



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

RENATO FERNANDES RIBEIRO

**UM ESTUDO DA “INTENCIONALIDADE SEM CONTEÚDO” PROPOSTA PELO
ENATIVISMO RADICAL E SUA RELAÇÃO COM A INTELIGÊNCIA ARTIFICIAL**

FORTALEZA

2018

RENATO FERNANDES RIBEIRO

UM ESTUDO DA “INTENCIONALIDADE SEM CONTEÚDO” PROPOSTA PELO
ENATIVISMO RADICAL E SUA RELAÇÃO COM A INTELIGÊNCIA ARTIFICIAL

Dissertação apresentada ao Curso de do
Programa de Pós-Graduação em Ciências
da Computação do Centro de Ciências da
Universidade Federal do Ceará, como requisito
parcial à obtenção do título de mestre em . Área
de Concentração: Inteligência Artificial

Orientadora: Prof.^a Dr.^a Ana Teresa de
Castro Martins

Coorientador: Prof. Dr. Marcelino Cavalcante
Pequeno

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R372e Ribeiro, Renato Fernandes.

Um estudo da "intencionalidade sem conteúdo" proposta pelo enativismo radical e sua relação com a inteligência artificial / Renato Fernandes Ribeiro. – 2018.
55 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2018.

Orientação: Profa. Dra. Ana Teresa de Castro Martins.

Coorientação: Prof. Dr. Marcelino Cavalcante Pequeno.

1. Intencionalidade. 2. Conteúdo. 3. Inteligência Artificial. 4. Enativismo Radical. I. Título.

CDD 005

RENATO FERNANDES RIBEIRO

UM ESTUDO DA “INTENCIONALIDADE SEM CONTEÚDO” PROPOSTA PELO
ENATIVISMO RADICAL E SUA RELAÇÃO COM A INTELIGÊNCIA ARTIFICIAL

Dissertação apresentada ao Curso de do
Programa de Pós-Graduação em Ciências
da Computação do Centro de Ciências da
Universidade Federal do Ceará, como requisito
parcial à obtenção do título de mestre em . Área
de Concentração: Inteligência Artificial

Aprovada em: 30 de Novembro de 2018

BANCA EXAMINADORA

Prof.^a Dr.^a Ana Teresa de Castro Martins
(Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcelino Cavalcante Pequeno
(Coorientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Carlos Eduardo Fisch de Brito
Universidade Federal do Ceará (UFC)

Prof. Dr. Francicleber Martins Ferreira
Universidade Federal do Ceará (UFC)

Prof. Dr. Pedro Porfírio Muniz Farias
Universidade de Fortaleza (UNIFOR)

A todas as mentes dispostas a filosofar.

AGRADECIMENTOS

Ao Prof. Dr. Marcelino Cavalcante Pequeno, por me orientar em minha dissertação de mestrado, e por ter me inspirado a estudar filosofia da mente. À Prof.^a Dr.^a Ana Teresa de Castro Martins, por todo acompanhamento e apoio fundamentais para a conclusão deste mestrado. Agradeço também ao Prof. Dr. Carlos Eduardo Fisch de Brito, por ter sido grande fonte de conhecimento e inspiração nos temas que envolvem este trabalho, e cuja iniciativa de organizar debates, seminários e congresso dos quais participei foi crucial para meu progresso intelectual.

Aos demais membros da banca. Prof. Dr. Francicleber Martins Ferreira, por ser parte da organização e do enriquecimento de muitos dos debates aos quais estive presente, e pelas críticas e comentários que ajudaram a guiar a reta final do trabalho. E Prof. Dr. Pedro Porfírio Muniz Farias, pelas importantes observações escritas que contribuíram para o melhoramento desta dissertação.

A todos os demais professores que fizeram parte do meu aprendizado, alguns desde a graduação, e que me deram a base necessária para ir adiante.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo financiamento da pesquisa de mestrado por meio de bolsa de estudos.

Agradeço, também, ao meu irmão Rodrigo Fernandes Ribeiro, e a todos os meus amigos, entre eles: Dimitri-Alexander, Taynan Cid, Vitor Cabral, Nichollas Fonseca, Arthur Lima, Roberto Evangelista e Giovani Loiola. Pessoas estas que me apoiaram durante todo o processo, e com os quais pude diversas vezes compartilhar meus estudos e descobertas.

Agradeço, em especial, a minha namorada Ingrid de Sales Rabelo, que me acompanhou e me assistiu em todas as fases importante do mestrado. Que me incentivou nos momentos difíceis, e com quem pude compartilhar os momentos de felicidade.

E por fim, quero agradecer aos meus pais, Sueli Fernandes Ribeiro e José Rusney Silva Ribeiro, pelos cuidados, pelo carinho, por acreditarem em mim e investirem no meu potencial, e principalmente, por sempre terem me educado com base nos princípios do respeito, da razão e do pensamento crítico.

“[...]from so simple a beginning endless forms
most beautiful and most wonderful have been,
and are being, evolved.”

(Charles Darwin)

RESUMO

O problema de se construir uma inteligência artificial genuína continua em aberto. O paradigma que definiu as tentativas iniciais baseava-se em dois pressupostos filosóficos: o representacionismo, que afirma que a cognição se dá através de uma réplica interna do mundo na mente do observador; e a teoria computacional da mente, que diz que o raciocínio é uma computação dessas representações simbólicas no cérebro do observador. Críticas a esses pressupostos vieram à tona, e uma das alternativas mais recentes, o enativismo radical, propõe que cognição envolvendo conteúdo nasce com as práticas culturais humanas, e que a cognição em mentes mais básicas não envolve conteúdo. Portanto, neste nível, deve existir “intencionalidade sem conteúdo”. Os objetivos deste trabalho são: apresentar de maneira breve as bases históricas e filosóficas da inteligência artificial, bem como as tentativas de fugir de paradigmas mais tradicionais; discutir os termos filosóficos intencionalidade e conteúdo; e apresentar um estudo da “intencionalidade sem conteúdo” proposta pelo enativismo radical. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Palavras-chave: Intencionalidade. Conteúdo. Inteligência Artificial. Enativismo Radical.

ABSTRACT

The problem of building a genuine artificial intelligence stays open. The paradigm that defined the first attempts was based in two philosophical assumptions: the representationalism, which claims that cognition happens through an internal replica of the world in the mind of the observer; and the computational theory of mind, which affirms that thinking is a computation of those symbolic representations inside the brain of the observer. These assumptions have been criticized, and one of the most recent alternatives, the radical enactivism, proposes that cognition involving content comes to life with human cultural practices, and that cognition in basic minds does not involve content. Therefore, at this level, "intentionality without content" must exist. The goals of this academic work are: briefly presenting the historical and philosophical roots of artificial intelligence, as well as the attempts of escaping from more traditional paradigms; discussing the philosophical terms intentionality and content; and presenting a study of "intentionality without content" proposed by the radical enactivism. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Keywords: Intentionality. Content. Artificial Intelligence. Radical Enactivism.

LISTA DE TABELAS

Tabela 1 – Requerimentos de sistema - IA enativa	23
--	----

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Organização do trabalho	12
2	MÉTODO DE PESQUISA	14
2.1	Objetivos	14
2.2	Metodologia empregada	14
2.3	Resultados esperados	14
3	CONTEXTUALIZAÇÃO HISTÓRICA	15
3.1	Enativismo	16
3.2	IA corporificada (Embodied AI)	17
3.3	IA Enativa de Froese e Ziemke	19
4	O PROBLEMA	26
4.1	Intencionalidade	28
4.2	Conteúdo	31
5	ENATIVISMO RADICAL	33
5.1	The Hard Problem of Content	34
5.2	A origem natural do conteúdo	35
5.3	Enativismo Radical (REC) vs Enativismo Adaptativo Autopoietico (AAE)	38
6	“INTENCIONALIDADE SEM CONTEÚDO” - ENTENDENDO A UR- INTENTIONALITY	43
6.1	Reação vs Intencionalidade	48
7	PROXY CAUSE - UMA CONTRIBUIÇÃO AO ENATIVISMO RADICAL	50
8	CONCLUSÃO	53
	REFERÊNCIAS	54

1 INTRODUÇÃO

O problema de se desenvolver computacionalmente uma inteligência genuína, projeto por vezes referido ao longo da história da Inteligência Artificial (IA) como “IA verdadeira” ou “IA geral”, é um dos mais antigos sonhos da área e continua como um problema em aberto. O paradigma que definiu as tentativas iniciais baseava-se em dois pressupostos filosóficos: o representacionalismo, que afirma que a cognição se dá através de uma réplica interna do mundo na mente do observador; e a teoria computacional da mente (ou computacionalismo), que diz que o pensamento é uma computação dessas representações simbólicas no cérebro do observador.

Os esforços computacionalistas da IA fizeram parte do surgimento de um movimento interdisciplinar maior que veio a ser chamado de ciência da cognição, cujo objetivo é entender a natureza dos fenômenos cognitivos, e que, além da própria IA, inclui áreas como filosofia da mente, psicologia, linguística, antropologia, biologia e neurociência.

Críticas ao computacionalismo surgiram dentro das ciências cognitivas, e questionavam desde a ideia de que a mente era formada por representações simbólicas até a própria noção de que a cognição era um fenômeno puramente interno. Hoje, um dos principais movimentos dentro da ciência da cognição que defende uma narrativa externalista para a cognição (ou seja, que a considere um fenômeno que emerge da interação em si do agente com seu meio externo, e não um mero processamento interno de representações desse meio) é o enativismo.

O enativismo radical, em especial, definido de maneira mais extensa por Daniel Hutto e Erik Myin no livro *Radicalizing Enactivism - Basic Minds without Content*, (HUTTO; MYIN, 2012), defende que existe um nível básico de cognição onde não há conteúdo. Esta perspectiva serviu de motivação para este trabalho, pois dentro de uma lógica de pesquisa ascendente (bottom-up), o primeiro desafio rumo a uma inteligência artificial genuína seria o de pesquisar como surge o fenômeno da cognição em sua forma mais básica.

O conceito de intencionalidade é especialmente intrigante para o que propomos estudar. A maneira usual de se falar de "intenção" se refere a objetivo ou propósito. Para a filosofia, no entanto, este termo traz um conceito mais sofisticado. Ele se refere à propriedade do mental de ser sobre algo, de se dirigir ou apontar para um objeto, propriedade essa por vezes referida como aboutness ou directedness. Para citar alguns exemplos, sempre desejamos, imaginamos, buscamos ou rejeitamos algo. Segundo a visão clássica de intencionalidade resgatada da escolástica e redefinida por Franz Brentano, intencionalidade seria portanto a propriedade do mental de apontar para um conteúdo intencional, este que seria uma representação

interna de algo que pode ou não ser um objeto do mundo real. Percebemos que este conceito se faz importante dentro do estudo da IA quando artigos como o de T. Froese e T. Ziemke (FROESE; ZIEMKE, 2009) por vezes criticam tentativas de se desenvolver inteligência artificial por essas não possuírem o que eles chamam de "agência intencional", ou seja, a capacidade do agente de se acoplar ao mundo e agir dentro dele apontando para objetos deste mundo de uma maneira que faça sentido para o próprio agente (o que os enativistas chamam de sense-making). Mesmo em uma lógica de pesquisa ascendente, a intencionalidade é importante. Daniel Hutto e Erik Myin em seu mais recente livro *Evolving Enactivism: Basic Minds Meet Content* (HUTTO; MYIN, 2017) se recusam a abandonar o conceito de intencionalidade e, como uma continuação de sua obra de 2012, advogam que para formas básicas de cognição sem conteúdo deve haver portanto intencionalidade sem conteúdo. No entanto, eles admitem que, segundo uma abordagem clássica, isto se configura como uma contradição em termos. Investigar o fenômeno da inteligência passa por analisar a cognição em seu nível mais básico, e a chamada “intencionalidade sem conteúdo” – termo tão polêmico e ao mesmo tempo tão importante se considerarmos a possibilidade de mentes básicas sem conteúdo – está no cerne desta investigação, e é a grande motivação para este trabalho.

A finalidade do trabalho, portanto, é a de contarmos uma história breve da inteligência artificial a fim de contextualizar suas bases filosóficas mais tradicionais e suas tentativas de fugir de paradigmas representacionistas tradicionais; discutirmos termos como inteligência, intencionalidade e conteúdo; e por fim, analisarmos a proposta de “intencionalidade sem conteúdo” do enativismo radical, comparando-a com a maneira que o enativismo mais tradicional enxerga conteúdo e intencionalidade, e também com a vertente mais aceita do cognitivismo - a Teleossemântica.

1.1 Organização do trabalho

No tópico 1, introduzimos o trabalho, contextualizando e explicando as motivações. O tópico 2 abordará brevemente o método de pesquisa, apresentando os objetivos, a metodologia empregada, e os resultados esperados. O tópico 3 é uma contextualização histórica, que explicará como, partindo do projeto da inteligência artificial, chegamos ao enativismo. Este tópico também apresentará duas das alternativas à clássica inteligência artificial simbólica, são elas a IA corporificada e a IA enativa. Muito do pensamento enativista mais tradicional estará condensado nesta seção. O tópico 4 apresentará o problema: como questionamentos sobre

inteligência nos levam a investigar “intencionalidade sem conteúdo”? O que seria de maneira resumida “intencionalidade” e “conteúdo”? O tópico 5 abordará o enativismo radical, posição que defende a “intencionalidade sem conteúdo”, e o tópico 6 tratará da naturalização deste tipo de intencionalidade em si, proposta pelos enativistas radicais e chamada por eles de *ur-intentionality*. Por fim, o tópico 7 apresentará uma possível contribuição para o enativismo radical, as *proxy causes*, e o tópico 8 concluirá o trabalho.

2 MÉTODO DE PESQUISA

2.1 Objetivos

Os objetivos gerais deste trabalho são: i - Falar brevemente sobre a história da inteligência artificial, descrevendo seus objetivos iniciais, inspirações tomadas do enativismo e da fenomenologia, e tentativas de deslocamento de um paradigma clássico para novas formas de se fazer IA, como são os casos da IA corporificada e da IA enativa. ii - Discutir conceitos empregados nas ciências e na filosofia para tentar descrever cognição - em especial intencionalidade e conteúdo.

O objetivo específico deste trabalho é: i - Investigar e entender o fenômeno da “intencionalidade sem conteúdo” proposto pelo enativismo radical, e compará-lo com a visão mais tradicional do enativismo.

2.2 Metodologia empregada

Este trabalho visa realizar uma pesquisa de base cujo interesse a longo prazo é o de avançar o projeto de se construir uma inteligência artificial genuína, através da investigação do fenômeno da cognição. É uma pesquisa muito mais próxima da exploração científica e filosófica do que do uso de engenharia para a solução de um problema prático. Por conta disso, a pesquisa se dará de maneira exploratória, com o uso de levantamento bibliográfico e análise de exemplos. As conclusões e resultados levantados serão analisados de maneira qualitativa, com o objetivo de tentar traçar conexões entre as diferentes teorias e pontos de vista a cerca do assunto.

2.3 Resultados esperados

Esperamos com esta pesquisa fornecer conexões esclarecedoras entre vários elementos próprios do vocabulário das ciências da cognição (como intencionalidade, conteúdo, representação, etc), e em especial entender a proposta do enativismo radical de “intencionalidade sem conteúdo”, que se configura como um potencial ponto de partida para o estudo da cognição.

3 CONTEXTUALIZAÇÃO HISTÓRICA

Desde o surgimento das primeiras máquinas a realizar computação, o poder dessas máquinas de manipular símbolos já incitava especulações sobre a mente ser o produto de uma computação sofisticada realizada no cérebro e sobre o computador ser capaz de replicar a mente. Isto fica bem claro na cronologia apresentada por (HENDERSON, 2007), onde ele cita que o próprio Alan Turing na década de 1940 - quando os primeiros computadores reais estavam sendo criados - já lançava perguntas como: o que a mente faz pode ser expresso como uma computação? um computador pode se tornar tão avançado a ponto de as pessoas não identificarem que não se trata de um humano?

Continuando a cronologia, em 1956 John McCarthy cunhou o termo inteligência artificial, e junto de Marvin Minsky, organizaram a conferência de Dartmouth que marcou o início dos estudos nessa área. Cientistas como Allen Newell e Herbert Simon, que estavam presentes, já trabalhavam nessa época com um programa para prova automática de teoremas (O Logic Theory Machine, ou LTM). Eles criaram em 1960 o General Problem Solver (GPS), em que dados a especificação do domínio de um problema, um conjunto de operadores e diretrizes gerais para o uso dos mesmos, o programa desenvolvia uma solução para o problema usando esses operadores (HENDERSON, 2007). Também formularam a hipótese do Sistema Simbólico Físico, que, diziam eles, teria os meios necessários e suficientes para gerar ações inteligentes gerais (NEWELL; SIMON, 1976). Este conceito equiparava o cérebro a um computador, e por conta dele, a IA que se apoiava nessas bases ficou conhecida como IA simbólica.

Essa linha de pesquisa se apoia filosoficamente em dois conceitos. O primeiro deles é o representacionalismo. A mente é sempre vista como um conjunto de representações (ou réplicas) dos elementos contidos no mundo que são produzidas pelo cérebro. O segundo conceito é a da teoria computacional da mente (ou computacionalismo), que leva além o conceito das representações e as define como símbolos tal que o raciocínio sobre eles é uma computação realizada no cérebro.

Com o passar do tempo o projeto da IA simbólica para criação de inteligência geral foi se mostrando ineficaz, e ocorreu uma mudança nos objetivos dos estudiosos da área. A IA passou a ser usada para resolução de problemas reais que poderiam ser modelados formalmente e que poderiam se beneficiar com o uso de heurísticas e meta-heurísticas para achar soluções ou otimizações do problema. Esse tipo de IA passou a ser conhecida como Good Old-Fashioned Artificial Intelligence (GOFAI). Frente ao fracasso dos projetos de IA genuína, a GOFAI sofreu

críticas filosóficas, e segundo T. Froese e T. Ziemke, apesar das diferenças, todas essas críticas concordavam que faltava a esses sistemas a propriedade de agência intencional, e que não há sentido em falar que um sistema sabe o que está fazendo sem essa propriedade (FROESE; ZIEMKE, 2009).

A busca por novos paradigmas que pudessem definir bases para a IA diferentes dos tradicionais representacionalismo e computacionalismo convergiu com as ideias dentro da filosofia da mente que buscavam fugir dessas mesmas bases. A corrente mais proeminente a ir contra essas duas bases tradicionais é o enativismo. A seguir, faremos uma breve introdução ao enativismo, e seguiremos por apresentar duas formulações dentro da IA que se utilizaram de muitos de seus conceitos.

3.1 Enativismo

Segundo (FROESE; PAOLO, 2011), a abordagem enativista teve início no livro *The Embodied Mind* (VARELA *et al.*, 1991). Como objetivo deste trabalho, os autores se propuseram a apresentar um modelo de ciência cognitiva que se contrapusesse ao paradigma computacionalista que dominava a área, falando de conceitos como o corpo sendo pré-requisito para a experiência sensorial e motora; imersão desse corpo em contextos biológico, psicológico e cultural; e a autonomia do ser vivo como conceito central para a cognição. A base fundamental para a ideia de autonomia foi o conceito de autopoiese, neologismo criado em 1971 por Humberto Maturana e Francisco J. Varela para definir o conceito de vida: An autopoietic system is organized (defined as unity) as a network of processes of production (synthesis and destruction) of components such that these components: (i) continuously regenerate and realize the network that produces them, and (ii) constitute the system as a distinguishable unity in the domain in which they exist. (VARELA, 1992). (Um sistema autopoético é organizado (definido como unidade) como uma rede de processos de produção (síntese e destruição) de componentes de tal modo que estes componentes: (i) regeneram e dão forma física continuamente à rede que os produz, e (ii) constituem o sistema como uma unidade distinta no domínio em que existem). Todos esses conceitos relacionados a embodied mind e autopoiese serviram de base para moldar o enativismo como o conhecemos hoje.

Desde então, as sucessivas tentativas de deslocamento para longe dos paradigmas representacionalista e computacionalista renderam uma série de abordagens que transformaram o enativismo em um aglomerado de ideias organizadas em quatro conceitos. (ROWLANDS,

2010) atribui ao filósofo Shaun Gallagher o termo 4E conception of the mind (concepção 4E da mente) para se referir a esses conceitos. São eles:

i - Embodied: pode ser traduzida como corporificada. Desenvolvida primeiramente no já citado livro *The Embodied Mind*, é a ideia de que a cognição não está apenas presente como representação no cérebro, mas nos processos corporais, sensoriais e motores. ii - Embedded: pode ser entendida como a mente “embutida” ou “imersa”. É a ideia de que processos mentais, por serem resultado de evolução por seleção natural submetidos a uma pressão ambiental específica, funcionam de maneira ótima apenas dentro desse contexto ambiental (ROWLANDS, 2010). A capacidade de adaptação totalmente versátil do agente dentro do ambiente se dá pois todas as restrições do ambiente vêm primeiro, e portanto a evolução e a diferenciação do corpo só são possíveis dentro daquele ambiente, e estão atreladas ao mesmo. iii - Enacted: é a ideia de que a mente não só é constituída por processos neurais, mas pela maneira como o organismo age ou reage em relação ao mundo, e pela maneira que o mundo reage à ação do organismo (ROWLANDS, 2010). A mente é um fenômeno performado que compreende essas ações e reações. iv - Extended: é um conceito que afirma que a mente não é um fenômeno localizado unicamente dentro do sistema fechado que constitui o organismo, mas que também se estende de várias maneiras para o ambiente. Pode ser visto como em (CLARK; CHALMERS, 1998), que encara o uso de ferramentas como veículos extensores da mente.

3.2 IA corporificada (Embodied AI)

Uma nova tentativa de construir agentes em que a inteligência emerge de maneira intrínseca e genuína se deu no final da década de 80 e começo da década de 90, com Rodney A. Brooks. Em (BROOKS, 1991) o autor descreve quatro fundamentos usados para a construção de seus robôs, são eles: Situação, Corporificação, Inteligência e Emergência (Situating, Embodiment, Intelligence and Emergence).

Como ele explica em seu artigo, *situating* é a propriedade de o robô estar situado lidando diretamente com o mundo, e não com descrições abstratas dele. *Embodiment* significa que o robô tem corpo e tanto experimenta o mundo de forma direta quanto suas ações podem modificar o mundo e promover feedback na sua percepção. *Intelligence* seria a propriedade de o robô ter comportamentos percebidos como inteligentes, mas a fonte dessa inteligência não vem apenas da computação, mas da localização do robô no mundo e das modificações nos sensores. *Emergence* seria a propriedade de a inteligência emergir indiretamente do conjunto de interações

com o mundo e entre os componentes.

A maneira que Brooks encontra para atingir esses fundamentos é usando o que passou a ser chamado de laço sensório-motor. Basicamente os sensores do robô são proximamente relacionados com motores que modificarão a posição ou estado do robô no mundo sem ter que passar por uma representação centralizada. A esse conjunto de componentes sensor-motor é dada uma ordem de prioridade. O motor é programado para agir de uma determinada maneira a depender do que foi captado pelo sensor relacionado, modificando a localização e o estado do robô no mundo, e logo modificando o que é agora captado pelo sensor, reiniciando o ciclo.

Esse modelo passou a ser conhecido como embodied AI (ou Heideggerian AI segundo Dreyfus), que pode ser traduzido como IA corporificada, e que mais tarde ganhou um conjunto de princípios de design, que pode ser visto em (PFEIFER *et al.*, 2005). É importante frisar que o primeiro princípio: "Synthetic methodology - Understanding by building" (Metodologia sintética - Compreensão através da construção) é de especial importância não só para a embodied AI mas para o uso da IA dentro dos estudos da ciência da cognição como um todo. Como (FROESE; ZIEMKE, 2009) explica, esse princípio, apesar de não excluir totalmente essas novas abordagens para a IA como sendo esforços de engenharia, propõe como finalidade última o uso dessa IA de maneira científica, dado seu potencial de nos ajudar a entender os fenômenos naturais da vida e da mente.

A maior contribuição da IA corporificada foi ter projetado um modelo computacional que faz emergir comportamento inteligente sem representação explícita a priori do mundo. Como o próprio Brooks diz, "representations and models of the world simply get in the way. It turns out to be better to use the world as its own model." (BROOKS, 1991) (representações e modelos do mundo simplesmente ficam no nosso caminho. No final das contas é melhor usar o mundo como seu próprio modelo).

(FROESE; ZIEMKE, 2009) tece críticas a este modelo, propondo construir o framework para o que chama de IA enativa. De modo geral, Froese e Ziemke afirmam que falta à IA corporificada projetar agentes genuinamente autônomos, ou seja, cujos objetivos, organização e respostas sejam intrínsecos a eles próprios.

(HUTTO; MYIN, 2012) utiliza os avanços na IA corporificada como provas de conceito contra teorias intelectualistas para a mente. No capítulo 3, os autores descrevem o intelectualismo como a posição que considera que qualquer ação inteligente necessita, como mínimo requerimento, que haja pensamento baseado em representações com conteúdo. Eles

resumem o raciocínio da seguinte maneira: se toda ação inteligente envolve planejamento, e todo planejamento envolve computação e representação, então “temos uma má notícia para os defensores do enativismo radical”. Os robôs de Brooks, bem como outros projetos na área de IA corporificada, alegam conseguir produzir comportamento inteligente sem a necessidade de representações ou modelos internos, Hutto e Myin utilizam então esses avanços para argumentar que, diferente do que os intelectualistas acreditam, muitos - talvez a maior parte - dos fenômenos cognitivos obedeçam aos mesmos princípios da inteligência de tais robôs.

Devemos ter cautela antes de assumirmos como verdade que os robôs da IA corporificada são prova suficiente para descartar a necessidade da existência de representação com conteúdo nas mentes básicas. Primeiro porque o debate sobre se há representação implícita ou não nesses robôs ainda não é ponto pacífico, pois uma máquina invariavelmente deverá ter que transmitir informação entre componentes utilizando corrente elétrica e transistores. Segundo porque os laços sensório-motores da máquina foram refinados extrinsecamente a fim de produzir comportamento inteligente do ponto de vista de seu projetista (portanto, há conteúdo mental de um projetista sendo atribuído à máquina). No entanto, essa discussão é suficiente para sintetizar a importância do princípio de design de “Metodologia sintética - Compreensão através da construção”, citado anteriormente, em que a IA pode ser usada como um importante instrumento para fornecer novas percepções sobre a inteligência.

3.3 IA Enativa de Froese e Ziemke

Em 2009, Tom Froese e Tom Ziemke publicaram o artigo *Enactive artificial intelligence: Investigating the systemic organization of life and mind*, que definiria a Inteligência Artificial Enativa. Apoiados sobre uma base teórica advinda da filosofia fenomenológica e da biologia teórica, eles argumentam que a IA corporificada, apesar de se estabelecer como uma alternativa válida à GOFAI, ainda falha em produzir cognição genuína por não possuir agência intencional como a observamos nos seres vivos. Segundo eles, os problemas resolvidos pelos agentes na IA corporificada continuam sendo relevantes para seus desenvolvedores, e não para os agentes em si. Uma analogia abordada no artigo que ilustra o ponto é o funcionamento de um termostato clássico, que funciona de maneira acoplada ao ambiente com seus componentes físicos produzindo um comportamento orientado a um objetivo através de feedback negativo. A maneira que interpretamos tal comparação seria a de que projetar o mecanismo que gera tal feedback e em consequência tal comportamento, assim como avaliar se o comportamento

atinge os objetivos pretendidos é papel do engenheiro (portanto, os objetivos são impostos extrinsecamente). Seguindo este raciocínio, concluímos que a IA corporificada em linhas gerais seguiria o mesmo tipo de lógica, mas com laços sensório-motores mais sofisticados. E mesmo abordagens que utilizam algoritmos evolucionários para ajustar o funcionamento da dinâmica interna dos componentes sensório-motores ainda avaliam os objetivos a serem atingidos de maneira extrínseca.

(FROESE; ZIEMKE, 2009) faz menção às críticas de (DREYFUS, 2007) para dizer que a IA corporificada não chegou a resolver o Frame Problem como proposto por (DENNETT, 1978). Eles escrevem: “Dreyfus claims that the field still has not managed to properly address the problem of meaning in AI. Moreover, as long as there is no meaningful perspective from the point of view of the artificial agent, which would allow it to appropriately pick up relevance according to its situation in an autonomous manner, such a system cannot escape the notorious ‘frame problem’ as it is described by Dennett” (Dreyfus afirma que a área [da IA corporificada] ainda não conseguiu endereçar propriamente o problema da significância na IA. Além disso, enquanto não houver uma perspectiva significativa do ponto de vista do agente artificial, que o permitisse captar, de maneira autônoma, relevância de acordo com sua situação, tal sistema não escapa ao notável “frame problem” como descrito por Dennett). O que faz sentido, pois como poderíamos afirmar que um agente está de fato limitando o que é relevante para ele de ser atualizado após uma ação, se nada é relevante para tal agente (e sim para seu projetista)? Se por um lado, o problema criado pela IA simbólica — de que existem infindáveis representações simbólicas de fatos dentro de um contexto que não são relevantes para o ser cognitivo — é resolvido pela IA corporificada simplesmente não lidando com representações simbólicas do mundo (mas com o próprio mundo), por outro lado, essa forma de lidar com o mundo continua não sendo relevante para o agente em si, e o que é relevante ou não continua sendo escolhido arbitrariamente por um projetista, apenas de maneira diferente de como faz a IA simbólica. Como fazer emergir significância de maneira intrínseca é o problema para o qual a IA enativa busca encontrar resposta.

Para os autores, e para o projeto enativista tradicional, duas condições necessárias para a agência intencional são autonomia e adaptatividade. Froese e Ziemke resgatam (VARELA, 1992), em que, como já explicamos na seção sobre enativismo, atribui aos sistemas autopoieticos o surgimento da intencionalidade, e (VARELA, 1979), em que ele define uma noção mais geral de autonomia: We shall say that autonomous systems are organizationally closed. That is, their

organization is characterized by processes such that 1. the processes are related as a network, so that they recursively depend on each other in the generation and realization of the processes themselves, and 2. they constitute the system as a unity recognizable in the space (domain) in which the processes exist. (Diremos que sistemas autônomos são organizacionalmente fechados. Isto é, a organização deles é caracterizada por processos tal que: 1. Os processos são relacionados como uma rede, tal que eles dependem recursivamente um do outro na geração e realização deles próprios; 2. Eles constituem o sistema como uma unidade reconhecível no espaço (domínio) no qual os processos existem). Froese e Ziemke se referem a sistemas organizacionalmente fechados como constitutivamente autônomos, o que seria uma generalização do conceito de autopoiese. Portanto, para o projeto enativista, intencionalidade tem como condição necessária a atividade de um sistema constitutivamente autônomo.

Os autores explicam que o enativismo utilizar o conceito de sistema constitutivamente autônomo da biologia capturaria a ideia de (JONAS, 2001), em que é explicado que a diferença entre um mecanismo de feedback (como um termostato por exemplo) e um ser vivo é que, se o mecanismo para de funcionar, ele continua existindo. Esse mecanismo existe materialmente por si só, o que ele chama de being by being (ser por ser). Já um ser vivo é a rede de processos em si. Se os processos param de funcionar a vida deixa de existir, e isto ele chama de being by doing (ser por fazer). Logo um ser vivo teria teleologia intrínseca.

Teleologia seria, em termos aristotélicos, a causa final. Por exemplo, a causa da existência de uma faca é sua finalidade. Ela só existe pois existia uma finalidade (cortar), e um objeto (faca) foi construído para satisfazer essa finalidade. Esta seria uma teleologia extrínseca, pois a finalidade não pertence ao objeto. Já nos seres vivos há uma circularidade causal, pois a causa de uma rede de processos constitutivamente autônoma existir é a própria existência da rede de processos. A existência de um agente autônomo e a constante criação da sua própria identidade são sinônimos, o que leva os autores a chegarem à conclusão de que a autonomia constitutiva é necessária para a teleologia intrínseca. Em outras palavras, para que emerja em um agente objetivos próprios a este agente, sua existência deve ser vinculada a um contínuo processo que tem por finalidade continuar existindo. Objetivos secundários seriam todos derivações indiretas da contínua luta pela própria existência.

Um vez explicada a necessidade da autonomia para a agência intencional, nos voltamos agora para a adaptatividade. O artigo apresenta o argumento de que o conceito de autonomia constitutiva não possui gradação, de modo que bastaria um evento não destruir

por completo um sistema autônomo para que este evento fosse classificado como bom para este sistema, quando na verdade, para evitar degradação, um sistema deve ter a capacidade de lidar com alterações internas e externas e melhorar sua condição de sobrevivência de maneira espontânea.

Tomando o exemplo clássico da bactéria em um gradiente de sacarose de (VARELA, 1992), (PAOLO, 2005) explica que se a bactéria não for exposta a um evento que a destrua completamente, ela não precisa consumir a sacarose pois continuará existindo. No entanto, não consumir a sacarose a levará a uma gradual degradação que a destruirá irremediavelmente em um futuro próximo. Perceba que o fato de a bactéria utilizar sacarose como nutriente faz com que nós apontemos a relevância da sacarose para a bactéria. No entanto, a conservação da identidade proposta pela teoria da autopoiese seria insuficiente para explicar o que leva a bactéria a buscar seu nutriente, já que como explicado, para manter sua identidade é suficiente que ela fique inerte enquanto gradualmente se degrada. Seria preciso algo a mais para explicar tanto intencionalidade quanto o comportamento da bactéria (por que a bactéria apontaria para a sacarose como relevante de uma maneira tão específica?). A resposta, segundo Di Paolo, seria adaptatividade. Organismos teriam uma extensão de estados em que sua sobrevivência é viável (o viability set, ou conjunto de viabilidade), e têm a capacidade de reagir a perturbações não-fatais que alterem suas condições de existência dentro desse espaço de viabilidade, de modo a melhorar espontaneamente essas condições. (FROESE; ZIEMKE, 2009) recorrem a este trabalho de 2005 de Di Paolo, citando sua definição para adaptatividade: A system's capacity, in some circumstances, to regulate its states and its relation to the environment with the result that, if the states are sufficiently close to the boundary of viability, 1. Tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence, 2. Tendencies of the first kind are moved closer to or transformed into tendencies of the second and so future states are prevented from reaching the boundary with an outward velocity. (A capacidade de um sistema, em certas circunstâncias, de regular seus estados e suas relações com o ambiente com o resultado de que, se os estados estiverem próximos o suficiente dos limites de viabilidade, 1. As tendências são distinguidas e ações são tomadas sobre elas a depender se os estados se aproximam ou se afastam do limites e, como consequência, 2. Tendências do primeiro tipo são movidas para perto de ou transformadas em tendências do segundo tipo de modo que estados futuros são impedidos de atingir o limite com uma velocidade para fora). É concluído que esta definição de adaptatividade é condição necessária para sense-making (o “fazer sentido” que um

agente tem de elementos do ambiente).

Essas duas condições são resumidas no artigo pela tabela:

Tabela 1 – Requerimentos de sistema - IA enativa

	System requirements (SR)	Implicações	Normatividade
SR-1	Autonomia constitutiva	Teleologia intrínseca	Uniforme
SR-2	Adaptatividade	Sense-making	Gradual

Fonte: Froese and Ziemke 2009

O artigo combina as duas implicações afirmando que ter autonomia constitutiva é formar a própria identidade que constitui uma relação com o ambiente, delimitando o domínio de ações possíveis nele, estabelecendo assim as bases para a regulação adaptativa. A autonomia é portanto a base para construir uma “perspectiva” do ambiente, logo também é necessária para sense-making.

Os autores finalizam a base teórica citando os cinco passos para a cognição de (THOMPSON, 2004): 1. Vida = autonomia constitutiva + adaptatividade 2. Autonomia constitutiva implica emergência de uma identidade 3. Emergência de uma identidade adaptativa implica emergência de um mundo 4. Emergência de identidade adaptativa e mundo = sense-making 5. Sense-making = cognição.

E completam: “...as Thompson points out, these steps amount to an explicit hypothesis about the natural roots of intentionality. In other words, they form the basis of the claim that the ‘aboutness’ of our cognition is not due to some presumed representational content that is matched to an independent external reality (by some designer or evolution), but is rather related to the significance that is continually enacted by the precarious activity of the organism during its ongoing encounters with the environment.” (“Como Thompson aponta, esses passos correspondem a uma hipótese explícita sobre as raízes naturais da intencionalidade. Em outras palavras, eles formam a base da afirmação de que a nossa cognição ser ‘sobre algo’ não se deve a um suposto conteúdo representacional pareado a uma realidade externa independente (por algum projetista ou pela evolução), mas está na verdade relacionado à significância que é continuamente “performada” pela atividade precária do organismo durante seus contínuos encontros com o ambiente) (nota: enact é uma palavra de difícil adaptação para o português, assim performar pode não traduzir completamente o sentido pretendido pelo autor). Importante notar por essa citação que já se falava dentro do enativismo de uma história natural para a intencionalidade sem conteúdo representacional. No entanto é importante também deixar claro que em momento ne-

nhum o projeto enativista tradicional fala de um tipo de intencionalidade que não possua nenhum tipo de conteúdo. Sense-making é um termo que trata da emergência de significado, portanto da emergência de conteúdo, porém um conteúdo que não seja necessariamente representacional. É neste ponto que está a principal discordância entre o enativismo radical e o enativismo de Varela, Thompson e Di Paolo (que (HUTTO; MYIN, 2017) vem a chamar de “enativismo autopoietico adaptativo”). Discutiremos com mais detalhes essas diferenças nas sessões adiante.

A partir dos dois requerimentos SR-1 e SR-2, os autores formulam dois princípios para o projeto da IA enativa, são eles: Enactive AI design principle 1 (EAI-1): O sistema deve ser capaz de gerar sua própria identidade sistêmica em algum nível de descrição; Enactive AI design principle 2 (EAI-2): O sistema deve ter a capacidade de regular ativamente sua interação sensorimotora em andamento em relação a uma limitação de viabilidade (viability constraint). EAI-1 traduz o princípio da autonomia. Para um sistema ser verdadeiramente autônomo, o artigo descreve que deve existir uma engenharia de emergência de segunda-ordem, ou seja, deve-se projetar um ambiente que faça surgir um agente que por sua vez, quando acoplado com esse mesmo ambiente, produz o comportamento desejado. Como o surgimento de um ser é anterior à sua evolução, design explícito e algoritmos evolutivos estão fora de questão, sendo por este motivo chamado pelos autores de the hard problem of enactive AI. EAI-2 traduz o princípio da adaptatividade. O artigo explica que este princípio poderia ser interpretado de duas maneiras. A primeira maneira é com uma viability constraint imposta extrinsecamente, de modo que um agente artificial construído dessa maneira poderia ser usado para isolar e estudar fenômenos dinâmicos de adaptabilidade, como a homeostase. Mas caso o sentido seja o de uma viability constraint que emerja de maneira intrínseca ao agente, essa constraint deriva de EAI-1 e da identidade do agente, complementando o problema anterior e tornando o hard problem of enactive AI ainda mais difícil. Os autores afirmam que estes princípios e os fundamentos nos quais se baseiam (autonomia e adaptatividade) são necessários para cognição genuína, mas nada garante que sejam suficientes.

Se prosseguirmos de onde esses fundamentos param e continuarmos a comparação com a biologia observamos que funções biológicas que regulam os processos de reprodução parecem contradizer o fundamento da adaptatividade. Se por um lado a simples autonomia não explica as funções biológicas que parecem otimizar o comportamento do ser vivo em busca de se evitar sair do seu espaço de viabilidade (e portanto a capacidade adaptativa precisa ser levada em consideração), por outro lado apenas autonomia e adaptação parecem não ser suficientes para

explicar comportamento reprodutivo, já que por vezes a reprodução exige gastos de energia que invariavelmente tendem a empurrar o ser vivo em uma direção que, nas últimas consequências, o colocaria para fora de seu espaço de viabilidade. No entanto, o comportamento reprodutivo é perfeitamente explicado pela seleção natural, pois um ser só vem a existir se seu gerador e semelhante tiver as capacidades de produzi-lo. E por ser um semelhante, este ser que foi gerado muito provavelmente herdará tais capacidades reprodutivas do seu gerador (e se não herdar, eventualmente sua particular incapacidade reprodutiva será apagada da existência).

Note, no entanto, que explicar as origens naturais da cognição e buscar fundamentos necessários e suficientes para que o fenômeno da cognição aconteça são coisas diferentes. Podemos supor por exemplo um modelo virtual de mundo que possua a propriedade de gerar continuamente e espontaneamente uma gama de seres autônomos e adaptativos que seguem a lógica do being by doing sem necessariamente terem que se reproduzir. Se o hard problem of enactive AI como proposto por (FROESE; ZIEMKE, 2009) fosse resolvido, esse poderia ser um dos exemplos de simulações possíveis. Por isso que precisamos não só de uma história natural para a origem da intencionalidade, mas também esquematizar a composição da mesma, de modo que possamos chegar o mais próximo possível de reproduzi-la. Por esse motivo que a ideia deste artigo de Froese and Ziemke é importante. Dentro do contexto deste trabalho, suas ideias nos fazem lembrar que mesmo que nos debruçemos sobre o enativismo radical em busca da história natural do surgimento das formas mais básicas de cognição, o trabalho da IA não acaba aí. Precisamos pinçar dessa história natural quais características são dispensáveis e quais são indispensáveis para que o fenômeno da cognição aconteça. O artigo também se mostra útil pelo compilado de informações e história do pensamento enativista.

4 O PROBLEMA

Agora que já fornecemos uma base histórica e filosófica da IA, do enativismo, e da relação entre ambos, podemos iniciar a abordagem do problema apresentando mais detalhadamente a motivação por trás do trabalho. Começaremos pelo questionamento mais básico: o que significa ser inteligente? Esta não é uma pergunta fácil de ser respondida. O modelo de inteligência que utilizamos quase que intuitivamente como parâmetro é o humano, nossa capacidade de adaptação, raciocínio, análise, ponderação, produção, etc. De maneira geral, podemos dizer que a inteligência humana é constatada ao se observar todo o aparato cognitivo que possuímos à nossa disposição para resolver problemas, sendo provavelmente o aparato mais complexo entre todos os seres ou agentes que tivemos a chance de analisar.

No entanto, isolar um tipo de agente (no caso, o mais complexo que temos conhecimento), e afirmar que o mesmo manifesta inteligência não é suficiente para responder a algumas de nossas perguntas: i - Se partimos da única certeza de que o ser humano é o padrão de agente inteligente, isso significa que a busca por uma inteligência artificial genuína estaria destinada a tentar se aproximar cada vez mais de uma cópia da inteligência humana? ii - É conceitualmente possível utilizar um computador para simular uma cópia perfeita de inteligência humana, sabendo que o mesmo é estruturado de maneira fundamentalmente diferente de nossos corpos e cérebros? E mesmo que seja possível, é viável? iii - Caso a busca não se limite a copiar a inteligência humana, mas a tentar desenvolver uma inteligência tão adaptativa e complexa quanto, como saberemos que um agente diferente é também inteligente, se não temos uma definição propriamente dita para inteligência? Estaríamos fadados a reconhecer a inteligência apenas onde há resolução de problemas úteis para nós? iv - Se pudermos reconhecer níveis menos complexos de inteligência em outros tipos de agente, poderíamos encarar a adaptatividade de outros seres biológicos como formas de inteligência, pois os mesmos resolvem seus problemas vinculados à luta pela sobrevivência? Uma árvore seria inteligente? E uma bactéria?

Como mostramos, os primórdios da IA já contavam com a ideia de poder replicar a mente humana usando um computador, ideia que foi atacada ao longo da história de diferentes maneiras, com críticas sendo feitas às bases filosóficas que apoiavam a maneira com que as pesquisas eram conduzidas. Em (DREYFUS, 2007), por exemplo, após o autor afirmar que a melhor tentativa de se criar uma IA genuína até então foi com a IA corporificada, ele conclui que consertar essa IA corporificada para torná-la mais Heideggeriana ao ponto de simular com eficácia a inteligência humana significaria programar além de uma simulação cujo modelo

cerebral reagisse com exatidão ao que é relevante para nós, também programar um agente cuja maneira de estar imerso e corporificado experimente o que é significativo para nós da mesma maneira particular que experimentamos. Para ele, isso significaria o seguinte: “we would have to include in our program a model of a body very much like ours with our needs, desires, pleasures, pains, ways of moving, cultural background, etc” (Nós teríamos que incluir no nosso programa um modelo de corpo tal qual o nosso, com nossas necessidades, desejos, prazeres, dores, maneiras de se mover, antecedentes culturais, etc). Como ele mesmo conclui, isto parece improvável de ser realizado. Mesmo as técnicas mais avançadas de IA atuais, incluindo deep learning, ainda possuem um problema de anterioridade, o que significa que os objetivos dos algoritmos e os resultados considerados úteis ou inteligentes estão invariavelmente ligados ao que nós humanos consideramos que fazem sentido para nós. Não estamos perto de uma inteligência artificial que seja autônoma, robusta, flexível e complexa como a humana.

Consideramos que esta visão de Dreyfus sobre como seria construir uma cópia da inteligência humana concorda com o que defendem paradigmas enativistas para a cognição, na medida em que para os enativistas, a cognição trata-se de como e porque um corpo interage com um ambiente da maneira particular que interage. O próprio termo “inteligência” escapa ligeiramente do vocabulário enativista usual, que discorre mais sobre adaptatividade, autonomia e cognição.

Portanto, respondendo às perguntas: i - Precisamos de uma história natural de como surgem ações inteligentes para não ficarmos presos a achar que um agente que manifesta inteligência genuína precisa agir como um humano, precisamos saber onde e como nasce a cognição em sua forma mais básica; ii - Fugiremos do funcionalismo de propor um caminho de pesquisa que tente simular por computador uma cópia exata da mente humana, mas sim usar a inteligência artificial como ferramenta para em primeiro lugar entender como fenômenos cognitivos funcionam; iii - Não só temos que encontrar uma história natural para o surgimento da cognição, mas também definir que condições são necessárias e suficientes para que um agente seja considerado cognitivo a despeito do que nós humanos consideramos inteligente ou útil para nós (devemos recorrer portanto ao que seriam objetivos intrínsecos); iv - Dessa forma, a luta pela sobrevivência de um ser vivo é um objetivo intrínseco, e suas funções biológicas realizam ações inteligentes nesse sentido. Entender este fenômeno está entre os objetivos do enativismo.

Tentaremos entender a cognição e como um agente possuir cognição está relacionado ao mesmo possuir objetivos intrínsecos. Motivo pelo qual não faz sentido falar em inteligência

genuína para agentes cuja estrutura é determinada por objetivos definidos de maneira extrínseca, como no exemplo clássico de um programa que joga xadrez e é desenvolvido de maneira tal a cumprir seu objetivo (definido por um humano) de ganhar o jogo. Por esse motivo, o paradigma enativista, que parte de uma ideia bastante esclarecedora relacionada a autonomia (a autopoiese), foi escolhido como um dos pilares deste trabalho, e quando nos referirmos a “inteligência genuína” no âmbito da IA, estaremos falando da possibilidade de fazer emergir cognição genuína ou de usar a IA para melhor entendê-la.

Obedecendo a uma abordagem ascendente (bottom-up), o ponto de partida da pesquisa deve ser o de investigar qual o fenômeno observável menos complexo possível na natureza conseguiríamos categorizar como possuindo cognição. Este é o mesmo ponto de partida de (HUTTO; MYIN, 2012), no livro *Radicalizing Enactivism: Basic Minds without Content*. Neste livro em tom de manifesto, os autores advogam pela existência de um nível de cognição básica onde não há conteúdo (portanto, de uma intencionalidade sem conteúdo).

Tendo em vista que o ponto de partida do enativismo radical coincide com o que consideramos o ponto de partida mais lógico ao investigar o fenômeno da inteligência. E também considerando que “cognição sem conteúdo” e “intencionalidade sem conteúdo” são posicionamentos instáveis, intrigantes e polêmicos nas ciências cognitivas. Consideramos de grande importância a análise desse posicionamento. Como já explicado na introdução, entender essas ideias é a motivação e o objetivo deste trabalho. Daremos prosseguimento buscando um entendimento mais claro do que seria intencionalidade, conteúdo, e o que se quer dizer com “intencionalidade sem conteúdo”.

4.1 Intencionalidade

A maneira que a palavra “intencionalidade” é usada naturalmente ou vulgarmente tem relação com intenção, propósito ou objetivo, no entanto não é desse significado que trata a filosofia. Como falado na introdução deste trabalho, a definição mais ampla e abstrata na filosofia fala da intencionalidade como sendo a propriedade de ser sobre ou de alguma coisa (sendo essa propriedade a marca dos fenômenos mentais).

(JACOB, 2014) cita Franz Brentano como o responsável por recuperar da escolástica esse conceito, e o coloca como precursor das discussões a cerca da definição de intencionalidade, que acontecem desde o final do século 19 até a atualidade. (BRENTANO, 1874) fala da inexistência intencional de um objeto, ou seja, a mente faria referência a um conteúdo ou

objeto não necessariamente material. O mental portanto seria definido por conter esses objetos intencionalmente dentro de si.

Diversas linhas que sucederam Brentano formulam a intencionalidade como a marca do mental, e tentam relacionar estados mentais a objetos ou conteúdos a que esses estados se referem. O problema com linhas de pensamento desse tipo é que ou não há uma definição apropriada do que seria mente ou estado mental para servir de apoio à definição de intencionalidade, ou a definição de mente em si se apoia na de intencionalidade. É o caso do representacionalismo, que parte do princípio de que “representações são estados mentais com intencionalidade” (PITT, 2017).

Uma definição natural de intencionalidade é importante pois se a intencionalidade é a marca do mental, uma inteligência artificial genuína terá agência intencional. Por ser uma definição natural, qualquer linha que não tente chegar no intencional/mental a partir de um vocabulário não-intencional não interessa a esse trabalho.

A tarefa de partir de um vocabulário não-intencional configura-se como um desafio pois um dilema surge ao tentar naturalizar intencionalidade: estados mentais parecem ter poder causal sobre o mundo material. (DEACON, 2012) discute esse problema chamando a atenção para o que Aristóteles classificou como causa final, ou teleologia. A causa da construção de uma ferramenta, por exemplo, é sua finalidade. Em outras palavras, uma ferramenta só é construída, ela só existe, porque servirá a um propósito, e este propósito é um estado mental. Isto configura uma espécie de inversão temporal de causa e efeito (a causa neste exemplo é um estado futuro ainda não existente, um estado em que a ferramenta está sendo usada para o propósito a qual foi projetada). Para Deacon, foi da vida e posteriormente da mente que emergiu essa inversão temporal de causa e efeito. No entanto ele não se aprofunda na discussão filosófica de intencionalidade, e chega a cunhar outro termo mais geral - entencionalidade - para designar esse regime de causa final observado nos seres vivos (como em funções biológicas e atividade celular por exemplo), e que abarca o conceito de intencionalidade. Apesar do termo entencionalidade parecer uma fuga de uma discussão mais profunda sobre intencionalidade, ele nos dá uma indicação de que a busca por naturalizar o conceito deva se apoiar na biologia (tanto para buscar intuições quanto para definir que qualquer que seja a definição de intencionalidade, seres biológicos devem possuí-la). Deacon também traz uma visão que seguiremos na pesquisa, que vai contra os casos mais extremos de eliminativismo materialista (que abandona por completo qualquer ideia de intencionalidade). Ele fala: “We need to explain the ententional nature

of our own existence, not explain it away” (DEACON, 2012) (em tradução livre, seria algo como: "Precisamos explicar a natureza intencional de nossa própria existência, não explicar sua eliminação"). Negar o eliminativismo materialista e cunhar o termo intencionalidade para designar de maneira geral a teleologia presente nos seres vivos são dois movimentos que enxergamos como importantes pelos seguintes motivos respectivamente.

Primeiro, intencionalidade é um termo antigo, com raízes em um pensamento dualista como pode-se verificar pelas definições clássicas acima. A primeira reação de um cientista dedicado a estudar fenômenos naturais, portanto, poderia ser de negar a utilidade de tal termo, classificando-o como anti-científico ou sem significado relevante, afinal de contas as ciências naturais não trabalham sob a ótica de "causa final", como se houvesse um motivo por trás da existência de objetos e fenômenos do mundo que não pudesse ser explicado puramente como "causa eficiente"(para colocar em termos aristotélicos), ou seja, como resultado de uma configuração de mundo submetida às leis da física. No entanto, o argumento contra o eliminativismo aponta que há um fenômeno no mental que é indiscutivelmente diferente do mundo inorgânico, e que parece motivar os agentes possuidores de mente a direcionar o mundo para configurações que não têm necessariamente uma relação direta com o jogo de causa e efeito estudado pela física, mas sim com objetivos ou estados mentais que parecem ausentes do mundo físico. O ponto é que não precisamos por obrigação nos prender a definições antigas de intencionalidade que não teriam espaço dentro da filosofia naturalista ou da ciência, mas sim entender que o fenômeno observado, que deu origem ao termo, existe, e nós podemos tomar para nós a tarefa de tentar explicá-lo. Também podemos nos dar a liberdade de continuar chamando-o de intencionalidade. Naturalizar intencionalidade é apontar para o fenômeno que primeiramente originou as intuições que a definiram na filosofia e explicar tal fenômeno de maneira consistente com as ciências naturais.

O segundo motivo, relacionado ao termo "intencionalidade" usado por (DEACON, 2012), é o de que este termo foi criado para abarcar toda a teleologia observada no mundo natural. O fenômeno de causa final presente na mente humana apresenta as mesmas características dos fenômenos observados em todo tipo de vida terrestre, mais precisamente quando tratamos de funções biológicas, e não está presente nos fenômenos inorgânicos. Uma função biológica parece ter uma finalidade que diz respeito à sobrevivência do ser vivo. Esta é a base da intuição para buscar as origens da intencionalidade nos seres vivos, e atrelá-la a toda forma de cognição (e não só à mente humana).

Seguindo esse deslocamento para a biologia, um trabalho importante foi (VARELA, 1992), que foca na biologia da intencionalidade e argumenta que o acoplamento da unidade autopoietica com o ambiente é responsável pelo surgimento de intencionalidade. Para Varela, um sistema autopoietico ao se definir como uma unidade cria ao mesmo tempo uma perspectiva do que está fora dessa unidade. Essa perspectiva ele chama de mundo e o diferencia de ambiente no sentido que o mundo para a unidade autopoietica foi “definido no mesmo movimento que originou a identidade [dessa unidade] e que só existe nessa definição mútua”. Ele cita o exemplo de um ambiente que consiste de um gradiente de sacarose, que, sem uma bactéria, está sujeito apenas às leis físico-químicas, mas com a presença da bactéria passa a ter uma significância especial no sentido de que a existência da bactéria em si aponta para a sacarose como um item importante no acoplamento bactéria-ambiente para a manutenção de tal unidade autopoietica (pois a sacarose serve como alimento). Para Varela, esse tipo de relação é responsável pelo surgimento da intencionalidade. Tal qual Varela, os proponentes do enativismo radical também veem na biologia a chave para explicar o surgimento da intencionalidade. No entanto, o enativismo autopoietico e o enativismo radical não concordam quanto à questão do conteúdo. A posição de que existe intencionalidade sem conteúdo é unicamente do enativismo radical. Nos aprofundaremos nessas questões mais adiante. Por ora, é suficiente compreender de onde nasce o conceito de “intencionalidade”, que motivos temos para querer investigar propostas de naturalização para este conceito, e o que nos faz buscar na biologia tal naturalização.

4.2 Conteúdo

Conteúdo mental se refere diretamente ao sentido clássico de intencionalidade. A entrada para Externalism about mental content (externalismo sobre conteúdo mental) da enciclopédia de filosofia de Stanford (LAU; DEUTSCH, 2016) coloca que um estado mental intencional seria aquele que possui um conteúdo mental. Logo, “querer que esteja chovendo” e “acreditar que está chovendo” seriam estados mentais psicologicamente diferentes, mas com o mesmo conteúdo. Esta visão clássica de conteúdo é indissociável das visões clássicas para a intencionalidade e para o mental. O enativismo radical se opõe frontalmente a esta posição representacionista de que toda e qualquer atividade mental pressupõe conteúdo representacional, e parece fazer isso não dissociando conteúdo de representação, mas sim atacando o conceito mais aceito de intencionalidade, de modo a reformulá-la para que suas manifestações mais básicas se apresentem sem que haja conteúdo sendo representado. (HUTTO; MYIN, 2017)

apresenta a seguinte visão resumida para o que concebem por “conteúdo”: “to be in a contentful state is to take (‘represent,’ ‘claim,’ ‘say,’ ‘assert’) things to be a certain way such that they might not be so”. Tradução livre: “estar em um estado de conteúdo é tomar (‘representar’, ‘afirmar’, ‘falar’, ‘declarar’) coisas como sendo/estando de uma certa maneira, tal que elas podem não ser/estar de tal maneira”. Em outras palavras, o enativismo radical não necessariamente discorda do cognitivismo quanto à maneira mais básica de entender conteúdo: seria uma normatividade semântica que necessita de um veículo físico para ser representada, e que por ser uma normatividade, possui condições de verdade e possibilidade de ocorrer misrepresentation. Os pontos de discordância são em como conteúdo se manifesta (e por conseguinte, por que, e em que ponto da cadeia evolutiva, se manifesta), e em como devemos definir intencionalidade.

Três pontos importantes para essa discussão são: i - *The Hard Problem of Content*, discutindo em (HUTTO; MYIN, 2012), em que discorrem sobre o abismo existente entre o veículo físico da representação e o seu significado (o seu conteúdo); ii - Um esboço para a história natural da origem do conteúdo, presente em (HUTTO; SATNE, 2015), em que argumentam que o conteúdo é produto de um conjunto de práticas culturais que antes de tudo se manifestam externamente, para só depois serem internalizadas. iii - A diferença entre as concepções para “conteúdo” do enativismo radical e o enativismo tradicional (chamado por Hutto e Myin de “enativismo adaptativo autopoiético”). Abordaremos detalhadamente cada um destes tópicos na seção seguinte, que tratará propriamente do enativismo radical.

5 ENATIVISMO RADICAL

Daniel D. Hutto e Erik Myin, no livro intitulado *Radicalizing Enactivism: Basic Minds Without Content*, de 2012, fazem uma classificação do espectro de ideias dentro da ciência da cognição, e em especial, dentro do enativismo. Abaixo, algumas dessas classificações e explicações presentes no livro:

- **Content Involving Cognition (CIC)**. Considera que a cognição deve envolver conteúdo. Os CICs irrestritos também acreditam que a mente se trata do uso e manipulação de representações, e tomam a posição intelectualista de que toda representação necessariamente envolve conteúdo. Segundo (HUTTO; MYIN, 2012), os conceitos 4E são vistos pelo CIC como indesejados e retrógrados pois tratariam da cognição como mero comportamento.

- **Enactive/Embodied Cognition (CEC)**. Este posicionamento seria daqueles que admitem que estados e processos corporais podem assumir papéis cognitivos. (HUTTO; MYIN, 2012) encara este posicionamento como conservador, argumentando que esses papéis ainda são vistos pelo CEC como representações ou processos computacionais, com a diferença de que são exercidos por veículos corporais.

- **Sensorimotor Enactivism**. Considera que a atividade neural é necessária mas não suficientemente responsável pela cognição, e que percepção e ação estão intrinsecamente conectados na explicação da mesma. Estão preocupados em analisar a cognição como um engajamento concreto, estendido e interativo do organismo, e rejeitam o conceito de representação interna (HUTTO; MYIN, 2012).

- **Autopoietic Enactivism**. É a ideia de que o mental emerge das atividades de auto-organização e auto-criação dos seres vivos. (HUTTO; MYIN, 2012) cita (VARELA *et al.*, 1991) como o trabalho precursor desse paradigma. A grande contribuição dessa visão seria apresentar uma posição totalmente anti-representacionista e colocar processos organizacionais do organismo e pressões ambientais como igualmente importantes na emergência da cognição.

- **Radically Enactive/Embodied Cognition (REC)**. Esta é a posição em que (HUTTO; MYIN, 2012) se coloca. Assim como a visão sensorimotora, o REC entende a cognição como um fenômeno corporificado que organismos desempenham ativamente junto ao ambiente. E assim como a visão autopoietica (da qual é mais próxima), ela tenta abandonar totalmente o representacionismo.

Com a intenção de apresentar as diferenças do REC, (HUTTO; MYIN, 2012) argumenta que o enativismo sensorimotor não abandona o intelectualismo, pois esse tipo de

enativismo afirma que existe conhecimento implícito na experiência perceptiva, e que essa experiência tem conteúdo representacional. Também argumenta que diferente do enativismo autopoietico, o REC abandona a noção de sense-making (“fazer sentido”) para o que ele chama de basic mind (mente básica).

5.1 The Hard Problem of Content

Em sua obra de 2012, Daniel Hutto e Erik Myin apresentam o Hard Problem of Content (HPC), ou, em tradução livre, o problema difícil do conteúdo. Geralmente os problemas difíceis na filosofia, especialmente no que tange a naturalização de conceitos, são aqueles que demonstram que há um abismo tão grande entre a formulação de um conceito e a sua explicação, que no final das contas a melhor estratégia seria esquivar-se do problema reformulando tais conceitos, ou declarar que o problema é intratável.

Hutto e Myin escrevem que a formulação do HPC visa atingir três objetivos: i - Minar o desafio do processamento de informação, que seria o argumento de que rejeitar que mentes básicas não possuem conteúdo seria o mesmo que rejeitar que o mental é essencialmente processamento de informação; ii - Mostrar que os defensores do CIC ainda não conseguiram aterrar sua noção de conteúdo a nenhuma proposta naturalista para o surgimento do mesmo. iii - Colocar as bases para uma proposta positiva do REC sobre mentes básicas.

Como já argumentamos, o enativismo radical não nega que haja um caráter representacional no conteúdo, mas é exatamente por não rejeitarem esta ideia que os defensores do REC discordam dos cognitivistas quanto ao surgimento e à manifestação deste conteúdo representacional. O HPC seria então uma tentativa do REC de pressionar os cognitivistas para uma posição onde não haveria saída para a naturalização de conteúdo a não ser aceitar a proposta do REC.

Na primeira parte do problema, os autores usam duas noções defendidas pelo CIC. A primeira, retirada de (DRETSKE, 1981) seria a de que informação é uma commodity objetiva. Ou seja, a geração, transmissão e recepção de informação acontece no meio físico e não pressupõe nenhum processo interpretativo. A segunda noção seria a de que informação é covariação, e é colocada da seguinte maneira: “s’s being F ‘carries information about’ t’s being H iff the occurrence of these states of affairs covary lawfully, or reliably enough.” (HUTTO; MYIN, 2012). Traduzindo livremente, seria algo como: s’s sendo/estando F “carrega informação sobre” t’s serem/estarem H se a ocorrência desses estados apresenta covariância de forma certa, ou

suficientemente confiável. O argumento é de que não há nenhuma implicação que em algum momento explique de maneira naturalizada como informação-como-covariância se transforma em conteúdo informacional. Informação-como-covariância explicaria apenas o veículo físico em que o conteúdo está representado, mas não como aquele estado físico venha a ter sentido, propósito, ou condições de verdade para algum processo que o interpreta. Ou seja, não bastaria processar informação para possuir conteúdo. Eles nomeiam esta primeira parte do problema de “Covariance Doesn’t Constitute Content Principle” (Princípio da covariância não constituir conteúdo).

A segunda parte do problema seria a de admitir que conteúdo não é redutível a covariação, e portanto, o projeto para naturalizar conteúdo seria o de encontrar que outros fenômenos físicos além da covariância contribuem para o surgimento de conteúdo.

Eles prosseguem por afirmar que uma maneira de se esquivar do dilema seria o de postular uma definição para informação mais robusta do que simples covariância. Os autores citam o exemplo de informação-como-indicação, de Dretske, mas alegam que um estado indicar outro estado ainda se constituiria como mera relação entre estados físicos. Eles prosseguem afirmando que mesmo que falássemos de “indicação” como algo subjetivo, esta explicação ainda necessitaria de todos os detalhes de como um estado físico pode subjetivamente indicar algo a alguém.

Após críticas de que o HPC estaria usando dois pesos e duas medidas, pois o problema na verdade minaria até mesmo as tentativas do REC de naturalizar conteúdo, (HUTTO; MYIN, 2017) respondem a essas críticas afirmando que reduzir conteúdo a fenômenos físicos pode não ser frutífero quando apenas um naturalismo estrito, baseado apenas nas ciências duras, é levado em consideração. Os autores alegam que a alternativa seria basear-se em um naturalismo relaxado, que busca munção explanatória nas práticas sócio-culturais. Eles afirmam que o surgimento de conteúdo na natureza não é totalmente contínuo, mas sim um salto de capacidades justamente por conta de tais engajamentos sociais e culturais. A este fenômeno, eles dão o nome de Kinky Cognition. As bases desse pensamento foram postas em (HUTTO; SATNE, 2015), que discutiremos a seguir.

5.2 A origem natural do conteúdo

O artigo *The natural origins of content* (HUTTO; SATNE, 2015) que tenta desenvolver as bases para explicar conteúdo, segue a linha enativista, e começa com a seguinte pergunta:

“Assume some states of mind are contentful. Assume it is possible, for example, to think thoughts that refer to things beyond themselves, thoughts that can be true or false. Where, when and how did (and do) such contentful states of mind come on the scene? How is content possible in a natural world?” (“Assuma que alguns estados mentais tem conteúdo. Assuma ser possível, por exemplo, ter pensamentos que se referem a coisas para além deles mesmos, pensamentos que podem ser verdadeiros ou falsos. Onde, quando e como tais estados mentais com conteúdo apareceram (e aparecem)? Como é possível haver conteúdo em um mundo natural?”). O autorres prosseguem afirmando que para responder a essa pergunta, é necessário (i) não pressupor conteúdo, e (ii) ter credenciais científicas reconhecidas. A estrutura do artigo consiste em usar a divisão proposta por (HAUGELAND, 1990) para as linhas de pensamento a cerca do surgimento de conteúdo no mundo natural em três grupos: Neo-cartesianos, neo-behavioristas e neo-pragmatistas. São então tecidas críticas a essas linhas, buscando maneiras alternativas com que cada uma delas pode contribuir para a construção de uma teoria do conteúdo robusta o suficiente.

A primeira linha de pensamento seria a dos neo-cartesianos, que defendem que a intencionalidade é propriedade exclusiva de estados mentais com conteúdo. Para os autores, a mais promissora teoria neo-cartesiana atualmente é a teleossemântica, dos trabalhos de Millikan (MILLIKAN, 1984; MILLIKAN, 2005), que tenta mostrar como a seleção natural seria a responsável por transformar a teleologia intrínseca dos seres vivos em conteúdo representacional com condições de verdade. Para eles, a seleção natural não só cria uma normatividade biológica (por "projetar" ao longo da história da evolução disposições orgânicas que, se relacionando com itens do mundo, trabalham em função da sobrevivência do organismo e da perpetuação da espécie), mas no mesmo movimento cria os meios de representação internas os quais o organismo utiliza para perceber o mundo e atuar no mesmo. O artigo explica que na visão de Millikan, um aparelho tem a teleofunção de representar X se for usado, interpretado ou consumido pelo sistema por ter tal função de representar a presença de X.

Como falha da teleossemântica, é apresentada a crítica de (FODOR, 1990) de que essa explicação é incapaz de projetar uma descrição fixa (entre as muitas alternativas possíveis) como o conteúdo de uma representação. Para ilustrar, temos seu icônico exemplo do sapo: “Darwin doesn’t care how you describe the intentional object of frog snaps . . . Darwin cares how many flies you eat, but not what description you eat them under” (“Darwin não se importa com como você descreve o objeto intencional da linguada do sapo. . . Darwin se importa com a

quantidade de moscas comidas, não sob qual descrição elas foram comidas”). Este problema que (FODOR, 1990) apresenta, mais conhecido como Disjunction Problem (problema da disjunção) será exemplificado com mais detalhes na seção sobre ur-intentionality, em que analisaremos a proposta de (HUTTO; MYIN, 2017) para naturalização de intencionalidade sem conteúdo. É suficiente por agora entender que o argumento de Fodor, em outras palavras, é: explicar com base na história da evolução porque um organismo aponta para um objeto não é suficiente para mostrar como esse organismo representaria esse objeto precisamente, já que uma representação de uma disjunção qualquer de fatores "mosca ou X ou Y ou Z ou..." se encaixaria tão bem na explicação quanto a representação de "mosca".

O que Hutto e Satne argumentam é que não devemos descartar a teleossemântica por completo, mas sim reconceitualizá-la e transformá-la em teleossemiótica, o que significaria desconsiderar todas as suas tentativas de explicar conteúdo mas ainda assim utilizar a ideia de que a seleção natural cria um tipo de normatividade. Não há uma prova consistente de que casos mais simples de normatividade biológica originada da evolução por seleção natural não envolvam conteúdo, mas há uma constatação de que a (considerada) melhor tentativa de naturalizar conteúdo representacional (a teleossemântica) não é suficiente para os autores, pois não resolve algumas questões, entre elas o problema da disjunção de Fodor. Assume-se então que conteúdo não aparece necessariamente na normatividade biológica, mas que essa normatividade é a marca da intencionalidade.

O artigo segue por apresentar os neo-behavioristas, descritos como aqueles que negam conteúdo e afirmam que o que acontece é que seres como os humanos possuem a capacidade de olhar para certos sistemas ou processos e atribuir-lhes conteúdo. Hutto e Satne alegam que não fica bem explicado o que exatamente faz surgir a capacidade de atribuir conteúdo. A correção que os autores propõem seria a de diferenciar a ur-intentionality (intencionalidade sem conteúdo) da intencionalidade ativa (aquela em que existem agentes possuidores de práticas de atribuição de conteúdo).

A terceira linha, a neo-pragmatista, é apresentada como a que defende que os estados mentais com conteúdo derivam de práticas sócio-culturais já estabelecidas. Primeiro teriam surgido formas primitivas de linguagem para que então essas representações linguísticas fossem internalizadas. Segundo (HAUGELAND, 1990) a conformidade sócio-cultural também seria responsável pelo surgimento de um novo tipo de mecanismo aberto de seleção natural, que faria emergir novas práticas e instituições culturais, normas sociais e intencionalidade original.

Uma das críticas feitas por Hutto e Satne é que participar de práticas sócio-culturais requer intencionalidade, ou seja, requer que os seres apontem para certos elementos do mundo e para as práticas relacionadas àqueles elementos fazendo sentido desses elementos e práticas. Mas como a intencionalidade nos moldes tradicionais teria que envolver conteúdo, ela não poderia servir de suporte para a explicação do surgimento de conteúdo. A resolução para esse problema para os autores é novamente considerar que existem níveis de intencionalidade sem conteúdo (*ur-intentionality*), e esta intencionalidade seria a base para o surgimento de agentes cuja intencionalidade passa pela atividade de atribuir conteúdo, sendo essa atribuição originada de práticas sociais. O termo usado em inglês para descrever aspectos culturais como suportes na emergência de conteúdo é *cultural scaffolding*. Este é um termo de difícil tradução. Scaffold significa “andaime”, portanto *cultural scaffolding* guarda o sentido de que práticas culturais servem de andaime para outras práticas, ou seja, uma prática se constrói sobre a outra caracterizando um tipo de evolução aberta (*open-ended*).

Este artigo foi importante pois, além de esclarecer em que posição se encontra o enativismo radical quanto ao conceito de “conteúdo”, também monta o alicerce para o início de uma retórica positiva por parte do REC (que teve seu prosseguimento em (HUTTO; MYIN, 2017)), já que uma das críticas ao livro de 2012 de Hutto e Myin foi o excesso de ataques a outros posicionamentos em detrimento da construção de uma posição alternativa mais sólida.

Para resumir nosso entendimento sobre o artigo, temos que, para o REC: i - Há intencionalidade sem conteúdo; ii - A intencionalidade surge com a normatividade biológica originada da evolução por seleção natural, em outras palavras ela surge com as funções biológicas; iii - O conteúdo emergiria de práticas culturais muito específicas (*cultural scaffolding*), e portanto uma visão de intencionalidade antes desse contexto não possuiria conteúdo. Logo, o artigo aponta que a intencionalidade sem conteúdo está em toda atividade de organismos que pressuponha normatividade biológica mas não pressuponha normatividade cultural.

5.3 Enativismo Radical (REC) vs Enativismo Adaptativo Autopoiético (AAE)

A visão utilizada por (FROESE; ZIEMKE, 2009) em sua proposta de IA enativa, de que *sense-making* emerge nos organismos com autonomia constitutiva e capacidade de adaptação, é referida em (HUTTO; MYIN, 2017) como *Autopoietic-adaptive Enactivism (AAE)*. Para os autores, o AAE traz grandes contribuições ao identificar cognição no direcionamento que organismos apresentam para alguns itens do mundo, e ao tentar explicar tal cognição em termos

de normatividade biológica, excluindo do vocabulário toda a bagagem trazida por cognitivistas, como "mental contents, prior intentions, directions of fit"(conteúdos mentais, intenções anteriores, direções de ajuste) e similares.

O REC, apesar de não fazer uso da teoria da autopoiese, em nenhum momento critica a AAE quando a mesma coloca que a autonomia é responsável por criar um propósito natural inerente ao organismo (teleologia intrínseca). Hutto e Myin no entanto dão ênfase à colocação de que autonomia não é suficiente para explicar o comportamento dos organismos de apontarem para itens do mundo com objetivos (portanto, não explicaria cognição). A autopoiese nos diria por que sobreviver seria um propósito natural, mas não se encarregaria de explicar o comportamento de buscar condições favoráveis e evitar condições adversas que possam destruir o organismo gradualmente (pois não ser destruído imediatamente significaria cumprir o propósito de se manter vivo). Ambos REC e AAE concordam que uma característica essencial para a cognição é a adaptatividade. (HUTTO; MYIN, 2017) inclui então a seguinte citação: “if autopoiesis (or autonomy) suffices for generating a ‘natural purpose’ ..., adaptivity reflects the organism’s capability — necessary for sense making — of evaluating the needs and expanding the means towards that purpose” (FROESE; PAOLO, 2011) ("se basta autopoiese (ou autonomia) para gerar um 'propósito natural' . . . , adaptatividade reflete a capacidade do organismo — necessária para sense making — de avaliar as necessidades e expandir os meios para esse propósito"). Apesar de não criticarem diretamente a teoria da autopoiese e seu uso na AAE, Hutto e Myin, ao não utilizarem essa teoria como suporte para suas propostas e ao darem ênfase à adaptatividade como condição necessária para cognição, parecem assumir que a adaptatividade regulada pelas funções biológicas é não apenas necessária, mas suficiente para explicar cognição. E mais, a seleção natural seria o suficiente para explicar a origem dessas funções biológicas.

Os proponentes do REC também dão crédito ao AAE no que diz respeito a reconhecer que formas de cognição mais sofisticadas envolvendo conteúdo representacional só podem surgir por meio de práticas normativas muito especiais, como as práticas culturais humanas, e nada têm a ver com a normatividade biológica de seres mais simples. Qualquer tentativa de interpretar o pareamento entre itens do mundo e metabolismo de organismos menos complexos — como se as "peças" do corpo do organismo que reagissem a tais itens cumprissem a função de ser uma representação do mundo com um conteúdo significativo para o organismo — corre o risco de ser apenas uma antropomorfização desse organismo.

O ponto de desacordo entre REC e AAE está no termo "sense making". Primeiro por

que o termo parece assumir que há uma continuidade no reino biológico do fenômeno de construção de sentido, mas não explica como, partindo do "fazer sentido" anti-representacionista de seres mais simples, conseguimos chegar ao "fazer sentido" que envolva por completo representação e conteúdo presente nos seres humanos. Sense making, para o REC, seria adotar um vocabulário obscuro e não científico que parece projetar em organismos mais simples o sentimento humano de ver sentido no mundo, ao mesmo tempo que nega que seja um fenômeno idêntico. Parar neste termo, portanto, não seria o bastante para se construir uma base teórica sólida. (HUTTO; MYIN, 2017) cita a proposta de (JESUS, 2016) de prosseguir a partir de "sense making" do AAE, e dar suporte teórico a essa ideia por meio da biossemiótica.

Antes de dar prosseguimento, devemos fornecer uma breve noção do que seria a biossemiótica. A página da International Society for Biosemiotic Studies (BIOSEMIOTICS.ORG, 2018) cita algumas definições para a biossemiótica. Entre elas estão as seguintes: i - O Oxford Dictionary of Biochemistry and Molecular Biology (1997) apresenta a seguinte definição de biossemiótica: "The study of signs, of communication, and of information in living organisms" ("O estudo dos sinais, da comunicação, e da informação em organismos vivos"). ii - (SEBEOK, 1991) explica, em linhas gerais, que a biossemiótica enxerga os processos de troca de mensagens (semiose) como indispensáveis para os organismos vivos. Ele alega que, com exceção dos computadores, nenhum ser não-vivo tem a capacidade de conter, replicar e expressar mensagens, ou extrair seu significado. O estudo da semiose seria portanto um ramo indissociável do estudo dos seres vivos.

A proposta de (JESUS, 2016) é resumida por (HUTTO; MYIN, 2017) citando afirmações de que deve haver uma "reconceitualização não-antropocêntrica da noção de cognição como o processo ativo e criativo de bio-semiose por sistemas bio-semióticos" ("non-anthropocentric reconceptualisation of the notion of cognition as the active and creative process of bio-semiosis by bio-semiotic systems.") que consistiria em um fenômeno de interpretação de sinais onde há um "processo triádico irreduzível em que A interpreta B como 'correspondendo a' C" ("irreducible triadic processes in which A interprets B as 'standing for' C"). Em outras palavras, a cognição seria o processo de transformação de simples sinais oriundos do meio-ambiente em mensagens de fato, capazes de direcionar o estado interno daquele organismo e seu comportamento atual em função do item do mundo responsável por emitir tal sinal. Hutto e Myin seguem por analisar essa proposta, alegando que o modelo de sinais naturais "correspondendo a" itens do mundo não funciona pois é de fato completamente convidativo a análises representacionistas. A maneira

que interpretamos este posicionamento é de que os autores fogem dessa ideia de correspondência entre o sinal que chega ao ser vivo e o item do mundo por encararem a discussão como uma dicotomia "anti-representacionalismo" vs "representação + conteúdo", e portanto temem que falar de "sinais correspondendo a itens do mundo" traga toda a carga cognitivista (de que cognição necessariamente envolve conteúdo) a qual o REC tenta negar.

Ainda em relação a esta visão dicotômica de conteúdo, temos o segundo motivo para que haja um desacordo entre REC e AAE quanto a termo sense-making, e este motivo é que de fato sense-making se refere a conteúdo na visão da AAE, mas um tipo de conteúdo não-representacional. Quando (VARELA, 1992) discorre sobre a diferença entre o mundo de um organismo e ambiente em que o mesmo se encontra, ele afirma que esta diferença é o surplus of signification (ou "excedente de significado"), e que é exatamente este excedente que é a "mãe da intencionalidade". Este excedente de significado seria a normatividade semântica que emerge quando um organismo interage de diferentes com itens do mundo a depender de suas necessidades. Uma resenha do livro de Hutto e Myin de 2017, escrita por Evan Thompson (um dos grandes pensadores da assim chamada AAE), corrobora com esta visão. (THOMPSON,) discorda da visão de "conteúdo" dada por Hutto e Myin, e argumenta que nem todo conteúdo precisa ser interpretado em termos de condições de corretude, sendo a forma mais básica de conteúdo intencional uma que não está representada (codificada) em um "veículo" físico. (HUTTO; MYIN, 2017) discorre sobre formas de interpretar conteúdo de maneira não-representacional, mas rapidamente alega que o dito conteúdo não-representacional seria apenas o caso em que o sentido e a referência não se distinguem, ou seja, o objeto intencional seria objetivamente o item do mundo em si com o qual o ser cognitivo estivesse interagindo. Thompson também discorda deste ponto em sua resenha, argumentando que descrever um conteúdo intencional não-representacional é mais do que descrever puramente o objeto físico (item do mundo) em questão, mas entender como tal objeto se apresenta ao ser cognitivo, e porque o ser interage com o objeto da maneira particular que interage. Também argumenta que a diferença de sentido-referência de Frege trata de linguística, e que formas básicas de cognição não-linguísticas ainda necessitam de uma investigação para averiguar se há modos de apresentação não-representacionais.

Claramente, assumir a posição enativista, questionar as bases do que se entende por intencionalidade, mas não questionar o que majoritariamente se entende por conteúdo (caindo na já citada dicotomia "anti-representacionalismo" vs "representação + conteúdo") pode causar alguns problemas. Certamente, o próximo caminho a seguir seria o de dissociar conteúdo e

representação. O artigo de Carlos E. F. de Brito, *Intelligence with Representation but No Content* (2017), discorda dessa dicotomia, mas propõe uma discussão um pouco diferente. Se por um lado, a contenda entre REC e AAE está no conteúdo representacional versus conteúdo não-representacional, (BRITO, 2016) defende a possibilidade de haver representação sem conteúdo. Mais adiante analisaremos este artigo, mas antes detalharemos a proposta de Hutto e Myin para a naturalização de intencionalidade sem conteúdo.

(HUTTO; MYIN, 2017) prossegue com os trabalhos iniciados em (HUTTO; MYIN, 2012) e posteriormente em (HUTTO; SATNE, 2015), e propõe a construção de um projeto para a naturalização de intencionalidade sem conteúdo. Como discutido, os autores advogam por uma visão dupla da cognição em que existe um salto de capacidades entre mentes sem conteúdo e mentes com conteúdo, a esse salto eles dão o nome de *Kinky Cognition*. A seguir discutiremos a naturalização proposta em (HUTTO; MYIN, 2017) para a *ur-intentionality*.

6 “INTENCIONALIDADE SEM CONTEÚDO” - ENTENDENDO A UR-INTENTIONALITY

Uma proposta de naturalização para intencionalidade sem conteúdo foi apresentada por Daniel Hutto e Erik Myin no livro *Evolving Enactivism: Basic Minds Meet Content*, de 2017. Chamada de ur-intentionality (ur-intencionalidade), a ideia é seguir com o que foi proposto em (HUTTO; SATNE, 2015) e utilizar a Teleossemântica de Millikan para explicar, através da seleção natural, por que os seres vivos apresentam uma direcionalidade a certos itens do ambiente.

Os autores iniciam esta proposta de naturalização buscando definições clássicas de intencionalidade. Entre tais definições, algumas das quais já abordamos na seção que explica o termo e suas raízes filosóficas, eles relatam que Brentano disseminou a ideia de que a intencionalidade é a única marca do mental e não é redutível em termos de fenômenos naturais, o que posteriormente foi amenizado para "a intencionalidade é uma das marcas do mental, e também é uma marca naturalizável"(ROY, 2015) (“intentionality is one of the marks of the mental, as well as a naturalizable mark”). Este relaxamento do termo permitindo sua naturalização é bem vindo, mas um problema ainda reside quando falamos sobre o paradigma Brentaniano. Em um contexto em que a filosofia analítica trata intencionalidade como envolvendo conteúdo por definição, defender que existem níveis básicos de intencionalidade onde não há conteúdo soa contraditório. Um dos objetivos seria então desacoplar as duas ideias de modo a pensar em uma forma de intencionalidade primitiva, sem conteúdo, não representacional, mas ainda assim tendo a principal propriedade do fenômeno intencional, o de ser uma "atitude direcionada a um objeto"(HUTTO; MYIN, 2017).

O primeiro passo para a naturalização é entender por onde começar a busca. Citando (MULLER, 2013), os autores mostram que a perspectiva naturalista para explicar um fenômeno consiste usualmente em enxergá-lo como algo construído por estruturas e capacidades mais simples, e do ponto de vista biológico, essas estruturas mais simples podem ser entendidas observando organismos mais simples e buscando respostas no processo amplamente aceito de evolução Darwiniana por seleção natural. Ganhamos mais por aplicar nossas intuições nesses exemplos menos complexos em oposição a pensar no fenômeno como tendo surgido totalmente formado e sendo irredutível. Em resumo, os autores apelam para o modo de pensar científico/naturalista para justificar o porquê de o estudo ser bottom-up ao invés de top-down. Para eles, a melhor tentativa bottom-up de explicar conteúdo utilizando a evolução Darwiniana foram

as teorias teleossemânticas de Millikan, e eles seguem por tentar desacoplar intencionalidade e conteúdo dentro da teleossemântica argumentando que essa teoria não é capaz de explicar conteúdo, mas que é possível extrair dela a explicação para a direcionalidade de organismos para com certos objetos do mundo. A essa reconceitualização eles dão o nome de teleossemiótica.

(HUTTO; MYIN, 2017) explica que a teleossemântica encara "representações mentais como estados internos que possuem a função biológica de habilitar organismos a acompanharem itens específicos do mundo" ("... mental representations as inner states that have the biological function of enabling organisms to keep track of specific worldly items"). Mas para esta visão, não bastaria que um item tenha a função biológica de relacionar sistematicamente um organismo a aspectos do ambiente. Para haver de fato representação mental, as seguintes condições deveriam ser atendidas: "i - O item ou resposta deve ser relacionalmente adaptado a alguma característica específica do mundo. ii - A relação em questão deve ser caracterizada por uma regra de mapeamento ou correspondência. iii - O item ou a resposta deve ter a função apropriada de orientar um dispositivo consumidor cooperativo no desempenho da função adequada daquele dispositivo consumidor com relação ao rastreamento ou atuação na característica de mundo visado em (i)".

Novamente vemos uma lógica tríade, semelhante à proposta de (JESUS, 2016) citada por (HUTTO; MYIN, 2017) e abordada na seção sobre o marco teórico do enativismo radical. O dispositivo consumidor (A) utiliza o item ou resposta do organismo (B) que representa uma característica do mundo (C) para realizar uma função apropriada. Relembremos que a proposta em (JESUS, 2016) consistia em interpretar o sense-making do AAE à luz da biossemiótica como um "processo triádico irreduzível em que A interpreta B como 'correspondendo a' C". Agora que detalhamos a proposta da teleossemântica, vemos por que o REC foge de um vocabulário que afirme que um item em um organismo "corresponda a" uma característica do mundo, e por que Hutto e Myin em sua obra de 2017 coloca essa interpretação biossemiótica como tão convidativa às teorias teleossemânticas.

É importante notar no entanto que os fatores que compõem ambas as tríades, em especial "B", têm significados diferentes para cada teoria. No caso da teleossemântica, B é de fato encarado como um item ou resposta do próprio organismo representando um conteúdo referente a uma característica externa do mundo. Para a biossemiótica, podemos interpretar B como qualquer sinal, interno ou externo, no fluxo de informação. Mesmo que esta visão biossemiótica evite usar "conteúdo" para explicar a relação cognitiva dos organismos com o mundo, Hutto e Myin

parecem tomar um cuidado extra de evitar que essa visão contribua para a posição cognitivista ao usar um vocabulário que remete ao representacionalismo. No entanto, se descrevermos o que acontece com os seres vivos em termos puramente de informação-como-covariância, corremos o risco de não diferenciar suficientemente as atividades dos organismos de outros eventos naturais do mundo inorgânico. Se enxergamos, portanto, sob o ponto de vista do REC, a biossemiótica pode ficar entre estar “perigosamente” próxima do cognitivismo ou ser insuficiente para explicar cognição.

Mas Independentemente de o vocabulário usado tratar de representação, conteúdo ou fluxo informacional, o que é observado é que organismos são dotados de aparatos responsáveis por criar uma correspondência entre suas ações e características específicas do mundo, e estas ações ocorrem como resultado de uma série de funções biológicas adequadas performando de maneira correspondente ao mundo. Isto é ponto pacífico para as três abordagens, AAE, REC e Teleosemântica. O AAE fala da importância da adaptatividade, sendo as funções biológicas as responsáveis por conferir tal adaptatividade, e tendo elas sido naturalmente “projetada” por seleção natural. REC e teleosemântica fundamentam suas visões nas funções biológicas, e como elas são responsáveis por fazer surgir um tipo de normatividade. Entender que a performance de uma função biológica é correspondente a um item do mundo é entender que esta função só é presente e só atua da maneira que atua em razão de ter servido historicamente em algum ponto para a sobrevivência e perpetuação de um ancestral. Por exemplo, a função desempenhada pelos flagelos de determinada bactéria de fazê-la nadar para longe de um determinado ambiente aquoso pode corresponder ao fato de o ambiente estar muito quente, pois historicamente a obtenção dessa função após certas mutações foi útil para a sobrevivência daquele organismo e para a perpetuação dessa função que permita aos seus descendentes a sobrevivência a esse tipo de situação. É isto que significa ter uma função adequada (proper function).

E é por isso que (HUTTO; MYIN, 2017) buscam na “RECTificação” da teleosemântica a naturalização que precisam para intencionalidade sem conteúdo. Eles colocam: "Their concern is always with what consumer devices are supposed to do, not what they are disposed to do. Recognizing this is utterly crucial when it comes to understanding teleosemantic theories of content" ("Sua preocupação é sempre com o que dispositivos consumidores devem fazer, não o que estão organizados para fazer. Reconhecer isso é absolutamente crucial para entender teorias teleosemânticas de conteúdo"). O que a teleosemântica basicamente afirma é que como há uma correspondência entre itens/respostas orgânicas internas e características do mundo, e como

esses itens internos são usados por um dispositivo consumidor para desempenhar uma função adequada (ou seja, existe a possibilidade do desempenho da função falhar), então tais itens internos são representações com conteúdo sobre mundo, e se representados incorretamente acarretarão em um mal-funcionamento. Logo, há uma condição de verdade sobre se a representação do mundo está correta ou incorreta. Hutto e Myin capturam da teleossemântica a compreensão de que a evolução e as funções biológicas criam uma espécie de normatividade que é mais do que puramente fluxo de informação-covariância, mas ao mesmo tempo negam que tal normatividade seja semântica, ou que haja uma representação com conteúdo e condições de verdade. Para eles, o que acontece é um fluxo de informação regulado por uma normatividade mais simples que a semântica, uma normatividade biológica. Por isso o nome “teleosemiótica”.

Como já vimos na seção sobre conteúdo, uma crítica à classificação de um item ou resposta orgânica como conteúdo representacional é o Disjunction Problem (FODOR, 1990). Retornando ao exemplo do sapo esticando a língua para capturar uma mosca, o argumento segue na linha de que não podemos afirmar que o que está sendo representado pelo aparelho sensorial do sapo é "mosca" e que esta seja uma condição de verdade (no sentido de que se algo que não for uma mosca for percebido como mosca pelo sapo, esta seria uma representação falsa). A representação poderia na verdade ser da proposição "mosca ou X ou Y ou Z ou...", onde X, Y, Z etc, seriam objetos do mundo "confundíveis" com mosca pelo sapo. Neste caso a teleossemântica falharia em explicar como se dá a representação de maneira precisa, pois qualquer disjunção de fatores faria o trabalho de ser uma representação apropriada para a função de pegar moscas. Segue que a teoria não dar conta de naturalizar definições precisas de representação a tornaria inútil para o propósito a que se coloca. Estendendo a explicação, interpretamos que o argumento é de que a rede de processos de um sapo, selecionada por meio de seleção natural, que aponta para pontos pretos no ambiente esticando a língua para alimentar-se desses pontos, não tem uma condição de verdade precisa, pois a seleção natural não trabalharia dessa maneira. Não existiria uma representação interna de moscas tal que esticar a língua para uma pintura de um ponto preto seria uma apreensão falsa e pegar de fato uma mosca seria uma verdadeira. Esticar a língua para moscas e pontos pretos são apenas consequências de um sistema que no esquema geral funciona imerso no ambiente em que moscas e pontos pretos existem, pois foi selecionado naturalmente por funcionar nesse ambiente. É como uma máquina projetada para cortar grama a uma certa altura, e que cortará tudo em tal altura, até o que não for grama. Ou um prédio cujas colunas têm capacidade para aguentar mais peso do que é preciso. No final das contas, independente

de fazer ou não mais do que o necessário, a grama vai ser cortada, o prédio se manterá erguido, e o sapo comerá moscas. Os sistemas funcionam para o que foram projetados a fazer (com a diferença óbvia que os objetivos de um ser vivo são intrínsecos a ele e foram "projetados" pela seleção natural). Apenas a seleção natural não é suficiente para explicar conteúdo.

O que Hutto e Myin argumentam é que, mesmo que não haja conteúdo em um nível mais básico, funções biológicas são suficientes para explicar a origem natural da intencionalidade, pois a seleção natural cria uma normatividade, de maneira que as funções biológicas apontam para itens do mundo de acordo com sua capacidade de fornecer/perpetuar as condições de realização de seus objetivos intrínsecos de sobrevivência. Em outras palavras, um sapo pode não representar o conteúdo referente a "mosca" internamente, mas certamente seu comportamento veio a se desenvolver daquela maneira pois alimentar-se de moscas contribuiu para a espécie evolutivamente. Para os autores, o problema da disjunção afetaria somente a nível de representação de conteúdo pois, por mais que qualquer disjunção de fatores que inclua o item de mundo correspondente à função adequada a ser desempenhada seja uma representação apropriada (impossibilitando à teoria de fornecer uma explicação de como representações precisas aparecem), não é qualquer desses fatores ou disjunções que importam para a manutenção de tal função ao longo da história evolutiva. Um sapo só possui tal ato de esticar a língua para tudo que se pareça com mosca porque a mosca especificamente serviu em algum ponto como alimento para um ancestral e se perpetuou como alimento ao longo do processo evolutivo. (HUTTO; MYIN, 2017) coloca: "the frog's perception-action routines were forged through a long process of selection by consequences in order to get flies, and not anything else, into its belly ... Hence, the biological function of the frog's tongue-snapping behavior is not disjunctive at all" ("As rotinas de percepção-ação do sapo foram forjadas por um longo processo de seleção por consequências a fim de obter moscas, e não qualquer outra coisa, para sua barriga ... Portanto, a função biológica do comportamento de esticar a língua do sapo não é nada disjuntivo"). O apontamento é que por mais que a evolução não explique a criação de uma normatividade semântica, ela teria poder para explicar o surgimento de uma normatividade mais simples, biológica. É essa normatividade que permitiria dizer que os organismos, por meio de funções biológicas adequadas, apontam para itens do mundo como objetos que servem a propósitos intrínsecos. (HUTTO; MYIN, 2017) vai além em sua crítica à teleosemântica e afirma que explicar via seleção natural por que os organismos apontam para itens do mundo da maneira que apontam não é suficiente para explicar como se dá especificamente a representação de conteúdo. Utilizar a explicação

do porquê para chamar atenção para as correspondências entre "características do mundo" e "consumo de itens/respostas orgânicas para desempenho de funções", e a partir disso saltar para atribuir a tais itens/respostas orgânicas status de conteúdo representacional coloca a causalidade puramente nos processos biológicos, tal qual qualquer teoria eliminativista faria. Isso quer dizer que explicar conteúdo nesses termos tira qualquer poder causal que o mental possa ter, mesmo em seres humanos.

Uma observação importante é que os autores fazem questão de esclarecer que esta proposta de naturalização não se resumiria a respostas a estímulos cegas e fixadas "de fábrica" (hardwired), pois uma das características da adaptatividade dos seres vivos explicada pela evolução é a capacidade de aprendizado, ou seja, a capacidade de, mediante o contato com estímulos prévios, se reorganizar internamente para responder melhor a tais estímulos. Portanto a maneira que um organismo aponta para um objeto do mundo permite a modificação por meio do aprendizado. Outra observação dos autores é a de que falar de adaptatividade não implica que o organismo possua a representação interna prévia do estado favorável a que pretendem chegar quando estão em um estado, local ou configuração desfavorável. Não existe uma noção prévia de um estado que se quer alcançar, mas sim uma constante interação organismo-ambiente que empurra o organismo contra desvios que fogem da norma de viabilidade (normatividade essa que, como já vimos, explica-se pela seleção natural).

6.1 Reação vs Intencionalidade

A concepção que se tem de intencionalidade é a de um direcionamento ativo do ser intencional para com o objeto intencional. Portanto, podemos dizer que a sentido da direcionalidade de intencionalidade é "de fora para dentro". Um leitor mais observador poderia se perguntar, baseando-se no conceito de telefunção, se de fato qualquer função apropriada de um organismo que o direciona para uma maior estabilidade – que lhe permita mais chances de sobrevivência – possa ser encarado como sendo intencional. Dessa maneira, estaríamos igualando reação à intencionalidade.

Observando novamente a definição para representação mental da teleossemântica abordada em (HUTTO; MYIN, 2017), temos o item 3 que diz: "O item ou a resposta deve ter a função apropriada de orientar um dispositivo consumidor cooperativo no desempenho da função adequada daquele dispositivo consumidor com relação ao rastreamento ou atuação na característica de mundo visado em (i)". E ao lado deste item, há um número que nos direciona

para a seguinte nota dos autores na parte final do livro: “For this reason, the mechanisms that distribute pigmentation so as to match the skin of chameleons to their immediate environments are not candidates for being intentionally directed. This is because such devices fail to satisfy the final condition.” (Por esta razão, o mecanismo que distribui pigmentação de modo a igualar a pele dos camaleões a suas imediações não são candidatos para serem intencionalmente direcionados. Isto por que tais aparelhos falham em satisfazer a condição final”. Podemos ver por esta nota que a ur-intentionality não abarca todo e qualquer tipo de reação orgânica, e não é toda teleofunção que atua como fator principal em uma ação intencional. Não é qualquer reação (como curar-se de um ferimento ou resistir a um choque físico) que se pode conferir caráter intencional, e Hutto e Myin concordam com isso. A pergunta que fica é, se a ur-intentionality quer negar conteúdo para mentes básicas, que sentido faz manter em sua teoria as condições da teleossemântica para conteúdo representacional?

Transformar teleossemântica em teleossemiótica implica em dizer que a teleossemântica falha frente ao Disjunction Problem e ao HPC, mas não implica em se desfazer das condições que a compõem. A obra de Hutto e Myin não trata desta questão de forma clara, mas podemos perceber pela supracitada nota dos autores que RECTificar teleossemântica também passa por tomar as três condições para existência de representação com conteúdo e afirmar que as exatas mesmas condições na verdade regulam a existência de intencionalidade. Seguindo a lógica do que foi proposto pelo HPC, os itens ou respostas dos organismos não seriam representações de conteúdos, mas sim puramente fluxo informacional. Alguns desses fluxos informacionais seriam responsáveis por desencadear a realização de uma função apropriada de rastreamento ou atuação sobre um item do mundo, e por motivos de seleção natural, esta relação entre o fluxo informacional específico e o desempenho da função permaneceram e se proliferaram, resultando no surgimento de uma normatividade biológica.

7 PROXY CAUSE - UMA CONTRIBUIÇÃO AO ENATIVISMO RADICAL

Como já vimos, uma das críticas ao REC é de não dissociar representação e conteúdo. Um dos trabalhos recentes que busca iniciar esse trabalho é o artigo de 2016, *Intelligence with Representation but No Content*, de Carlos Brito, onde ele relata como a discussão em torno dos fenômenos cognitivos giram em torno de uma dicotomia. De um lado, os anti-representacionistas que tomam a cognição como uma relação de causalidade constante entre os estímulos do ambiente e o aparato sensório-motor do organismo, e neste jogo de pareamento dinâmico não há espaço para representação. Do outro lado, os cognitivistas, que enxergam a cognição por definição como envolvendo representações internas com conteúdo sobre o mundo, e alegam que existem fenômenos cognitivos em que ocorrem processamentos de fatores que não estão presentes naquele exato momento, e que portanto pressupõem que tais representações devem existir internamente para serem processadas. Enquanto os anti-representacionistas encaram como mera antropomorfização tomar organismos tão simples quanto bactérias como sendo capazes de entender e representar o que está no mundo, os cognitivistas negam que seja possível conceber a cognição como algo diferente de processamento de símbolos internos que correspondem a itens do mundo. Esta dicotomia é ponto central em toda a discussão do enativismo radical. Aterrando sua concepção de informação em um conceito de constraint tomado da física, (BRITO, 2016), com o objetivo de explicar como alguns organismos parecem tomar decisões baseando-se no "ausente" ou "abstrato" (CLARK; TORIBIO, 1994) sem sair da ideia de regime causal, usa tal concepção para cunhar o termo proxy cause, que ele advoga por ser interpretado como uma representação sem conteúdo, transformando a dicotomia em uma tricotomia.

(BRITO, 2016) cita a definição de constraint de (MORENO; MOSSIO, 2015): “constraints are local and contingent causes, exerted by specific structures or processes, which reduce the degrees of freedom of the system on which they act, ... constraints are usually introduced as external determinations that contribute to determining the behavior and dynamics of a system” (“Constraints são causas locais e contingentes, exercidas por estruturas ou processos específicos, que reduzem os graus de liberdade do sistema no qual elas atuam, ... constraints são geralmente introduzidas como determinações externas que contribuem para determinar o comportamento e a dinâmica de um sistema”). Em outras palavras, quando aquecemos um balão de borracha com ar dentro, por exemplo, as partículas gasosas recebem energia e se agitam, e o processo dinâmico de várias moléculas agitadas colidindo entre si e contra as paredes de

borracha do balão impõem uma limitação ao formato do balão, que não pode secar mais do que um determinado limite permitido pelo ar aquecido em seu interior, então o mesmo infla. De maneira semelhante, as moléculas que constituem o material da borracha do balão impõe limitações de movimentação e organização entre si, permanecendo ligadas umas nas outras em um estado de maior organização do que o gás em seu interior, admitindo um estado sólido em um formato específico. Este formato também impõe uma limitação na expansão do gás sendo aquecido em seu interior, impedindo que a tal temperatura ele se expanda mais do que um certo limite, e impedindo que o ar de dentro vaze com velocidade pelos poros da borracha. Todas essas limitações impostas que diminuem graus de liberdade são o que tomamos por constraints. O artigo de Carlos Brito segue por citar os trabalhos de (COLLIER, 2011; PATTEE, 1977) como relacionando informação a constraint. (BRITO, 2016) aponta que uma proxy cause é uma constraint física.

(ILLARI; RUSSO, 2016) apresenta de maneira concisa o modelo de informação por covariância, e credita este modelo a estudos feitos na área de filosofia da informação desde os anos 90 por pesquisadores como Collier, Dretske, Floridi, Barwise e Seligman. Segue a definição: "The covariance model of an infomorphism states that if two systems a and b are coupled in such a way that a's being (of type, or in state) F is correlated to b being (of type, or in state) G, then such a correlation carries for the observer of a the information that b is G." ("O modelo de covariância de um infomorfismo afirma que se dois sistemas a e b são pareados de tal maneira que 'a' ser/estar (de um tipo / em um estado) F está correlacionado a 'b' ser/estar (de um tipo / em um estado) G, então tal correlação carrega para o observador de 'a' a informação de que 'b' é/está G."), onde, segundo (COLLIER, 2011), um canal de informação é uma família de infomorfismos. Uma possível interpretação para esse modelo é a de que, para existir tal correlação entre 'a' e 'b', deve haver uma limitação física (constraint) imposta pelo estado F de 'a', tal que F diminua os graus de liberdade de 'b', colocando-o em um estado G. Portanto, informação como constraint não contradiz a teoria da informação assumida pelo REC de informação-covariância. Ao contrário, ela complementa o conceito de informação.

(BRITO, 2016) introduz a noção de proxy cause como: "something that stands in for something else as a cause of a specific effect on a third element, typically a functional or adaptive mechanism. We propose to interpret a proxy cause as a contentless representation." ("algo que substitui outro algo como a causa de um efeito específico em um terceiro elemento, tipicamente um mecanismo funcional ou adaptativo. Propomos interpretar uma proxy cause como uma

representação sem conteúdo."). O artigo então faz o esforço de apontar para as ações de organismos cuja causalidade e a troca energética é imediata e diferenciá-las da proxy cause, que seria um elemento que desacopla o regime causal de limitações de tempo e espaço, pois a energia para reagir a tal estímulo vem de dentro do próprio organismo que reage. As proxy causes abririam a possibilidade de reações muito mais complexas do que apenas respostas cegas imediatas sensório-motoras, pois seria possível a um organismo combinar proxy causes oriundas de diferentes fontes para orientar a sua rede interna de sistemas a desempenhar uma função mais robusta (obviamente, a capacidade de uma rede de sistemas interna combinar diferentes proxy causes dessa maneira adaptativa foi "projetada" por seleção natural).

Este trabalho possui dois aspectos extremamente importantes. O primeiro é o de explicar adaptatividade, aprendizado e 'abstração' por meio de um regime causal aterrado na física, descartando assim a necessidade de explicar os fenômenos com base em representação com conteúdo na cognição mais simples (que não envolve o típico conteúdo mental humano, envolto em uma carga cultural e linguística). O segundo aspecto é o defender o conceito de que nem sempre a informação carregará um conteúdo representacional, pois ela pode servir a uma função responsável pelo know-how do organismo, e não pelo know-that.

Certamente, um defensor do REC dificilmente aceitaria que as proxy causes fossem definidas como representações. No entanto, como (BRITO, 2016) já parte de um ponto de vista de que as proxy causes não envolvem conteúdo, uma possível RECtificação dessa ideia seria puramente terminológica, de classificá-las como nada mais do que regimes muito especiais de fluxo informacional. Explorar essa ideia pode ser um grande trunfo para o REC, pois, apoiando-se apenas na física e sem recorrer a conteúdo, o conceito tem o poder de refinar a explicação de fenômenos intencionais que se direcionam para itens do mundo não de maneira imediata e explicitamente reativa, mas de maneira desacoplada, sem graves limitações de tempo e espaço.

8 CONCLUSÃO

Concluimos que o enativismo radical é um passo importante para questionar as bases do que é proposto hoje para a cognição. No entanto, o termo “intencionalidade sem conteúdo” – que deveria ser mudado para “intencionalidade sem conteúdo representacional”, ainda suscitará muitos debates, principalmente porque o projeto enativista mais aceito, chamado de AAE pelos proponentes do REC, advoga por um tipo de conteúdo não-representacional (algo que é descartado pelo REC). Dialogar a cerca dessas diferentes noções de conteúdo certamente é um trabalho futuro a ser feito com mais cautela. No que tange outros aspectos do pensamento enativista, REC e AAE não apresentam graves discordâncias. Em suma, podemos dizer que a contribuição da AAE foi perceber que, para a agência intencional existir, deve haver autonomia e adaptatividade. A teleossemântica contribuiu mostrando que apenas faz sentido falar de adaptatividade quando temos uma explicação natural (seleção natural) para o surgimento de uma normatividade por meio de funções adequadas que regulam tal adaptatividade. O REC por sua vez, apoiando-se nessas contribuições, afirma que tal normatividade não precisa necessariamente envolver representações conteúdo, e sua teoria de *kinky cognition* aborda o conteúdo como um salto de capacidades que envolve práticas culturais. Finalizamos o trabalho com uma possível contribuição ao REC. O artigo (BRITO, 2016), que trata de um tipo particular de fluxo informacional desacoplado, se propriamente “RECTificado”, fornece munição explanatória para aterrar na física uma série de fenômenos orgânicos classificados por cognitivistas como envolvendo representações com conteúdo.

REFERÊNCIAS

- BIOSEMIOTICS.ORG. **Biosemiotics introduction**. 2018. Disponível em: biosemiotics.org/biosemioticsintroduction. Acesso em: 10 jul. 2018.
- BRENTANO, F. **Psychology from an Empirical Standpoint**. [S.l.: s.n.], 1874. ISBN 978-1138019171.
- BRITO, C. E. F. d. **Intelligence with Representation but No Content**. 2016. Não publicado.
- BROOKS, R. Intelligence without representation. **Artificial Intelligence**, [S.l.], v. 47, p. 139–159, 1991.
- CLARK, A.; CHALMERS, D. J. The extended mind. **Analysis**, [S.l.], Oxford University Press, v. 58, n. 1, p. 7–19, 1998.
- CLARK, A.; TORIBIO, J. Doing without representing? **Synthese**, [S.l.], Springer, Springer Nature, v. 101, n. 3, p. 401–431, 12 1994. ISSN 0039-7857.
- COLLIER, J. Information, causation and computation. In: **Information and Computation**. [S.l.: s.n.], 2011. p. 89–105.
- DEACON, T. **Incomplete Nature: How Mind Emerged from Matter**. [S.l.]: W. W. Norton, 2012. ISBN 9780393049916.
- DENNETT, D. C. **Brainstorms**. [S.l.]: MIT Press, 1978.
- DRETSKE, F. **Knowledge and the Flow of Information**. [S.l.]: MIT Press, 1981. ISBN 9780262040631.
- DREYFUS, H. L. Why heideggerian ai failed and how fixing it would require making it more heideggerian. **Philosophical Psychology**, [S.l.], v. 20, n. 2, p. 247–268, 2007.
- FODOR, J. A. **A Theory of Content and Other Essays**. [S.l.]: MIT Press, 1990. ISBN 9780262061308.
- FROESE, T.; PAOLO, E. A. D. The enactive approach: Theoretical sketches from cell to society. **Pragmatics and Cognition**, [S.l.], v. 19, n. 1, p. 1–36, 2011.
- FROESE, T.; ZIEMKE, T. Enactive artificial intelligence: Investigating the systemic organization of life and mind. **Artificial Intelligence**, [S.l.], v. 173, n. 3–4, p. 466–500, 2009.
- HAUGELAND, J. The intentionality all-stars. **Philosophical Perspectives**, [S.l.], Wiley-Blackwell, v. 4, p. 383–427, 1990.
- HENDERSON, H. **Artificial Intelligence: Mirrors for the Mind**. [S.l.]: Chelsea House, 2007. (Milestones in discovery and invention). ISBN 9780816057498.
- HUTTO, D.; SATNE, G. The natural origins of content. **Philosophia**, [S.l.], Springer, v. 43, n. 3, p. 521–536, 2015. ISSN 0048-3893.
- HUTTO, D. D.; MYIN, E. **Radicalizing Enactivism: Basic Minds Without Content**. [S.l.]: MIT Press, 2012.

- HUTTO, D. D.; MYIN, E. **Evolving Enactivism: Basic Minds Meet Content**. [S.l.]: MIT Press, 2017.
- ILLARI, P.; RUSSO, F. Information and causality. In: FLORIDI, L. (Ed.). **The Routledge Handbook of Philosophy of Information**. [S.l.]: Routledge, 2016.
- JACOB, P. Intentionality. In: ZALTA, E. N. (Ed.). **The Stanford encyclopedia of philosophy**. Winter 2014 ed. [s.n.], 2014. Disponível em: <<https://plato.stanford.edu/entries/intentionality/>>.
- JESUS, P. D. From enactive phenomenology to biosemiotic enactivism. **Adaptive Behavior**, [S.l.], v. 24, n. 2, p. 130–146, 2016.
- JONAS, H. **The Phenomenon of Life: Toward a Philosophical Biology**. [S.l.]: Northwestern University Press, 2001.
- LAU, J.; DEUTSCH, M. Externalism about mental content. In: ZALTA, E. N. (Ed.). **The Stanford encyclopedia of philosophy**. Winter 2016 ed. [s.n.], 2016. Disponível em: <<https://plato.stanford.edu/archives/win2016/entries/contentexternalism/>>.
- MILLIKAN, R. **Language, Thought, and Other Biological Categories: New Foundations for Realism**. [S.l.]: MIT Press, 1984.
- MILLIKAN, R. **Language: A Biological Model**. Oxford: Oxford University Press, 2005.
- MORENO, A.; MOSSIO, M. **Biological Autonomy: A Philosophical and Theoretical Enquiry**. [S.l.]: Springer Netherlands, 2015. ISBN 978-94-017-9837-2.
- MULLER, H. D. 10 naturalism and intentionality. In: MULLER, B. B. H. (Ed.). **Contemporary Philosophical Naturalism and its Implications**. [S.l.]: Routledge, 2013. p. 13–155.
- NEWELL, A.; SIMON, H. A. Computer science as empirical inquiry: symbols and search. **Communications of the ACM**, [S.l.], v. 19, n. 3, p. 113–126, 1976.
- PAOLO, E. A. D. Autopoiesis, adaptivity, teleology, agency. **Phenomenology and the Cognitive Sciences**, [S.l.], Springer, v. 4, n. 4, p. 429–452, 2005.
- PATTEE, H. Dynamic and linguistic modes of complex systems. **International Journal Of General System**, [S.l.], v. 3, p. 259–266, 01 1977.
- PFEIFER, R.; IIDA, F.; BONGARD, J. New robotics: Design principles for intelligent systems. **Artificial Life**, [S.l.], v. 11, n. 1-2, p. 99–120, 2005.
- PITT, D. The paraphenomenal hypothesis. **Analysis**, [S.l.], v. 77, n. 4, p. 735–741, 08 2017. ISSN 0003-2638. Disponível em: <https://doi.org/10.1093/analys/anx123>.
- ROWLANDS, M. **The New Science of the Mind: From Extended Mind to Embodied Phenomenology**. [S.l.]: Bradford, 2010.
- ROY, J.-M. Anti-cartesianism and anti-brentanism: The problem of anti-representationalist intentionalism. **The Southern Journal of Philosophy**, [S.l.], v. 53, p. 90–125, 09 2015.
- SEBEOK, T. **A Sign is Just a Sign**. Bloomington: Indiana University Press, 1991.

THOMPSON, E. [sem título]. **Notre Dame Philosophical Reviews**, [S.l.], 11 jan. 2018. Resenha da obra de: HUTTO, D. D.; MYIN, E. *Evolving Enactivism: Basic Minds Meet Content*. [S.l.]: MIT Press, 2017. Disponível em: <https://ndpr.nd.edu/news/evolving-enactivism-basic-minds-meet-content/>. Acesso em: 10 mar. 2018.

THOMPSON, E. Life and mind: From autopoiesis to neurophenomenology. a tribute to francisco varela. **Phenomenology and the Cognitive Sciences**, [S.l.], v. 3, 12 2004.

VARELA, F. **Principles of Biological Autonomy**. New York, NY: Elsevier North Holland, 1979.

VARELA, F. J. Autopoiesis and a biology of intentionality. **B. McMullin, N. Murphy (Eds.), Proc. of Autopoiesis and Perception: A Workshop with ESPRIT BRA 3352, Dublin City University, Dublin, Ireland**, [S.l.], Oxford University Press, p. 4–14, 1992.

VARELA, F. J.; THOMPSON, E.; ROSCH, E. **The Embodied Mind: Cognitive Science and Human Experience**. [S.l.]: MIT Press, 1991.