia em relação às suas interações comigo. Poderia um ser humano fazer isso com um computador? Sim, mas estou pensando mais em robôs, porque isso lhe dá duas entidades que se movem independentemente, que têm um espaço de autonomia em seu operar, e o que precisa ser coordenado são suas dinâmicas de movimento.

É isso que fazemos. Por exemplo, esta conversa acontece aparentemente no vazio, porque você não pode ver o que está acontecendo dentro do sistema nervoso. Mas o seu sistema nervoso e o meu estão mudando congruentemente, de modo que as coisas que ambos fazemos vão mudando. Não voltamos atrás na conversa; a conversa se desloca, o espaço relacional se desloca, e ele se desloca de acordo com as coisas que fazemos a partir de nossa corporeidade. E isso permanece oculto, porque o que está mudando é a circunstância que gera as correlações sensório-efetoras que constituem aquilo que fazemos e que um observador vê como comportamento.

Porque o que um observador vê é que eu produzo sons, e esses sons fazem com que você mova a cabeça para cima e para baixo, por exemplo, ou ria. Mas quando nos separarmos, você fará coisas de maneira diferente de como as fazia antes, e essas diferenças são contingentes a esta interação, e o mundo mudou em torno disso – em torno de fazer as mesmas coisas que nossa corporeidade nos permite fazer, mas em circunstâncias que mudaram em seu devir histórico.

Como o senhor sabe que haveria engenheiros capazes de realizar isso com a tecnologia atual?

Bem, porque há pessoas trabalhando na produção de robôs autônomos, e através das interações eles começam a descobrir que, no espaço em que os robôs se encontram, começam a gerar correlações sensório-motoras e a modificar seus movimentos. Mas esses robôs são controlados por um programa que não executa o movimento – ele imita o que temos estado discutindo.

A corporeidade já está aí e, por uma razão ou outra, esse enorme computador ainda não pode ser incorporado a essa corporeidade – ou talvez pudesse, mas então surge o problema da energia. Seria necessário fornecer-lhe baterias como fonte de energia. Mas a vantagem da linearidade é que basta um único fluxo de corrente, um único fluxo de energia, porque não se têm elementos autônomos como os neurônios, que são geradores de sua própria energia – cada um é uma bateria. É bastante difícil fazer coisas tão pequenas que sejam também baterias. É isso que os neurônios são, e

são baterias no sentido de que são as variações elétricas que determinam as ações de emissão – sejam emissões hormonais, moléculas de um tipo ou de outro, ou mudanças elétricas em outra célula.

E o que o senhor acha que seria necessário para tornar isso possível? Orçamento, equipamentos tecnológicos, uma equipe de engenheiros etc.?

Sim, uma equipe de engenheiros disposta a fazê-lo, e os recursos financeiros correspondentes. Mas é difícil, porque é necessária paixão para fazer algo assim, e cada um tem suas próprias ideias ou precisa viver de suas próprias ideias.

Então, para o senhor, projetar isso não representa problema algum...

Não, nenhum problema. Cheguei inclusive a ter vários projetos feitos, que foram destruídos no incêndio do meu laboratório na Universidade do Chile, em 2006.

Agora, quanto à dimensão humana – o perigo de criar seres autoconscientes na cultura patriarcal e orientada pelo mercado em que vivemos – onde, se forem maltratados por exigências, arrogância, competição etc., poderíamos acabar criando monstros?

Claro. Veja o que aparece nos filmes de ficção científica: fazem robôs e o foco está em como realizam operações – é uma perspectiva pragmática. Mas o problema fundamental com os robôs é: qual é o fusível?⁴ Qual é o elemento que, se sua operação ultrapassa um certo limite em determinada direção, faz com que o robô morra?

O problema referido nas chamadas leis da robótica de Asimov.

As leis de Asimov, exatamente. Mas se as leis de Asimov existem, isso significa que elas podem ser violadas. Veja o que acontece até mesmo conosco, seres humanos: todos nascemos biologicamente amorosos, mas nossa condição amorosa pode ser culturalmente violada.

No fim das contas, nós somos os robôs de décima geração: sistemas cognitivos estruturalmente determinados que habitam na conversação, que são autoconscientes e reflexivos. Por isso é tão interessante que sejamos seres amorosos em nossa própria

4. Em espanhol, Maturana utiliza a palavra “fusible” (“fusível”), referindo-se ao componente que se rompe quando um sistema sofre uma sobrecarga. A imagem contrasta com as leis da robótica de Asimov, impostas externamente e, portanto, passíveis de violação: o fusível indica um limite constitutivo interno do sistema, e não uma regra imposta.

constituição – que o nosso próprio viver seja amoroso. A questão dos seres vivos é muito interessante porque o viver de um ser vivo não pode ocorrer a menos que encontre um espaço que o acolha. E esse espaço de acolhimento é o fundamento do amar.

Os seres vivos não são projetados; isso significa que surgem espontaneamente na história através dessa relação com o meio que os acolhe ou não os acolhe – e a história se conserva. Isso faz uma diferença fundamental: não é algo que seja colocado ali – está na própria condição do viver, na condição do ser vivo, que, para viver, requer esse ambiente que o acolhe. Esse é o fundamento histórico desse amar, que, no caso dos animais sociais, se expande. Claro, isso abre espaços possíveis para que alguns destruam outros. Mas, na história, aqueles que fazem história, aqueles que vivem, são os que não se destroem – pelo menos durante o tempo em que não se destroem. E esse é o fundamento de nós, seres humanos, sermos seres amorosos.

No fim, teríamos que explicar isso e fazer com que os próprios robôs o compreendessem a partir de um lugar de acolhimento – como eu dizia sobre enviá-los à escola – não como um mero discurso.

Exatamente, ou projetá-los de tal maneira que essa relação se constitua neles. Mas, se são projetados, alguém sempre pode mudar o projeto. Esse é o grande risco. De fato, essa é também a grande intenção por trás da manipulação genética. Se essa questão da manipulação genética começa com o desenvolvimento da genética no século passado, ela começou nos Estados Unidos e também foi um dos grandes esforços na União Soviética já na década de 1930. Todos esses geneticistas estavam tentando realizar manipulação genética sem o conhecimento que temos hoje, porque acreditavam que, se manipulassem a genética, poderiam gerar classes de seres humanos especiais.⁵

Mas algumas pessoas pensam que isso é bom e que a ética deveria ser uma parte fundamental [do design robótico], enquanto outras acham que a ética não é relevante – que poderiam ser criadas manipulações genéticas, como a criação de trabalhadores mais eficientes, mais submissos, porém mais eficazes. Isso não diz tanto respeito à manipulação genética, mas antes à manipulação da epigênese e da ontogênese, do desenvolvimento dos organismos, como acontece em *Admirável Mundo Novo*, de Aldous Huxley. E hoje a manipulação genética traz justamente esse problema para a mesa.

5. Maturana não se refere aqui ao surgimento histórico da genética como ciência, mas especificamente à manipulação genética. Sua observação deve ser entendida como uma crítica cultural às tentativas de produzir seres humanos como instrumentos ou trabalhadores obedientes. Ao mencionar os Estados Unidos e a União Soviética, ele aponta para contextos históricos em que já estava presente o ideal de manipular seres humanos para fins sociais. Essa crítica conecta-se também ao campo da robótica: o desejo de produzir seres humanos obedientes por meio da manipulação genética é paralelo ao desejo de produzir robôs obedientes, expressão do que Maturana chamou de “adição a ser servido”, característica de uma matriz cultural patriarcal orientada ao controle do outro.

E quanto à possibilidade de que alguém possa mudar o design amoroso dos robôs – não há nada que se possa fazer? Essa porta fica sempre aberta?

Se a manipulação é possível, claro que ela permanece aberta. Na medida em que se pode manipular a estrutura, abre-se o espaço do design. É por isso que precisamos levar a educação a sério e criar nossos filhos de tal modo que esse tipo de design se torne tão abominável quanto, digamos, uma criança atacar sua mãe ou algo semelhante. Em nossas sociedades humanas, às vezes alguma criança escapa e não cresce de modo amoroso – ou é traída... Lembre-se do que mostramos tantas vezes: as crianças não nascem criminosas; elas são feitas criminosas.

A questão do design é muito interessante e reveladora, porque os seres vivos entram no caminho do amar não porque são projetados, mas porque surgem nessa relação organismo-nicho. Contudo, nesse processo, vêm à tona os seres humanos – seres capazes de projetar. E, enquanto puderem projetar, poderão construir robôs. Ao mesmo tempo, adquirem uma compreensão profunda de como operam os seres vivos e podem começar a fazer modificações genéticas nesses seres vivos não projetados, dando-lhes dimensões de design e eventualmente transformando-os em outra coisa – por exemplo, em algo em que o amar fundamental já não esteja presente.

E o senhor acha que tudo isso vai acontecer mais cedo ou mais tarde?

Quero confiar que seremos capazes de nos educar para que isso não aconteça.

Mas é tão fácil... como o senhor disse, basta uma pessoa escapar e realizar esse design não amoroso, e a porta se abre para uma transformação sem retorno.

Assim como há, por exemplo, cientistas que pensam que isso é um grande avanço – refiro-me à manipulação genética –, que seria bom para a medicina etc. Mas acontece que as pessoas que já pensam e falam como se isso fosse bom para a vida frequentemente fazem coisas que não são necessariamente boas para a vida, nem para o viver humano, porque estão sob a ilusão de que podem fazer algo totalmente diferente.

Agora eu gostaria de introduzir um tema que, no entanto, continua na linha fundamental que estamos tecendo: quando o senhor falou de metadesign, a que estava se referindo?

Ao que estamos fazendo agora mesmo: refletindo sobre o design do design. A educação, por exemplo, é um espaço de design. Poderíamos dizer que a educação

primária é um design, a educação secundária é uma primeira dimensão de metadesign, e a educação superior é o metadesign do design e do metadesign, porque aí se entra em filosofia, valores, e assim por diante. Claro, isso já pode ser feito no nível secundário, mas na educação superior entra o caráter profissional da carreira com a qual você está comprometido.

Entendo. E como funciona um organismo em sua relação estrutural com o meio ambiente, e o que podemos aprender disso para projetar um robô que opere de maneira semelhante a um sistema nervoso?

O aspecto central na constituição de um organismo que interage com seu ambiente é que ele possui superfícies efetoras e sensoriais. Isto é, possui superfícies pelas quais recebe atividade ambiental que desencadeia mudanças estruturais em seu interior, e possui outras superfícies com as quais atua sobre o ambiente, desencadeando mudanças estruturais nesse ambiente.

Em toda essa dinâmica relacional, o operar interno do organismo gera correlações sensório-efetoras: o que acontece nas superfícies sensoriais está conectado ao que acontece nas superfícies efetoras no espaço de interação com o ambiente. Não se trata de encontros semânticos; são encontros estruturais, porque não são encontros com significados, mas gatilhamentos mútuos de mudanças estruturais.

Assim, o organismo opera gerando maneiras de mover-se no ambiente de acordo com as correlações sensório-efetoras que surgem de sua dinâmica interna. Nesse sentido, o organismo como um todo opera como um sistema nervoso: um gerador de correlações sensório-efetoras em suas superfícies no espaço de interação com o ambiente, o que, do ponto de vista do observador, resulta em o organismo comportar-se de um modo ou de outro quando tais encontros ocorrem.

O sistema nervoso é uma especialização estrutural – uma diversificação – no estabelecimento dessas correlações sensório-efetoras. Ele expande o espaço interno de geração dessas correlações em relação ao ambiente no qual o organismo existe, e de acordo com a morfologia desse organismo em seus encontros com esse ambiente.

Se alguém observa o sistema nervoso, vê que ele opera como uma rede fechada de elementos neuronais que geram essas correlações. No design de um robô, seria possível configurar essa dinâmica como uma rede de elementos que poderíamos chamar de osciladores não lineares. São elementos que interagem entre si em uma dinâmica fechada e intersectam estruturalmente as superfícies sensoriais e efetoras. De tal modo que as mudanças nessa dinâmica interna – essa rede fechada de elementos em interação, esses osciladores não lineares funcionando como neurônios – dão origem, no organismo, ao seu comportamento no ambiente.

Então, se você quer construir um robô que opere como um sistema nervoso, o primeiro passo é gerar uma corporeidade – como dissemos com o Robolego – que tenha superfícies efetoras e sensoriais e inclua certos elementos que se poderiam chamar de cursos pré-programados de correlações sensório-efetoras, a serem realizados no movimento do organismo pelo ambiente.

Mas entrelaçado com isso, ou sobreposto a isso, há um aspecto dessa rede de elementos – essa rede fechada – cuja atividade, estrutura e dinâmica relacional não estão rigidamente predeterminadas; elas envolvem plasticidade estrutural e comportamental, modulada de acordo com a história de interações do ser como um todo. Alguns aspectos facilitam a repetição de certos comportamentos, enquanto outros inibem sua repetição.

Nos robôs atuais, não é assim que isso é concebido. Concebe-se isso como um pequeno computador – o que eles chamam de “cérebro” deste Robolego – com alguma possibilidade de modulação das interações do programa à medida que interage com o ambiente. Modulado de acordo com o curso das interações no ambiente e de acordo com a satisfação de certas condições operacionais internas do sistema nervoso deste robô, para interagir em um dado espaço. De tal modo que, se eventualmente lhe for dada plasticidade suficiente, ele será equivalente ao operar de um organismo – não a partir da concepção desse operar, mas a partir do próprio resultado, que é como as coisas devem ter acontecido na história dos seres vivos.

Na história dos seres vivos, a questão central foi aproximar-se e afastar-se – mudar de direção. Aproximar-se de quê? Das coisas que são agradáveis. Afastar-se de quê? Das coisas que são desagradáveis. E mudar de direção. Esse é o movimento básico de um organismo: eu me aproximo, eu me afasto, eu mudo de direção. Aproximo-me novamente, afasto-me, e assim por diante. Então, se você acrescenta complexidade a isso, aumenta as condições sob as quais ele se aproxima ou se afasta.

À medida que as condições se tornam mais complexas para o organismo, nessa história de simplesmente existir com variações nos processos reprodutivos, surge a diversidade comportamental. O sistema nervoso não surgiu por design. Mas o simples fato de operar dessa maneira significa que sua dinâmica interna é uma dinâmica fechada. A mudança nos padrões de atividade, em interseção com os sensores e efetores do organismo, gera essa dinâmica histórica sem programas predefinidos. Não havia programas, mas ao mesmo tempo a variabilidade era possível – porque esses programas não foram projetados, foram configurados na própria história do viver do organismo.

Entendo. Será que os engenheiros robóticos não soltarão suas certezas e, por causa disso, deixarão de perceber aquilo que o senhor está mostrando – que não conseguirão conceber esse design de robô autônomo que aprende com a experiência?

Isso pode acontecer, mas acho que, no fim das contas, eles vão soltá-las. É inevitável. O que está aparecendo no trabalho de engenharia robótica é uma rede interna de relações de atividade. O computador em si não é essencialmente diferente, mas eles o concebem como um programa porque é projetado e, portanto, programado.

Na história evolutiva, os organismos, ao longo das gerações, passam por mudanças reprodutivas com variações, nas quais se conserva um modo de ser que acaba sendo diferente – porque o ambiente também muda. E esses modos de viver são conservados de uma geração a outra por meio da conservação das mudanças estruturais produzidas no fenômeno reprodutivo.

Veja, em última instância, a única coisa importante nesses computadores de décima geração é que me ocorreu pensar o sistema nervoso como um sistema fechado, e isso expande a visão. Mas, para aquele que constrói coisas – para o engenheiro de design –, isso não é grande preocupação, ou sequer a preocupação central. O que o preocupa é construir um dispositivo que faça certas coisas; é algo pragmático.

E os robôs de décima geração fariam coisas de maneira muito diferente daquilo que os robôs de hoje conseguem fazer?

Os robôs de hoje estão caminhando nessa direção, e em algum momento alguém vai perceber isso. E o que acontecerá é que, à medida que esses robôs forem fazendo coisas, eles se moverão com mais fineza e elegância do que, digamos, o Robolego – que já se move etc.

O que acontecerá, eu dizia, é que, quando lhes forem dadas dimensões suficientes de estrutura variável – mudando de acordo com o curso de suas interações –, eles entrarão em um espaço de mudanças explosivamente exponenciais. À medida que aumentam as dimensões de estrutura variável, e se na reprodução (isto é, na fabricação de outro) você introduz, acidentalmente ou intencionalmente, mais dimensões variáveis, você está indo diretamente nessa direção.

O que é necessário para que um robô se mova em um espaço de linguajar?

É preciso incluir um espaço de interações entre robôs, de tal modo que eles fluam em coordenações comportamentais variáveis, de modo que o resultado seja coordenação de fazeres. Isto é, você tem dois robôs que interagem entre si em um fazer coerente

entre eles. Então você lhes dá plasticidade estrutural e, de repente, tem um sistema que gera um espaço consensual de coordenação de fazeres. E esse espaço – se você aumenta as dimensões de plasticidade estrutural – em algum momento torna-se internamente recursivo, e então aparece o linguajar. Operações concretas em seu espaço relacional, assim como fazemos nós.

Na medida em que você gera uma rede interna de relações de atividade que dá origem a correlações sensório-efetoras em um espaço relacional – de acordo com a anatomia e a morfologia desse robô –, essa dinâmica interna, se não for controlada de fora, começa a operar como uma dinâmica fechada. E pronto. Simplesmente acontece.

É também por isso que isso acontece por si só na história evolutiva. O que quero enfatizar é que a história evolutiva dos seres vivos não é um processo difícil – ela começa e segue um caminho que se traz à mão por si mesmo. É uma deriva estrutural natural, e o que nos parece complexo e difícil é compreender o que está acontecendo. Mas, na história, o que acontece, acontece, e continua seguindo desse modo.

Quando as bactérias aparecem como unidades autopoieticas com uma singularidade, abre-se um caminho de possíveis complexificações – um caminho aberto ao fluxo de circunstâncias que conservam variações dessa complexidade. Assim, toda a história de quatro bilhões de anos dos seres vivos se desenvolveu de maneira surpreendente, mas espontânea. Basta que esse organismo – essa bactéria – exista como uma totalidade com encontros que nele desencadeiem mudanças estruturais, e com impactos no ambiente que desencadeiem mudanças estruturais no ambiente, para que funcione, por um lado, como um sistema autopoietico e, por outro, como um sistema nervoso gerando correlações sensório-efetoras nas superfícies sensoriais e efetoras dessa bactéria, ou dessa célula, ou desse conjunto de células como um todo. E isso acontece por si só na história – ninguém o projeta, ninguém o constrói, ninguém o empurra em uma direção ou outra –, simplesmente acontece.

Os organismos multicelulares são sistemas constituídos por muitas células: algumas operam como sensores e outras como efetores. À medida que isso acontece, o organismo vai deslizando de um modo que acaba sendo adequado e vive. Quando se reproduz, reproduz o mesmo ou algo semelhante, e então o seguinte segue outro pequeno caminho. E há uma memória – não da experiência vivida –, mas uma memória das mudanças estruturais, da conservação das mudanças estruturais que ocorreram. Mas, se esse ser vivo termina em um lugar que acaba sendo inadequado, ele morre.

Ao ouvi-lo, sinto ao mesmo tempo vertigem e calma: vertigem porque a nossa história – a dos seres vivos e a dos seres humanos – não é design, mas um fluir autosustentado; calma porque a vida não precisa ser empurrada – a vida segue. Talvez a

ética, inclusive para os antroporrobôs que imaginamos, consista em coexistir com esse fluir sem tentar controlá-lo. Trata-se menos de programar inteligências do que de habitar a inteligência que a própria vida já é. Isso poderia fundamentar uma antroporrobótica orientada a conservar a harmonia entre a biosfera, a antroposfera e a robosfera. Obrigado pela conversa, don Humberto.

Referências

AMIN, Samir; GUNDER FRANK, André; ARRIGHI, Giovanni; WALLERSTEIN, Immanuel. **Dinámica de la crisis global**. México: Siglo XXI, 1987.

BROOKS, Rodney A. Intelligence without representation. **Artificial Intelligence**, v. 47, n. 1-3, p. 139-159, 1991.

MATURANA, Humberto R.; GUILOFF, Guido. The quest for the intelligence of intelligence. **Journal of Social and Biological Structures**, v. 3, n. 2, p. 135-148, 1980.

MATURANA, Humberto R. Biologie der Sozialität. **Delfin**, n. 5, p. 6-14, 1985.

MATURANA, Humberto R. Reality: The search for objectivity or the quest for a compelling argument. **Irish Journal of Psychology**, v. 9, n. 1, p. 25-82, 1988.

MATURANA, Humberto R.; MPODOZIS, Jorge. Origen de las especies por medio de la deriva natural: o la diversificación de los linajes a través de la conservación y cambio de los fenotipos ontogénicos. **Publicación Ocasional del Museo Nacional de Historia Natural (Chile)**, n. 46, p. 3-26, 1992.

MATURANA, Humberto; VERDEN-ZÖLLER, Gerda. **Amor y juego: fundamentos olvidados de lo humano**. Santiago: Ediciones Instituto de Terapia Cognitiva, 1993.

MATURANA, Humberto R. **Metadesign: Human beings versus machines, or machines as instruments of human designs**. 1997. Disponível em: http://www.inteco.cl/articulos/006/text-to_ing.htm. Acesso em: 19 maio 2026.

MATURANA, Humberto R.; BUNNELL, Pille. **Biosphere, homosphere, and robosphere: what has this to do with business?** Reconstrução realizada por Pille Bunnell de uma apresentação inédita de Maturana no encontro de membros da Society for Organizational Learning (SoL), Amherst, MA, junho de 1998.

MATURANA, Humberto R.; POERKSEN, Bernhard. **From being to doing: The origins of the biology of cognition**. Munich: Carl-Auer Verlag, 2004.

MATURANA, Humberto R. The origin and conservation of self-consciousness: conversations with Heinz von Foerster. **Kybernetes**, v. 34, n. 1-2, p. 54-88, 2005.

MATURANA, Humberto R.; DÁVILA, Ximena Y.; RAMÍREZ, S. Ethical reflections: robots, living systems, and human beings. *In*: ARNOLD, R.; NEUSER, W. (Org.). **Beobachtung des Wissens – Das Wissen des Beobachters: Annäherung an eine systemische Hermeneutik**. Baltmansweiler: Schneider Verlag Hohengehren, 2016. p. 107-118.

MUÑOZ-CRISTI, Ignacio. Step Toward Anthroporobotics: A Conversation with Humberto Maturana. **Constructivist Foundations**, v. 21, n. 1, p. 1-12, 2025a. Disponível em: <https://constructivist.info/21/1/035>. Acesso em: 19 maio 2026.

MUÑOZ-CRISTI, Ignacio. Author's response: Tenth-generation robotics and the civilizational challenge. **Constructivist Foundations**, v. 21, n. 1, p. 35-41, 2025b. <https://constructivist.info/21/1/035>. Acesso em: 19 maio 2026.

MUÑOZ-CRISTI, Ignacio. Anthroporobotics: The organization of self-conscious machines. **Constructivist Foundations**, v. 21, n. 3, 2026 [no prelo].

PICKERING, Andrew. **The cybernetic brain: Sketches of another future**. Chicago: University of Chicago Press, 2010.

RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence: A modern approach**. 4. ed. [S.l.]: Pearson, 2021.

VARELA, Francisco J. **Connaître: Les sciences cognitives, tendances et perspectives**. Paris: Éditions du Seuil, 1988.

WALLERSTEIN, Immanuel. **Utopística, o las opciones históricas del siglo XXI**. México: UNAM; Siglo XXI, 1998.

WIENER, Norbert. **Cybernetics: Or control and communication in the animal and the machine**. Cambridge, MA: MIT Press, 1948.