Entre autômatos e línguas: memórias de um aprendizado

Uma certa tarde, lá pelos idos de 1970, eu estava na biblioteca da PUC Minas, situada no prédio do antigo seminário, procurando algum livro que aproximasse minhas ideias sobre a teoria dos autômatos e as línguas naturais. Encontrei, em um volume de Psicologia, um longo artigo de Noam Chomsky sobre estruturas sintáticas e gramáticas gerativas.

Como conhecia bem a teoria dos autômatos finitos e das linguagens livres e também sensíveis ao contexto, entusiasmei-me imediatamente. Comecei a estudar as concepções linguísticas de Chomsky, em particular a distinção entre estruturas superficiais e profundas, procurando relacioná-las às línguas que conhecia ou desejava aprender: primeiro as línguas latinas e, posteriormente, o russo e o alemão.

Fiquei tomado pela ideia de que o conhecimento dessas estruturas poderia facilitar o aprendizado de várias línguas, até mesmo simultaneamente. Imaginava encontrar correspondências que permitissem passar de uma língua a outra com maior segurança, reconhecendo semelhanças de organização sob a diversidade das palavras. Na linguagem da matemática que me era familiar, aproximava essa esperança da ideia de isomorfismo, embora as correspondências entre línguas não fossem necessariamente exatas.

Passei então a comprar, ao longo do tempo, os livros de Chomsky. Entusiasmei-me particularmente com *Syntactic Structures* e enfrentei também a leitura de sua tese de doutorado, que me pareceu de enorme dificuldade, mesmo com minha familiaridade com linguagens formais, autômatos, semigrupos e operações de concatenação.

Parecia-me que ali se esclarecia algo essencial: a possibilidade de descrever, por meio de regras e relações formais, uma parte da extraordinária capacidade humana de produzir e compreender enunciados.

Essa descoberta se ligava, em minha imaginação, a uma cena doméstica muito simples: uma criança em seu cercadinho, na cozinha, ouvindo as palavras carinhosas da mãe enquanto ela lavava os pratos. Sem receber aulas de gramática, a criança começava a participar daquele universo de sons, relações e significados.

Eu associava essa aquisição à competência linguística de que falava Chomsky. A distinção entre estruturas superficiais e profundas oferecia-me uma maneira de pensar a organização das frases para além de sua aparência imediata. Convém, porém, separar os planos: essas estruturas eram representações propostas pela teoria linguística; sua relação com o funcionamento concreto do cérebro constituía outra questão.

Naquela época, eu já havia lido os trabalhos de outro pioneiro, Frank Rosenblatt, sobre perceptrons. Também conhecia um livro de Karl Pribram sobre fisiologia cerebral, que me parecera uma excelente porta de entrada para o assunto. Essas leituras se aproximavam, em minhas inquietações, dos estudos de linguagens formais e de suas bases matemáticas.

Poderiam esses elementos oferecer um caminho prático para a tradução entre línguas e para o aprendizado simultâneo de vários idiomas?

Em minha experiência, a hipótese mostrou-se fecunda, inclusive no estudo do russo e do alemão, especialmente em seus aspectos sintáticos. Reconhecer padrões de construção ajudava-me a organizar o aprendizado e a estabelecer relações entre formas de expressão.

Mas uma pergunta persistia: e o significado? Onde entravam as ideias em sua amplitude, as metáforas, as gírias, o humor, as alusões e tudo aquilo que uma frase comunica sem declarar inteiramente?

Cheguei a comprar um livro sobre o assunto, que acabou perdido na poeira das estantes. A pergunta, entretanto, permaneceu.

Mais recentemente, já na casa dos oitenta anos, decidi ampliar meus estudos com o chinês, particularmente o mandarim. Encontrei então um desafio que meus hábitos anteriores de aprendizado não me haviam preparado para enfrentar: os tons.

No mandarim padrão, diferenças de altura e de contorno da voz podem distinguir palavras. Uma mesma sequência de consoantes e vogais, pronunciada com tons diferentes, pode corresponder a palavras distintas. Seus quatro tons principais, além do tom neutro, exigiam de mim uma atenção nova à escuta e à pronúncia. O *pinyin*, sistema de romanização do mandarim, ajudava-me a registrar essas diferenças.

O belo edifício que eu construíra em minha interpretação de Chomsky precisou, então, ser ampliado. O estudo das estruturas sintáticas continuava valioso, mas não bastava para os desafios que agora se apresentavam. Era necessário considerar, com igual atenção, a organização sonora da língua, o vocabulário, o contexto e os usos culturais da expressão.

Minha surpresa voltou a crescer ao observar a rapidez com que sistemas de inteligência artificial generativa produzem traduções entre idiomas tão diferentes quanto o português e o chinês.

As redes neurais profundas abriram outro caminho para essa tarefa: aprender, a partir de grandes quantidades de dados, regularidades e relações úteis à produção de uma tradução. Nesses sistemas, a passagem entre línguas pode ocorrer sem uma gramática transformacional explicitamente programada.

Esse resultado recolocou minhas antigas perguntas. Que relações entre forma e significado podem ser aprendidas dessa maneira? Quanto do contexto se preserva na tradução? E o que acontece quando o sentido depende de uma metáfora, de uma lembrança compartilhada ou de uma referência cultural?

A possibilidade de traduzir com rapidez não encerrou minha investigação. Deu-me novas razões para continuar.

Assim chego ao fim destes prolegômenos sobre meu caminho pelas estradas do conhecimento das línguas humanas, antigas e modernas, incluindo o egípcio, o copta e as línguas maias, cuja escrita tanto me fascina.

Desenvolvida no mundo mesoamericano, em uma história marcada também por contatos entre povos, a escrita maia combinava sinais de palavras e de sílabas. Registrava acontecimentos, dinastias, rituais e concepções do tempo.

Vejo nela uma das belas manifestações da capacidade humana de dar forma ao conhecimento e de relacionar a experiência da vida aos mistérios da existência e do Cosmos.

Talvez seja essa a inquietação que une minhas leituras de autômatos, gramáticas e redes neurais: compreender como chegamos, por meio de sons e sinais, a construir e compartilhar sentidos.

Prof. SylvioSS.com.br

---

1. **Estrutura profunda não é estrutura cortical.** Em Chomsky, “profunda” designava um nível de representação gramatical, sem corresponder diretamente a uma região cerebral ou à estrutura inata da mente. O próprio autor advertiu contra essa interpretação. Por isso, é incorreta a identificação entre estrutura profunda e redes do hemisfério esquerdo., algumas vezes encontrada na literatura (Esclarecimento de Chomsky).
2. **Tons e sintaxe pertencem a dimensões distintas da descrição linguística.** O tom lexical distingue palavras; não equivale simplesmente à ênfase ou à posição do acento em português. Nada disso, isoladamente, demonstra uma incompatibilidade com a gramática gerativa. Descrição dos tons do chinês; pesquisa sobre percepção tonal.
3. **A pergunta sobre significado merece permanecer aberta.** A sintaxe contribui para a interpretação, mas não esgota metáforas, referências culturais e usos situacionais. Tampouco seria correto sugerir que Chomsky nunca tratou da semântica: ele investigou explicitamente suas relações com as estruturas sintáticas. Texto de Chomsky sobre interpretação semântica.

4. **O êxito da tradução neural não prova a inutilidade da teoria linguística.** Demonstra que é possível construir sistemas eficazes por outro procedimento. O artigo que apresentou o Transformer, por exemplo, avaliou a arquitetura em tarefas de tradução. Isso sustenta a descrição do avanço técnico, mas não resolve, por si só, as questões sobre aquisição humana e competência linguística. *Attention Is All You Need*.
5. **A escrita maia não surgiu em isolamento regional.** Os povos maias mantiveram relações com outros povos mesoamericanos. Sua escrita combinava sinais de palavras e sílabas e registrava também história política, além de temas religiosos. A interpretação espiritual pode ser citada aqui como uma apreciação pessoal. Guia do Museu Peabody; inscrições dinásticas de Copán.