

## DE UMA SEMÂNTICA CONCEPTUAL PARA UMA NEUROSEMÂNTICA:

APRESENTAÇÃO E DISCUSSÃO CRÍTICA DA CONTENDA FODOR & LEPORE *vs* PAUL  
CHURCHLAND A PROPÓSITO DA NOÇÃO DE SIMILARIDADE DE SIGNIFICADO<sup>1</sup>

### I - introdução

O presente artigo assume como propósito fundamental apresentar a polémica/contenda mantida entre Fodor & LePore (F&L) e Paul Churchland (PC) a respeito da semântica dos estados mentais. Este tópico é, só por si, importante para a definição do que seja uma proposta de explicação da racionalidade humana em termos de atitudes proposicionais. Atitudes como crenças, desejos, dúvidas, etc. estão no centro da maioria das teorias acerca da racionalidade humana. Podemos esquematicamente esclarecer como a questão acerca da semântica dos estados mentais se relaciona com este propósito mais amplo acerca da racionalidade humana. Tomemos os seguintes três níveis devidamente exemplificados:

- 1 - Atitudes proposicionais: António *acredita* que a bola é vermelha
- 2 - Pensamentos: “A bola é vermelha”
- 3 - Conceitos: BOLA, VERMELHO

Uma semântica para os estados mentais pode ser equacionada como uma semântica para a psicologia (uma *psicosemântica*). A questão central de uma psicosemântica pode ser enunciada do seguinte modo: como é que conceitos como BOLA, VERMELHO, etc., adquirem o conteúdo semântico/intencional que é suposto terem.

A maior parte do artigo consiste numa exposição dos pontos de vista antagónicos de F&L e PC a respeito desta questão, registando uma dinâmica dialéctica de resposta/ contra-resposta (nem sempre diacrónica). Na secção final desenvolve-se um diagnóstico crítico acerca do conteúdo e pertinência da própria polémica.

---

<sup>1</sup> Trabalho realizado no âmbito do projecto *Rationality, Belief, Desire II – from cognitive science to philosophy*, (POCTI/FIL/5555/2004) do Gabinete de Filosofia Moderna e Contemporânea do Instituto de Filosofia da FLUP (Unidade I&D 502 FCT).

## II - Similaridade e semântica

Os primeiros dados são lançados por F&L na sua obra *Holism: a Shoper's Guide* (1992) onde avaliam criticamente uma série de projectos semânticos sugeridos por vários filósofos, que tentam responder à questão atrás enunciada acerca do modo como os conceitos adquirem o seu conteúdo semântico. Antes de mais, convirá esclarecer em que consiste uma teoria semântica pelos padrões de F&L. Fundamentalmente uma teoria semântica deverá satisfazer conjuntamente os seguintes critérios:

(1) - *Condições de satisfação*: deve fornecer um critério de satisfazibilidade dos itens de uma linguagem natural – palavras- como de uma psicologia – conceitos- de tal modo que frases/pensamentos que julgamos serem verdadeiras ou sejam segundo a teoria em questão. Eg. “A neve é branca” deverá ser verdadeira segundo a teoria

(2) - *Composicionalidade*: tem de explicar como a produtividade e sistematicidade do pensamento e linguagens humanas são possíveis

(3) - *Tradução*: o significado é aquilo que a boa tradução preserva. Uma teoria semântica deve fornecer uma noção de significado em conformidade com esta intuição

(4) - *Explicação intencional*: a teoria semântica deve reconstruir uma noção de conteúdo que seja adequada aos propósitos de uma explicação intencional (atitudes proposicionais)

O ponto geral enunciado por Fodor & Lepore (1992) é o de que, uma vez não aceite a distinção analítico/sintético (distinção posta em causa por Quine) qualquer projecto semântico que satisfaça estas quatro condições é inviável à partida. Muito breve e sucintamente, Quine mostrou no clássico artigo “Two Dogmas of Empiricism” (1951) como a clivagem entre enunciados analíticos e sintéticos não era tão óbvia e segura como a tradição fazia parecer crer. A base da crítica quineana à distinção analítico/sintético assenta na dificuldade, diagnosticada por Quine, em estabelecer critérios de identidade de significado entre duas expressões. Ora, precisamente, o critério de identidade de significado entre duas expressões consiste na pedra basilar de qualquer teoria semântica.

A alternativa parece ser a adoptar uma perspectiva holista acerca do significado. No entanto, tal concepção não reconstrói a noção de identidade de significado e, por essa razão, não consegue estabelecer uma teoria semântica (que satisfaça os quatro critérios enunciados). No entanto, F&L salientam que: «[É] largamente assumido que, mesmo que o holismo proíba uma noção robusta de *identidade* de conteúdo, ele possibilita, não obstante, uma noção robusta de *similaridade* de conteúdo» (Fodor & Lepore, 1992, 17, ênfases no original)

Mas em que consiste exactamente a similaridade de significado/contéudo? A noção de similaridade de significado e o projecto de reconstruir uma semântica a partir dessa noção poderá talvez ser ilustrada mediante uma analogia aritmética. Tome-se o símbolo matemático “ $\approx$ ” que se pode ler como “...é aproximadamente idêntico a...” e que pode ser ilustrado da seguinte forma:

$$2+2 = 4$$

$$2,001+2 \approx 4$$

Suponha-se agora que, por uma bizarra razão, um matemático desconfiava da legitimidade científica da relação entre números e que se proporia reconstruir toda a aritmética na base da relação “ser aproximadamente idêntico”. Assim, para este matemático impõe-se reconstruir a aritmética substituindo a relação de igualdade em  $2+2=4$  por  $2+2 \approx 4$ . *Mutatis mutandis* alguns autores propõem substituir a noção de identidade de significado pela de similaridade de significado, eg: em vez de se afirmar que “Celibatário” e “Homem não casado” têm o mesmo significado afirmamos que têm significados similares.

Colocam-se duas questões centrais acerca da noção de “similaridade” de significado:

1ª: É ela capaz de substituir a noção de “identidade de significado” na construção de uma semântica que satisfaça os critérios (1)-(4)?

A resposta de F&L a esta questão consiste num rotundo “Não” – omite-se aqui a suposta demonstração de F&L desenvolvida em Fodor & Lepore, 1999.

2ª: Já alguém propôs uma reconstrução positiva da relação de “similaridade de significado”? E foi bem sucedido?

À primeira parte da questão F&L respondem “Sim, Paul Churchland”, à segunda parte, de novo, “Não”.

### **III - Paul Churchland e a noção de “similaridade de significado”**

PC desenvolve um modo de tornar mais claras e objectivas as noções de similaridade de conteúdo de representações mentais ou conceitos. Em particular, PC fornece um *critério* de similaridade de conteúdo mediante a adopção de uma *métrica* de similaridade. Assim, o que era vago (a relação de similaridade) torna-se ‘objectivo’ e mensurável.

A proposta de PC parte da ideia de que o cérebro representa a realidade por meio de activações de populações neuronais. Para efeitos de ilustração de como tal activação é processada pelos neurónios PC recorre ao aparato formal fornecido pelas redes neuronais artificiais desenvolvidas informaticamente. Uma rede neuronal artificial é composta por dois tipos de componentes: as ‘unidades’ (que correspondem *grosso modo* aos próprios neurónios) e as ‘conexões’ (que correspondem às sinapses). Na figura 1 (retirada de Churchland, 1998, 82) as unidades são representadas pelos círculos e as conexões pelos traços que estabelecem ligações entre essas unidades. A rede neuronal está organizada por camadas de unidades, no caso em consideração distinguem-se três: uma camada de *input*, uma camada intermédia, e uma camada de *output*. Cada unidade recebe um determinado valor numérico (por exemplo, a cada unidade da

camada de input é atribuível um valor em função do input apresentado) e transmite uma determinada informação (também codificada numericamente) às unidades da camada subsequente às quais está ligada via conexões. Podemos ver isso ilustrado na figura 1 onde cada célula da camada intermédia (X, Y e Z) adquire um certo valor de activação quando um determinado item é apresentado na camada de *input*. Por outro lado, a figura ilustra como cada camada (neste caso a camada intermédia) pode ser representada como definindo um espaço de estado em que cada uma dessas células (da camada intermédia) corresponde a um eixo/dimensão do espaço de estado que a camada como um todo define. O valor adquirido pela activação de cada uma das unidades (X,Y e Z) da camada corresponde a um triplo ordenado (<.7, .3, .5>) que consiste num vector ou ponto no espaço de estado.

Um determinado item/conceito corresponde a um ponto/vector num determinado espaço (o seu conteúdo é determinado pela posição que ocupa no espaço).

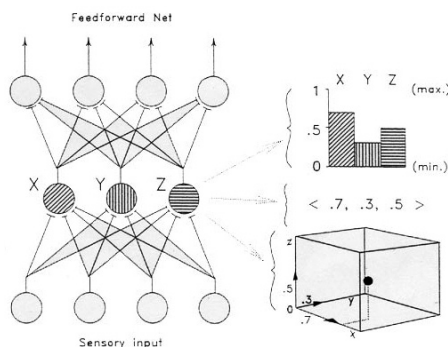


Figura 1

A partir deste enquadramento, PC desenvolve facilmente a noção de “Similaridade num espaço de estado”. O critério de similaridade num espaço de estado é ilustrado pela figura 2 (retirada de Churchland, 1998, 86): quanto mais próximos estiverem num espaço dois itens/conceitos, mais similares serão. Assim, por exemplo, o conceito HATFIELDS é mais similar em termos de conteúdo/significado do conceito ANDERSONS do que do conceito MCCOYS. O recurso aos espaços de estado definidos pelas camadas das redes neuronais fornece-nos um modo claro e intuitivo de definir ‘similaridade de conteúdo/significado’.

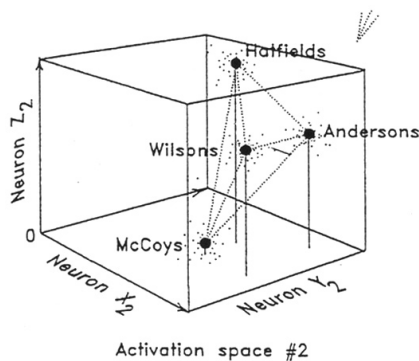


Figura 2

Muito sumaria e simplificada, é esta a proposta de PC para reconstruir a noção de ‘similaridade de significado’.

#### IV - Críticas de F&L à proposta de PC

De acordo com o critério de PC, dois conceitos são similares desde que ocupem posições similares no mesmo espaço de estado. A questão que se coloca de imediato é a de saber quando é que dois espaços (vamos supor, em pessoas diferentes) podem ser identificados. A resposta é a de que dois espaços são o mesmo se tiverem as mesmas dimensões. As dimensões correspondem a propriedades semânticas (eg. os conceitos GERORGE W BUSH e NELSON MANDELA ocupam posições próximas na dimensão “ser presidente”, menos próximas na dimensão “ser eleito” e extremamente distantes nas dimensões “ser inteligente” e “ser diplomata”). No entanto, vários são os problemas que F&L imputam à concepção de PC assim considerada.

##### 1 - Individuação/identificação das dimensões

Se as dimensões são (de certa forma) primitivos semânticos, como adquirem elas o seu conteúdo semântico? Em particular, se temos um critério de individuação/identidade da relação de similaridade invocando a identidade de espaços que, por sua vez é explicado em função das dimensões (têm as mesmas dimensões). Mas qual o critério de individuação/identidade das dimensões? Responder que duas dimensões são a mesma quando têm o mesmo significado/conteúdo semântico é retroceder ao problema de origem que, supostamente, seria contornada pela proposta de PC. Ele não responde a esta questão central parecendo tomar como adquirido o conteúdo semântico das dimensões.

Existe uma saída para esta dificuldade e que consiste em adoptar o ‘velho’ princípio empirista segundo o qual as dimensões corresponderiam a propriedades sensoriais simples (cujo conteúdo semântico seria garantido, por exemplo, mediante uma teoria acerca da relação causal com o meio exterior) e os conteúdos dos demais conceitos (complexos) corresponderiam à combinação (booleana ou, mais provavelmente, estatística) desses simples em espaços de estado definidos por esses simples. Assim, poderíamos ter, inclusive, um critério de identificação de espaços de estado (sem ser

*question begging*) e apelando directamente a órgãos sensoriais: dois organismos partilham o mesmo espaço semântico se os seus órgãos sensoriais forem idênticos em todos os seus detalhes (Fodor & Lepore, 1996a, 159)

A situação de PC corresponde a um dilema: ou escolhe a saída empirista e então a sua teoria é estéril, uma vez que é unanimemente reconhecido que o empirismo clássico é *falso* (como Quine uma vez afirmou a este propósito: “o impasse humeano é o impasse humano”) ou, alternativamente, não aceita essa posição empirista e acaba por ter nas mãos, de novo, o problema da identidade de significado o que significaria que a sua teoria estaria a cair numa petição de princípio.

## 2 - Escolha das dimensões: como determinar o que conta como um primitivo semântico?

Como é que Churchland decide o que é representado pelas dimensões dos espaços de estado e o que é representado por regiões nos espaços de estado que as dimensões definem (Fodor & Lepore, 1996a, 154)

Por exemplo: porque é que determinamos TOMATE por relação a um espaço de cores (os tomates são vermelhos) e não o contrário; determinamos VERMELHO por relação a um espaço de vegetais (VERMELHO tem um alto valor na dimensão “tomate”)? Em suma, como determinar o que conta como um primitivo semântico (dimensões)?

Temos de novo um dilema em tudo análogo ao anterior: a perspectiva empirista resolve este imbróglgio mas é falsa, a alternativa consiste em assumirmos uma situação em que esta questão não se resolve de uma forma que não seja totalmente arbitrária.

## 3 - Informação idiossincrática e número de dimensões

A relação de similaridade entre dois conceitos é, como vimos, definida como posições similares no mesmo espaço de estado. Um espaço de estado é idêntico a outro se têm ambos as mesmas dimensões. Mas é altamente implausível que sujeitos diferentes partilhem espaços com as mesmas dimensões dada a variação idiossincrática da informação que temos a respeito de um determinado item ou conceito.

Por exemplo: um veterinário tem muito mais informação acerca de cães do que eu. Assim sendo, o meu espaço de estado que define o conceito CÃO tem, em princípio, muito menos dimensões do que o do veterinário.

O problema está em que, se assim for, o critério de similaridade é apenas um critério de similaridade *token* (relativo a um determinado indivíduo num certo espaço de tempo) e não de similaridade *type* (intersubjectiva). Mas, neste caso, uma noção de similaridade token não é suficientemente robusta para construir uma teoria semântica.

## **V - Resposta de PC às críticas de F&L**

1 - PC reconhece como legítimas as críticas de F&L (pelo menos em Churchland, 1998)

PC reconhece que: a) Uma certa forma de empirismo humeano poderia estar implícita na sua formulação.

b) O modo de fornecer conteúdo às dimensões (em conformidade com o pressuposto empirista) seria o de estabelecer relações causais com o exterior (no caso de serem dimensões/neurónios sensoriais) ou assumindo micro-características de conteúdo desses neurónios sensoriais se não fossem neurónios/dimensões sensoriais.

No entanto, assumir este pressuposto conjugado com dois factos indisputáveis acaba por acarretar uma consequência fatal para o projecto construtivo de uma semântica. Esses dois factos são:

i) - Uma enorme idiosincrasia ao nível das dependências causais com o exterior (são diferentes de indivíduo para indivíduo)

ii) - O número de neurónios (mesmo aos níveis mais periféricos) varia muitíssimo também de indivíduo para indivíduo. Logo, o número de dimensões para cada espaço nunca será o mesmo intersubjectivamente.

O resultado de conjugar a estratégia enunciada em a) e b) com os factos i) e ii) é o de tornar impossível reivindicar, neste caso, uma relação de similaridade *type*, o que arruina o próprio projecto semântico.

2 - A solução para esta dificuldade passa por reformular o critério de ‘avaliação semântica’. O ponto fulcral e que marca a divergência fundamental com a tese implícita em a) e b), é o de retirar qualquer importância ao papel das dimensões. Assim, ao invés:

Um ponto no espaço de activação adquire um determinado conteúdo semântico *não* em função da sua posição relativa aos eixos constituintes mas em função de:

1 - A sua posição espacial relativamente a todos os outros pontos com conteúdo dentro desse espaço

2 - As suas relações causais com macro-características estáveis e objectivas do ambiente exterior (Churchland, 1998, 84-85)

A esta perspectiva, e sob sugestão de PC, podemos designar de *Dual Aspect Semantics* (DAS). Ela conjuga uma *Conceptual Role Semantics* (CRS) com uma teoria causal do conteúdo. Sendo que, para uma CRS um item mental/linguístico assume um determinado conteúdo semântico pelo papel inferencial que obtém com outros itens mentais/linguísticos. Para uma teoria causal um item obtém conteúdo mediante uma relação causal que estabelece com factos ou características do mundo exterior.

3 - Demonstração de como a DAS contorna o aparente problema i) (idiosincrasia das dependências causais)

Tomem-se duas redes artificiais idênticas treinadas para a mesma função como se ilustra na figura 3 (retirada de Churchland 1998, 86)

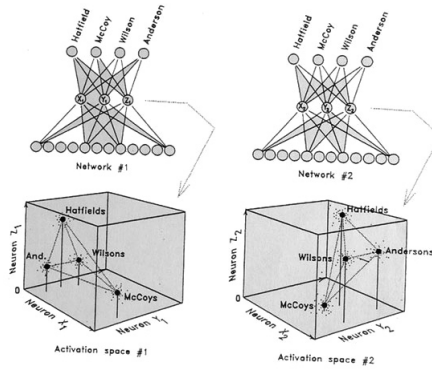


Figura 3

Estas duas redes foram treinadas para reconhecer faces de quatro famílias diferentes. Após o treino pedia-se à rede para identificar as fotografias como pertencendo a uma das quatro famílias (incluindo fotografias diferentes daquelas que tinham sido utilizadas no processo de treino). Ambas as redes apresentam um alto nível de sucesso (muito próximo entre as duas). É ao nível das camadas intermédias de cada rede que se explica esse sucesso: cada rede desenvolveu pontos/vectores protótipo para cada uma das famílias. Pontos que funcionam como ‘centros de gravidade’ conceptual (uma face que aproxime de um dos protótipos é identificado como pertencente a essa categoria). É a organização conceptual da rede. A ideia de ponto/vector protótipo é aqui decisiva e pode ser melhor ilustrada tendo em consideração a figura 4 (retirada de Churchland 1995, 83). Depois de sujeita à aprendizagem de discriminação de rochas e de minas a rede cria, por si mesma, protótipos desses dois tipos de entidades que não correspondem a nenhum item particular que tenham presenciado (o protótipo de ‘mina’, por exemplo, corresponde a uma espécie de somatório das características que a rede reconhece em todas as minas). Note-se a partição do espaço em dois sub-espacos: um para rochas, outro para minas

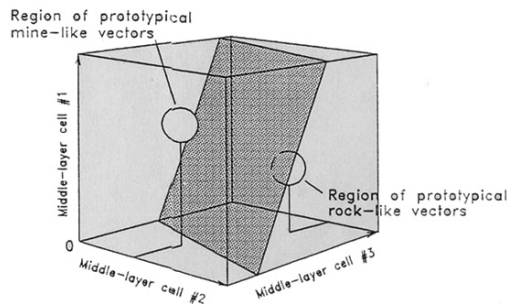


Figura 4

Como é perfeitamente visível na figura 3, cada uma das redes adoptou estratégias diferentes para resolução do mesmo problema (note-se as diferentes posições dos vec-



tores protótipos em cada um dos espaços). O que temos aqui ilustrado é um caso de idiosincrasia causal (o 'facto' i)). Cada uma das redes desenvolveu estratégias diferentes (diferentes conexões sinápticas e mesmo dimensões/eixos com conteúdos diferentes) e tal deve-se ao facto de cada uma das redes ter tido uma abordagem diferente com os objectos exteriores (por exemplo uma amostragem das fotografias em sequências diferentes para cada uma das redes no processo de aprendizagem).

No entanto, a performance computacional das redes é idêntica, isto porque:

1º - Os diversos vectores protótipos assumem distâncias e posições relativas idênticas (o sólido cujos vértices são os vectores protótipos é idêntico em ambos os espaços)

2º - Cada vector protótipo constitui a resposta à relação causal de cada item mostrado à camada de input (neste caso, as fotografias das faces dos elementos das famílias)

É, pois, deste modo que a DAS supera o problema da idiosincrasia das dependências causais

Generalização: é possível generalizar uma equação de medição de similaridade para cada tipo de hiper-sólido definido pelos vectores protótipos enquanto vértices para qualquer espaço de estado.

4 - Demonstração de como a DAS contorna o aparente problema ii) (o diferente número de dimensões)

Aqui a solução é ainda mais clara e simples de compreender. Como a figura 5 (retirada de Churchland, 1998, 90) ilustra, um espaço de D dimensões pode incluir sempre um hipsólido de n-D dimensões. E pode mesmo incluir um hiperplano de n-D dimensões onde esse hipsólido é definido. O que interessa é que as distâncias e relações entre os pontos/ vectores protótipos se mantenham entre si em qualquer dos casos. Ou seja, aplicando uma vez mais a DAS vemos como isso não cria qualquer tipo de problema (não interessam as dimensões, o seu número ou o seu conteúdo semântico particular. O que interessa é a relação entre os vectores protótipos e suas relações causais)

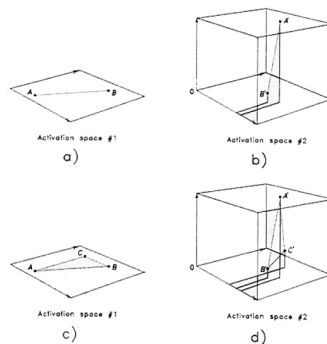


Figura 5

## 5 - As experiências Laakso-Cottrell

Estas experiências (utilizando um critério de similaridade conhecido como *Gutnam point alienation* (GPA)) mostram como existe similaridade mesmo nos casos em que os órgãos sensoriais são dissimilantes (relembre-se a objecção/observação de F&L no ponto 1 da secção IV).

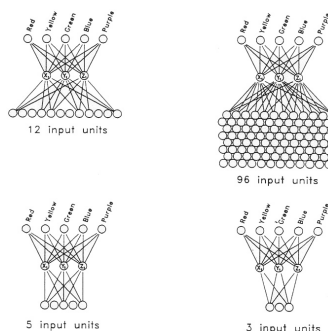


Figura 6

A figura 6 (retirada de Churchland, 1998, 100) ilustra quatro redes treinadas para a mesma função, a saber: reconhecer uma dada cor apresentada como estímulo. A grande diferença entre as quatro redes reside no diferente número de unidades de *input* (correspondendo, assim e *grosso modo* a órgãos sensoriais diversificados na sua estrutura). Estas redes não diferem apenas no número de unidades de input mas, sobretudo, no modo como esse input é processado. Utilizando o GPA Laakso e Cottrell conseguiram demonstrar que existe um assinalável grau de similaridade na organização interna (camada intermédia) das diversas redes. Tal resultado consiste em mais um ponto a favor da DAS e contra a perspectiva empirista clássica que F&L queriam imputar a PC.

6 - Finalmente, estes resultados acerca da mensurabilidade da similaridade podem ser generalizados a todo o tipo de redes (como as recorrentes e/ou com várias camadas intermédias)

## VI - Resposta de F&L à resposta de PC

1- Uma confusão por parte d PC marca toda a sua perspectiva e a convicção de que contornou definitivamente as críticas de F&L.

Essa confusão fundamental é entre *espaço semântico* e *espaço neuronal*

Espaço semântico: as suas dimensões/eixos correspondem a propriedades *semânticas*

Espaço neuronal: as suas dimensões/eixos correspondem a propriedades *físicas/neuronais*

A figura 7 (retirada de Churchland, 1995, 28 e 88) fornece-nos exemplos de cada um dos tipos de espaço: um espaço semântico de codificação facial e o espaço neuronal

correspondente á camada intremédia de uma famosa rede neuronal artificial desenvolvida e treinada para converter marcas escritas (palavras) em discurso verbal (a rede NET-talk). Note-se como no caso das faces as dimensões correspondem a propriedades semânticas (o tamanho da boca a largura do nariz, etc.). No caso do espaço da camada intermédia da rede NET-talk as dimensões correspondem a células e ao seu valor de activação e não a propriedades.

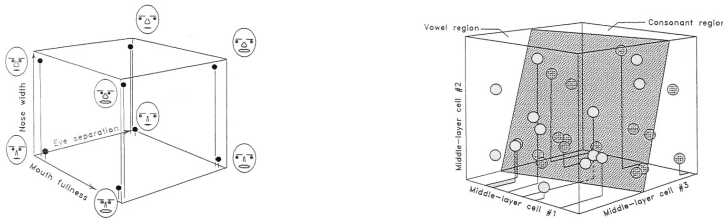


Figura 7

Esta confusão pode, de facto, inviabilizar todo o projecto de conceber uma semântica como F&L já sublinhavam no seu primeiro artigo que lançou a contenda:

Podemos estipular que as dimensões de um espaço de estado expressam sempre, e apenas, parâmetros psicofísicos (ou possivelmente neurológicos); de facto, pode bem ser isto que Churchland tem em mente. No entanto, se assim for o que ele nos está a oferecer não é de todo uma semântica. Pois uma semântica taxonomiza estados mentais pelo seu conteúdo e não pelas suas causas (Fodor & Lepore, 1996a, 154)

2 - Existem, de facto dois problemas paralelos (daí a confusão de PC) mas diferentes:

- 1 - Identificação de estados neuronais através de espaços neuronais
- 2 - Identificação de estados mentais/semânticos através de espaços mentais

A eventual solução para um dos problemas não pode, sem mais, ser considerada como uma solução para o outro (mesmo assumindo o fisicalismo):

Que a similaridade semântica seja preservada transdimensionalmente quando a similaridade neural o é não se segue de nenhum tipo de fisicalismo que tenhamos conhecimento (Fodor & Lepore, 1999, 400)

Seria necessário avançar com um argumento a favor de tal correlação e PC não o fornece. Na verdade, F&L fornecem um argumento que pretende provar exactamente o contrário:

Mesmo que se prove que existe identidade transdimensional de estados neurológicos, isso não prova que os conteúdos mentais também o sejam, na verdade, isso *não pode acontecer* porque a identidade dos conteúdos mentais *não é* preservada transdimensionalmente. Tome-se o seguinte exemplo:

Pessoa *a* – Espaço de três dimensões que define o seu conceito PEDRA

- 1ª dimensão: duro-mole
- 2ª dimensão: branco-negro
- 3ª dimensão: pesado-leve

Pessoa **b** – Espaço de quatro dimensões que define o seu conceito PEDRA. Além das três anteriores adiciona uma quarta: animado-inanimado. Para além disso a pessoa acredita que as pedras são entidades animadas.

O espaço da pessoa *a* pode ser incluído no espaço da pessoa *b* mas os seus conceitos PEDRA não são de todo idênticos nem mesmo similares. Esta situação é geral para todos os conteúdos mentais pois as suas condições de individuação são *não-monótonas*, i.e. a informação colateral/idiossincrática é fundamental nessa individuação

3 - No entanto, F&L vão ainda mais longe e demonstram que a proposta de PC não resolve sequer o problema da identificação transdimensional de estados cerebrais.

Os estados cerebrais tipo são individuados em termos da sua função, e a individuação funcional é, também ela, não-monótona. Por exemplo:

Um certo conjunto de *n* neurónios assume uma função *f*. Transferidos para um conjunto maior de neurónios (do qual acabam por constituir um sub-conjunto) não preservam a função *f* dadas, entre outros factores, as conexões que estabelecem com os outros neurónios (é pouco plausível que se preserve um sistema fisiológico separado apenas reservado aos neurónios iniciais).

## VII - Avaliação Crítica; Um Equívoco Fundamental

Limitou-se, até aqui, à apresentação da contenda mantida entre PC e F&L acerca da natureza da semântica, nomeadamente quanto à possibilidade de se construir uma semântica baseada na relação de similaridade de conteúdo. Para ilustrar tal polémica foram mobilizados os argumentos apresentados por uma e outra parte. No entanto, toda esta polémica parece, a meu ver, laborar sobre um equívoco a que os proponentes da contenda foram (consciente ou inconscientemente) alheios. Se tido em conta, esse equívoco questiona a própria pertinência do debate em consideração.

Parece-me que tanto F&L como PC se dão conta de uma diferença fundamental de perspectiva no modo como abordam o problema em causa mas acabam por remeter essa diferença para a periferia como se fosse um tópico marginal quando, na verdade, a sua importância é central para a avaliação de própria pertinência da discussão tal como ela é desenvolvida.

Essa diferença prende-se precisamente com o tipo de teoria semântica que cada parte preponente ambiciona dar conta. Para F&L, como se constatou na secção 2, uma semântica consiste numa teoria do significado que tem de satisfazer conjuntamente os quatro requisitos então enunciados (exibir condições de satisfazibilidade, composicionalidade, tradução e explicação intencional). De um modo mais explícito, os objectos de uma teoria semântica seriam itens linguísticos (palavras, frases) ou entidades mentais análogas a itens linguísticos (proposições e conceitos constituintes). PC é explícito quanto à sua diferente abordagem da semântica:

As abordagens clássicas do significado assumem como alvo principal a explicação de como o conteúdo se desenvolve ou está vinculado a itens *linguísticos* (i.e., em frases e nas suas palavras constituintes), ou nos análogos mentais directos de itens linguísticos (i.e., em crenças e nos conceitos constituintes). (...) [relativamente a uma semântica conexionista] estamos perante uma forma de economia cognitiva que tem pouco ou nada de linguístico, uma forma de cognição que é fundamental em todo o reino animal, uma forma de cognição que sustém a actividade linguística humana como apenas uma das últimas e talvez menos importantes das suas actividades cognitivas. (Churchland, 1996b, 282-283, ênfases no original)

PC encerra este considerando com a seguinte afirmação: “o conexionismo não tenta resolver o problema semântico de F&L focado na linguagem» (Churchland, 1996b, 283).

Infelizmente, tanto F&L como (um tanto surpreendentemente) o próprio PC parecem contornar esse facto ao longo da toda a sua polémica. F&L chegam a ser explícitos relativamente a esse alheamento; na sua última resposta a PC antecipam qual seria a reacção deste afirmando que «o que ele ripostaria, pensamos nós, seria em como nós falhámos em compreender quão radicais as suas propostas revisionistas realmente são; elas tencionam abstrair-se não apenas da ideia tradicional de que o constructo central da semântica é a identidade de conteúdo mas também do inventário de problemas tradicionais que as teorias semânticas supostamente solucionariam (...) mas nós não nos sentimos persuadidos em trocar o velho modo de conceber a semântica por um novo» (Fodor & Lepore, 1999, 391).

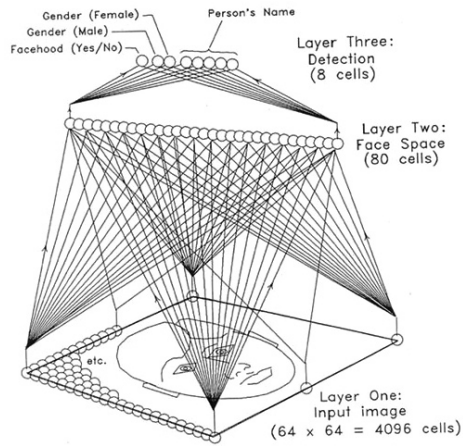
É pena que F&L não atentem com suficiente atenção a essa diferença de modo a notar que ela não consiste numa simples subtilidade. Se tivessem assumido essa diferença de perspectiva, verificariam que as suas críticas são estereis relativamente ao tipo de semântica sub-linguística advogada por PC. Para F&L «as posições ao longo das dimensões de um espaço semântico relativo ao qual um conceito está situado correspondem a atitudes proposicionais, a sustentação das quais é suposto ser constitutiva de ter o conceito.» (Fodor & Lepore, 1999, 395). Todo o problema com a atribuição de propriedades semânticas às dimensões é concebido neste quadro. Recordem-se as três críticas fundamentais movidas por F&L quanto a este problema e tal como enunciadas na secção IV: 1- A individuação/identificação das dimensões, 2- Como determinar o que conta como uma dimensão (determinar semanticamente TOMATE por relação a um espaço de cores ou, pelo contrário, determinamos semanticamente VERMELHO por relação a espaço de vegetais) 3- A questão da idiosincrasia da informação subjectiva na definição dos conceitos (recorde-se o exemplo do conceito PEDRA nas pessoas *a* e *b*). Explícita ou implicitamente estas observações críticas derivam da concepção linguística assumida por F&L a respeito dos espaços semânticos. No entanto, e como constatámos, o próprio PC reivindica uma abordagem diferente para esses espaços (nomeadamente o facto de que esse expediente se adequar a uma semântica de ordem sub-linguística).

Tentarei agora demonstrar como as críticas de F&L não colhem se adoptarmos o ponto de vista proposto por PC. No centro desta discussão estará a distinção fundamental entre codificação localizada e distribuída. Tome-se o seguinte exemplo para ilustrar e melhor compreender as diferenças entre os dois tipos de processamento de informação: um dicionário vulgar corresponde a caso típico de codificação localizada; a

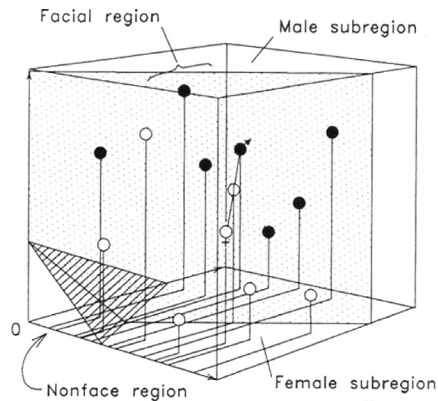
informação é armazenada localmente e do modo discreto. Se procurarmos informação a respeito da palavra 'avião' (por exemplo) ela estará localizada numa determinada entrada numa determinada página desse dicionário. Em que consistiria um dicionário em que a informação fosse codificada de um modo distribuído em vez de localizado? Seria seguramente um dicionário muito estranho; a informação de uma palavra como 'avião' não estaria localizada numa determinada entrada numa determinada página mas antes 'disseminada' por *todas* as páginas do dicionário. Uma forma de ilustrar melhor esta diferença será o de imaginar as consequências em cada um dos dicionários de se arrancarem algumas folhas: no dicionário 'standard' isso corresponderia à perda irreparável de alguma informação (por exemplo a definição da palavra 'avião' caso se arrancasse a página respectiva). No caso do dicionário distribuído essa perda de informação seria apenas parcial podendo mesmo, dependendo da extensão dos danos, ser desprezável.

É claro que um dicionário terá sempre de codificar informação de um modo localizado com entradas discretas ocupando uma localização precisa no espaço do livro. A ideia de codificação distribuída é extremamente contraintuitiva, não podendo nós imaginar como seria fisicamente possível conceber um dicionário que assumisse as características atrás enunciadas. No entanto, as redes neuronais podem, contrariamente aos dicionários 'standard', codificar a informação de um modo distribuído. Essa é uma das características que distinguem as redes neuronais de outros tipos de processadores de informação como os computadores clássicos que tratam a informação de um modo localizado, discreto e sequencial. As redes neuronais, por seu turno, adoptam um tipo de processamento paralelo e distribuído (PPD) (embora algumas redes possam processar a informação de um modo paralelo mas não distribuído).

Tome-se o exemplo de uma rede neuronal artificial desenvolvida por Cottrel e colaboradores que assume como função reconhecer faces humanas. A rede é constituída por três camadas: uma de input (com cerca de 4096 unidades) uma camada intermédia com 80 unidades, e a camada de output com cerca de 8 unidades (figura 8 retirada de Churchland, 1995, 40). Uma vez mais cabe à camada intermédia a tarefa de categorizar e classificar os inputs que lhe são apresentados. Esta camada define então um espaço a 80 dimensões (correspondendo cada dimensão a cada unidade) onde se localizam e classificam as diversas faces apresentadas em função da formação de vectores protótipos e consequente definição de sub-espacos categoriais que se constituem como partições do espaco original (ver figura 9, retirada de Churchland, 1995, 48).



*Figura 8*



*Figura 9*

No entanto, e ao contrário do caso ilustrado na figura 7, a cada um dos eixos ou dimensões desse espaço não corresponde nenhuma característica específica relativa a faces tal como o tamanho da boca ou a largura do nariz. Em vez disso, cada uma das unidades/dimensões codificam estruturas gerais e não traduzíveis de um modo linguístico. A figura 10 (retirada de Churchland, 1995, 48) documenta esta característica; exibe-se o campo receptivo de 6 das 80 unidades que compõem a camada intermédia (nenhuma delas assume claramente um estímulo discreto e identificável).



*Figura 10*

Mas, como PC salienta: «colectivamente estas dimensões abstractas sustêm uma estratégia de codificação que permite à rede estabelecer discriminações faciais até quase 100% de eficácia. Estas representações difusas, distribuídas e não composicionais não são únicas neste caso; elas constituem a codificação standard nas redes neuronais» (Churchland, 1996b, 279-280). Este tipo de codificação e processamento distribuído assume uma série de vantagens de adaptabilidade, poder de computação e resistência à degradação motivos que estarão na base da justificação evolucionária da sua adopção por parte dos sistemas nervosos biológicos (Cf Churchland & Sejnowski, 1992, 157-188).

Note-se o contraste entre este tipo de codificação distribuída com aquela proposta por F&L onde as dimensões de cada espaço assumiriam certos e determinados conteúdos semânticos bem delimitados (respeitando os pressupostos de um paradigma localizado e discreto a respeito do processamento e armazenamento de informação). É precisamente aqui que julgo poder detectar o falhanço das críticas movidas por F&L à proposta semântica de PC. As críticas de F&L não colhem porque, por e simplesmente, não faz qualquer sentido questionar pelo conteúdo concreto e linguístico de cada dimensão num espaço distribuído. Como atrás constatámos, as críticas de F&L repousam sobre a assunção de que os conteúdos semânticos das dimensões correspondem a conteúdos constituintes de atitudes proposicionais (corresponderiam a certos predicados como “ser vermelho”, “ser gordo”, “ser racional”, etc.).

Parece que as críticas de F&L a PC ficam sem efeito pois elas fazem mira sobre um outro alvo. Como testemunhámos já, PC assegura que «o conexionismo não tenta resolver o problema semântico de F&L focado na linguagem» (Churchland, 1996b, 283). A sua veledade é mais modesta se bem que não menos importante: fornecer as bases semânticas sub-linguísticas que sustêm a semântica linguística mais elevada. Estamos, pois, perante dois projectos semânticos distintos quanto ao seu alcance e domínio.

Lamentavelmente tal diferença é praticamente escamoteada ao longo da troca de argumentos entre as duas partes em contenda. Por um lado, e como se constatou já, F&L recusam-se a conceder qualquer tipo de crédito à proposta ‘radical’ de PC. Por outro lado, o próprio PC é, ao longo de toda a polémica, extremamente ambíguo



quanto às ambições de uma semântica reformulada em termos de redes neuronais. Na verdade, por um lado, como vimos, ele assume explicitamente que o tipo de semântica desenvolvido ao nível das redes neuronais não ambiciona colocar-se ao mesmo nível da semântica de estilo linguístico que F&L preconizam. No entanto, por outro lado, ele parece sucumbir à tentativa de entrar em diálogo com F&L quando na verdade, e pelas suas próprias palavras, tal diálogo é infrutífero e equívoco. PC parece por vezes desejar de um modo inconsistente o melhor de dois mundos, i.e.: por um lado uma semântica minimal e sub-linguística e, ao mesmo tempo satisfazer os critérios de uma semântica mais robusta que cumpra as exigências enunciadas por F&L.

Considero que PC não tem razões para tanto optimismo e ambição. Tomada por si só, uma teoria semântica fundamentada na activação de redes neuronais e traduzidas na posição vectorial num espaço de estado não consegue satisfazer os quatro critérios que, para F&L, definem uma teoria semântica (recorde-se sucintamente): 1- as condições de satisfazibilidade de palavras de uma linguagem natural, 2- o critério de composicionalidade, 3- o critério de tradução e 4- a capacidade de explicação intensional. O problema com a proposta de PC em satisfazer estes quatro critérios prende-se com o facto todos esses quatro critérios imporem desde logo como facto básico que a semântica em questão seja uma semântica que lida com conceitos. Ora, a semântica proposta por PC não chega a atingir esse nível conceptual (ela é sub-conceptual) uma vez que ela é meramente reactiva. A mera activação de um vector – correspondente à activação de uma rede *feedforward* – não equivale à posse do conceito cujo conteúdo corresponderia a esse vector. Aos conceitos é exigido que possuam a capacidade de representar itens por eles subsumidos e isso é algo que os vectores tal como concebidos até aqui, não fazem. Como afirma William Ramsey a este propósito «responder a um estímulo é uma coisa; representar esse estímulo é algo diferente» (Ramsey, 1997, 57). Os vectores num espaço de estado que represente a camada intermédia de uma rede neuronal limitam-se a classificar o input que lhes é apresentado de uma forma *on-line*; eles não assumem, só por si, nenhum atributo reflexivo e verdadeiramente representativo desse input, i.e.: eles não retêm a informação de modo a esta ser manipulada, lembrada ou evocada, tudo operações cognitivas que se requerem dos conceitos.

Uma semântica conexionista tal como PC a propõe, com activação de vectores em espaços de estado e posteriores transformações vectoriais ao longo de espaços subsequentes, constitui-se sem dúvida como um bom modelo do processamento mais básico de informação nos sistemas nervosos. E, como o próprio PC adverte, trata-se de uma semântica sub-linguística que é seguramente partilhada por formas de vida bastante rudimentares cujas capacidades cognitivas estão muito longe das capacidades racionais e linguísticas dos seres humanos.

A questão fundamental que subsiste será, pois, a de averiguar até que ponto existe um hiato qualitativo intransponível entre uma semântica de ‘nível baixo’ tal como a proposta por PC e a de ‘alto nível’ a que F&L fazem apelo. A resposta, não definitiva, a esta questão é a de que nada em princípio nos impede de adoptar uma perspectiva incrementalista e evolucionista segundo a qual a complexificação dos sistemas nervosos das criaturas mais simples resulta na emergência de capacidades semânticas mais desenvolvidas. Espera-se que num futuro próximo tal perspectiva incrementalista e evolucionista nos possa esclarecer a respeito de qual o melhor modo de desenvolver uma proposta de explicação da racionalidade humana.

João Fonseca  
IFL/ Universidade Nova de Lisboa

## Referências bibliográficas

- Churchland, Paul (1979); *Scientific Realism and the Plasticity of Mind*, Cambridge University Press
- Churchland, Paul (1989); "Some reductive strategies in cognitive neurobiology", in *A Neural Computational Perspective*, MIT Press, pp 77-110,
- Churchland, Paul (1995); *The Engine of Reason, The Seat of the Soul. A Philosophical Journey into the Brain*, MIT Press, Cambridge, Mass.
- Churchland, Paul (1996a); "Fodor and Lepore; state-space semantics and meaning holism", in *The Churchlands and Their Critics*, McCauley (ed.), pp 272-277
- Churchland, Paul (1996b); "Second reply to Fodor and Lepore" *The Churchlands and Their Critics*, McCauley (ed.), pp 278-283
- Churchland, Paul (1998); "Conceptual similarity across sensory and neural diversity", in Churchland & Churchland; *On the Contrary*, MIT Press, pp 81-112.
- Churchland, Patricia, & Sejnowski, Terrence (1992); *The Computational Brain*, MIT Press, Cambridge, Mass.
- Fodor & Lepore (1992); *Holism; A Shopper's Guide*, Blackwell, Cambridge Mass.
- Fodor & Lepore (1996a); "Paul Churchland and state space semantics", in *The Churchlands and Their Critics*, McCauley (ed.), pp 143-159.
- Fodor & Lepore (1996b); "Reply to Churchland", in *The Churchlands and Their Critics*, McCauley (ed.), pp 159-162
- Fodor & Lepore (1999); "All at sea in semantic space: Churchland on meaning similarity", in *Journal of Philosophy*, vol 96, pp 381-403
- Quine, W.V. O (1951); "Two dogmas of empiricism", in *From a Logical Point of View*, Harvard University Press, Cambridge Mass.
- Ramsey, William (1997); "Do connectionist representations earn their explanatory keep", in *Mind and Language*, vol 12, number 1