



Escrita por algoritmos: pode um avião voar?

Algorithmic writing: can an aeroplane fly?

Ana Marques Silva^(a)

^a Centro de Literatura Portuguesa da Universidade de Coimbra – ana.marques.silva@gmail.com

Resumo: Este artigo resulta do ciclo de seminários “Pessoa 3.0”, organizado por Diego Giménez, em 2023. A sessão que fui convidada a integrar, intitulada “Pessoa e Inteligência Artificial”, foi dedicada a uma reflexão acerca da automação da linguagem e suas relações com a criação poética. Este texto considera algumas dessas relações, tomando a poesia de Fernando Pessoa e seus heterónimos como referência. Num primeiro momento, procura-se dar resposta ao mote lançado pelo organizador do presente dossier, estabelecendo um contexto para a reflexão sobre a escrita de poesia e a geração automática de texto. Num segundo momento, considera-se a pertinência da mimese na automação da produção simbólica, por oposição a uma ideia de algoritmo literário.

Palavras-chave: Automação. Linguagem. Fernando Pessoa. Mimese. Algoritmo literário.

Abstract: This article is the result of the ‘Pessoa 3.0’ seminar series organized by Diego Giménez in 2023. The session I was invited to take part in, entitled ‘Fernando Pessoa and Artificial Intelligence,’ was dedicated to reflecting on the automation of language and its relationship with poetic creation. This article considers some of these relationships, taking the poetry of Fernando Pessoa and his heteronyms as a reference. Firstly, it seeks to respond to the motto launched by the organizer of this dossier, establishing a context for reflection on the writing of poetry and the automatic generation of text. Secondly, I consider the relevance of mimesis in the automation of symbolic production, as opposed to the concept of “literary algorithm”.

Keywords: Automation. Language. Fernando Pessoa. Mimesis. Literary algorithm.

“A inteligência artificial é capaz de gerar textos a partir de outros textos e imitando estilos. Em que medida podemos repensar a obra de Pessoa a partir da I.A.?", pergunta o mote que deu origem a este texto. Como pensar universos tão distintos? Um possível ponto de partida para repensar a

obra de Fernando Pessoa à luz da inteligência artificial é o conceito de algoritmo: um algoritmo como um conjunto de instruções. Um algoritmo pessoano como um conjunto de regras pré-definidas que contêm os traços distintivos de cada um dos diferentes heterónimos. Cada biografia um pensamento, um gosto, uma sensibilidade, um conjunto de experiências, de referências, uma voz. E uma coerência, uma unidade que estrutura as relações entre cada um destes aspectos. Cito Fernando Pessoa: “A cada personalidade mais demorada, que o autor destes livros conseguiu viver dentro de si, ele deu uma índole expressiva, e fez dessa personalidade um autor, com um livro, ou livros, com as ideias, as emoções, e a arte dos quais, ele, o autor real (...), nada tem, salvo o ter sido, no escrevê-las” (Pessoa, 1966, p. 95; Pessoa, 2008, p. 267).¹

A construção dos heterónimos – criteriosamente pensada, com as coerências entre a vida e obra em cada heterónimo – permite a analogia do jogo, do algoritmo que define *a priori* as regras do jogo: o modo como cada etapa se desenrola, o tipo de biografia que cada personalidade expressa, o tipo de voz que cada biografia produz. Cito ainda Pessoa:

Cada personalidade dessas – reparai – é perfeitamente una consigo própria, e, onde há uma obra disposta cronologicamente, como em Caeiro e Álvaro de Campos, a evolução da pessoa moral e intelectual do autor é perfeitamente definida. Vede como isto se dá em Caeiro. Da limpidez primitiva (que nunca, eu, logrei compreender ou sentir) da impressão nativa, a evolução é directa, adentro de “O Guardador de Rebanhos”, para a aprofundação filosófica. (...) Depois, com a vinda da doença, a perfeita lucilação imaginativa ou sensível se apaga, e temos, nos poemas fragmentários finais do livro, em certo ponto ainda a continuação do aprofundamento, pela evolução do espírito do poeta, em outros pontos uma turbação da obra, pela doença final, real como as minhas mãos, a que, com mágoa

1

Disponível em Arquivo LdoD:
https://ldod.uc.pt/fragments/fragment/Fr585/inter/Fr585_WIT_ED_CRIT_SC.

minha que chorei em lágrimas, o grande poeta sucumbiu (Pessoa, 2008, p. 270-271).²

Construções depuradas do intelecto, a cada um dos heterónimos pessoanos corresponde um conjunto de aspetos, de funções poéticas ou, se quisermos, de regras predeterminadas: cada heterónimo, um programa. Cada poema, uma realização, uma variação desse programa concebido por Fernando Pessoa, desse programa entendido como *ideia* particular de texto poético, como síntese, princípio orientador, enfim, como estética. Um algoritmo literário será essa estrutura simbólica que sintetiza e condensa uma teoria do texto, traduzida em termos computáveis.

Mas quando programas de computador geram textos à maneira de Dante, ou de Camões, ou de Pessoa, esse é um mero trabalho de imitação. Não se trata de literatura. Para gerar textos ao estilo de um dado autor, um programa de computador analisa estatisticamente bases de dados com textos desse autor e abstrai dessa amostra textual um modelo. Várias perguntas se colocam, sendo uma delas a seguinte: de onde provêm os dados, a base de informação na qual o programa – tomemos o Chat GPT como exemplo – é treinado? qual é o *corpus* para gerar, por exemplo, poemas à maneira de Caeiro? Que amostras textuais usa a máquina como modelo?

Ao contrário do que acontece noutras contextos, a empresa Open AI, detentora do Chat GPT, não dá a conhecer os dados com que a sua tecnologia é treinada. Apesar de a sua utilização ser cada vez mais transversal e de, virtualmente, todo o conhecimento acumulado na rede constituir a matéria prima para a geração de novos textos, além de a própria utilização contribuir para a aprendizagem dos algoritmos, apesar destes três factores existe uma cortina opaca que não permite mais do

²

Disponível em Arquivo LdoD:
https://ldod.uc.pt/fragments/fragment/Fr586/inter/Fr586_WIT_ED_CRIT_SC.

que uma utilização superficial. A opacidade não se aplica apenas aos processos algorítmicos: não decorre apenas das especificidades no que respeita à dimensão técnica, mas decorre também das escolhas no que respeita à dimensão social destes sistemas. Esta é, porém, uma questão que extravasa o contexto deste artigo. Outras duas questões, mais próximas do âmbito literário que aqui nos move, são as seguintes: 1) de que modo a máquina lê estes *corpora*? e 2) de que modo a especificidade da leitura maquinica influi na abstração do modelo criado a partir dessas referências?

Consideremos a primeira questão, relativa à leitura algorítmica: para um algoritmo, ler não consiste em descodificar signos linguísticos: a palavra “cadeira”, por exemplo, não corresponde ao significante-objeto, nem corresponde ao significado-ideia: do ponto de vista computacional, uma palavra corresponde apenas a uma dada sequência de caracteres, cuja probabilidade de ocorrência e distribuição é calculada em função da análise estatística.

No seu livro *My life as an artificial creative intelligence* (2022), Mark Amerika comparou o método de escrita modernista designado como *stream of conscience*, ou fluxo de consciência, ao método de geração automática de linguagem em sistemas como o Chat GPT, sistemas estes calculam e preveem, palavra a palavra, a palavra que se seguirá numa dada frase. Um exemplo dado pelo autor é o romance *Ulysses*, de James Joyce. Ao longo do romance lemos o monólogo interior de uma personagem, expresso em frases (invocadoras de imagens, memórias, impressões) que se encadeiam incompletas umas nas outras, numa sintaxe em permanente derivação. Tal como no GPT, as palavras surgem uma após a anterior, determinadas pela anterior. Mas se esta analogia é pertinente quanto à sequência de formas verbais num eixo, o mesmo não se aplica ao conteúdo dessas formas, isto é, às razões que determinam que uma dada sequência contenha

determinadas palavras e não outras. Se em *Ulysses*, aqui entendido como metonímia dos processos de associação em seres humanos, uma palavra invoca outra palavra em função do som, do contexto, ou por via de qualquer outra associação, para um algoritmo, porém, a cada caractere e a cada palavra segue-se um/a outro/a determinado/a pela probabilidade de ocorrer na sequência do/a anterior.

O que determina essa probabilidade é a quantidade de ocorrências em que determinadas palavras se sucedem, isto é, o critério é sempre limitado aos padrões que estão contidos nas amostras textuais e, logo, o processo de associação é predeterminado. Para um ser humano, esse processo é aberto e indeterminado: uma palavra é um significante ao qual está associado um significado, significado esse que invoca sempre um novo significante com um outro significado, num processo de sucessiva e contínua associação e derivação, como aquela que Joyce explora em *Ulysses*. Porém, do ponto de vista computacional, é indiferente se os caracteres são letras do alfabeto, ou valores RGB num espectro de cor, ou representações gráficas de notas musicais. A noção de linguagem depende da leitura, da modalidade de visão, se quisermos, do agente linguístico: onde um ser humano lê sentidos (no plural), um agente algorítmico lê padrões e probabilidades de distribuição.

Retomemos agora segunda questão: de que modo a especificidade da leitura maquinária influi na abstração do modelo criado? E façamos um exercício, para procurar responder a esta pergunta, pedindo a um sistema de geração textual para gerar poemas no estilo de cada um dos três heterônimos pessoanos. Primeiro pedido ao Chat GPT: “Escreve um poema no estilo de Alberto Caeiro”. Resultado:

Claro, vou tentar criar um poema ao estilo de Alberto Caeiro, o heterônimo de Fernando Pessoa conhecido pelo seu estilo simples e direto, ligado à natureza e à contemplação do mundo.

O Coração da Terra

*Eu não penso, apenas sinto.
O sol é o mesmo, o campo é o mesmo,
e o vento é o mesmo que me toca.
Nada mais há para saber.*

*As flores são flores,
e eu sou uma coisa na flor.
A vida é um corpo em movimento,
sem perguntas, sem respostas.*

*A terra respira o seu mistério,
e eu, parte dela, não preciso de mais.
A paz está nos campos abertos,
onde o saber é só o ver.*

(...)

Espero que tenha captado um pouco da simplicidade e do encanto da natureza que Caeiro tanto apreciava.

Segundo pedido ao Chat GPT: “Escreve um poema no estilo de Álvaro de Campos”. Resposta do Chat GPT:

Álvaro de Campos, outro heterônimo de Fernando Pessoa, é conhecido por sua poesia modernista e introspectiva, frequentemente explorando temas como a modernidade, o sentimento de alienação e a busca por sentido em um mundo em rápida transformação. Vamos tentar criar um poema que capture um pouco desse espírito:

O Infinito na Cidade

*Vivo na cidade que não para,
entre aço e vidro, no frenesi incessante,
onde os sons são uma sinfonia de motores
e os rostos são máscaras que passam.*

*A vida pulsa com uma força estranha,
como se o tempo fosse uma máquina
e eu, um grão de areia em sua engrenagem,
em meio à fumaça e ao concreto.*

*Eu sinto o peso do infinito nas ruas,
o eco das estrelas sob o asfalto,
a vastidão de um cosmos aprisionado
entre prédios que não sabem de horizontes.*

(...)

Espero que este poema capture o espírito de introspecção e a sensação de deslumbramento que são característicos do estilo de Álvaro de Campos.

Terceiro pedido ao Chat GPT: “Escreve um poema no estilo de Ricardo Reis”. Resposta do Chat GPT:

Ricardo Reis, um dos heterônimos de Fernando Pessoa, é conhecido pelo seu estilo clássico e contido, que valoriza a serenidade, o equilíbrio e a reflexão estoica. Seu trabalho frequentemente explora temas como a fugacidade da vida e a busca pela tranquilidade interior. Vamos tentar criar um poema que reflete esses temas e o estilo de Ricardo Reis:

A Quietude dos Dias

No silêncio do entardecer,
onde o crepúsculo se mistura ao ouro,
encontro a calma dos pensamentos,
e a serenidade dos dias passados.

O tempo é um rio lento que se desdobra,
e cada momento é um reflexo tranquilo,
como o vinho que amadurece na taça,
ou a brisa que se apaga na tarde.

Deixo para trás o turbilhão das horas,
e entrego-me à paz de uma existência
onde a alegria é uma flor efémera,
e a tristeza, um eco distante.

(...)

Espero que este poema tenha captado a calma e a reflexão que caracterizam o estilo de Ricardo Reis.

Uma primeira constatação, de menor importância, mas ainda assim relevante: a língua portuguesa que o GPT utiliza é a variante brasileira. Isto acontece porque esta é a variante com um maior número de amostras online. O mesmo se aplica aos tradutores automáticos que encontramos online, por exemplo. Mas a constatação que aqui mais nos interessa reter é que, apesar das enormes diferenças de estilo entre os três heterónimos, para o GPT apenas os temas mudam. O vocabulário dos textos gerados invoca vagamente o sensacionismo de Caeiro, o futurismo de Campos e o estoicismo de Reis, revelando competência quanto às temáticas dominantes e quanto à composição gramatical. Porém, a forma permanece sempre a mesma e a mais simples possível: quadras. Por que quadras? Pela mesma razão pela qual a variante da língua portuguesa é a variante do Brasil: sendo infindável a variedade de modelos de poesia, alguns estão mais estabilizados do que outros e alguns são mais comuns do que outros – das cantigas medievais à poesia popular e ao rap, a quadra é a mais abundante amostra de todo o universo de estruturas poéticas e, logo, é aquela que o GPT mais facilmente identifica e replica.

Mas reconhecer e replicar uma estrutura formal não é o suficiente para gerar um poema. Uma das razões para tal truismo é a própria noção, bastante elástica e escorregadia sobretudo desde o Modernismo, do que é um poema. Como poderá então a máquina – o grande modelo de linguagem que é o Chat GPT – abstrair um modelo de poesia a partir das amostras de que dispõe? Se, por um lado, um modelo é, por definição, uma simplificação da coisa modelada, e se, por outro lado a linguagem poética é caracterizada pela ambiguidade, pela plurissignificação e pela densidade semântica dos enunciados, então torna-se evidente o fosso entre o modelo abstraído pelo GPT, isto é, a mimese de poema, e a noção culturalmente estabelecida do que é um poema. Façamos mais um teste rápido, um quarto pedido ao Chat GPT: “Escreve um poema experimental”. Resposta do Chat GPT:

Claro, vamos explorar um poema experimental, jogando com a forma, o ritmo e a linguagem para criar uma experiência única. Aqui está um poema que mistura fragmentação, jogos de palavras e formas não convencionais:

Fragments de Realidade

1.

Eco na sala de espelhos,
pontos de luz derretidos
em um calendário de água,
o tempo se dissolve no teto.

2.

Palavras em salto triplo,
dançam sobre as cordas de um piano invisível,
escrevendo a história do vento
em notas que não se tocam.

3.

Nuvens de café e relógios derretidos,
escrevem mensagens no espelho quebrado,
o silêncio é uma pintura de respiração
onde o cheiro é o som da cor.

(...)

Este poema busca explorar o fluxo livre de ideias e a desconstrução tradicional da forma poética, criando uma experiência de leitura que reflete a fragmentação e a imprevisibilidade do pensamento moderno.

A razão para este pedido residiu na expectativa de que ao adjetivo “experimental” correspondesse uma forma poética menos previsível. Mas o insucesso deste resultado e dos anteriores sugere que, para a máquina, um poema é simplesmente um texto em verso. À luz da leitura maquinária, que consiste fundamentalmente na análise estatística, a palavra “experimental” estará provavelmente próxima, em termos de ocorrências identificadas pelo sistema, das palavras “livre”, “desconstrução”, “fragmentação” e “moderno”. Este pedido também serve para demonstrar que, mesmo sem pedir à máquina para gerar um texto à maneira de um dado autor, o que está em causa é ainda assim o princípio da imitação: a máquina processa fontes,

identifica as ocorrências mais próximas dos termos usados no pedido, neste caso “poesia experimental”, e gera resultados.

Na verdade, o que estes exercícios demonstram é a futilidade de esperar que um algoritmo, por complexo que seja, gere poemas sem que tenha para isso sido explicitamente programado. Um algoritmo literário, programado não em função da mimese, mas em função de princípios estéticos da poética moderna e pós-moderna, produz textos que classificamos sem hesitação de literários: porque nas suas premissas – nas funções programadas – está computacionalmente representada uma ideia do que um texto literário poderá ser. São muitos os poetas que exploram a computação e a automatização da linguagem numa ótica literária³. Essa não é, claro, a ótica programada no Chat GPT. A sua competência prosódica resulta da mimese. E se em de textos informativos ou argumentativos a mimese produz resultados suficientemente sofisticados para levantar a questão da indistinção e o problema do simulacro, já em textos poéticos a insuficiência da mimese enquanto técnica coincide com as suas limitações enquanto princípio estético, preferido, desde o Romantismo, pelo princípio da expressão.

A poesia, enquanto arte da linguagem e enquanto manifestação da singularidade humana, desempenha no contexto da geração automática de linguagem a função de Teste de Turing, validando o grau de sofisticação destes sistemas. Uma máquina passa no teste de Turing se um ser humano que com ela interaja for incapaz para distinguir se está a interagir com outro humano ou com uma máquina. Quando em 1950 Alan Turing concebeu este teste, chamou a atenção para a ilusão de inteligência que resulta do domínio coerente de signos linguísticos. O que o teste

³ Veja-se, a mero título de exemplo, o trabalho poético de Jhave, realizado com redes neurais artificiais, ou o trabalho de John Cayley, com um foco nos efeitos da mediação computacional nas estruturas linguísticas, entre muitos outros artistas e poetas que exploram a computação como médium expressivo e como ferramenta capaz de expandir a criatividade humana e a noção de texto literário.

indica é a percepção que temos da máquina, e não as suas capacidades cognitivas em si mesmas. Noam Chomsky, o pai da linguística moderna, pôs esta questão em termos simples: dizer que uma máquina “pensa” é como dizer que um avião “voa”, ou que um submarino “nada”. Em nenhum dos casos o verbo utilizado se adequa, porém, à falta de melhores palavras são estas que, metaforicamente, usamos.

Pouco depois do surgimento do primeiro computador, eram dados os primeiros passos no estudo da inteligência artificial, nomeadamente através das primeiras experiências com a automação da linguagem, demonstrando a relação entre a teoria gerativa da linguagem e o desenvolvimento dos primeiros computadores enquanto máquinas de processamento linguístico (Portela, 2017, p. 187). Em 1952, Christopher Strachey programou um gerador de cartas de amor com um simples sistema combinatório, substituindo num sintagma determinadas palavras por outras da mesma categoria morfológica, previamente selecionadas. Com esta experiência, Strachey demonstrou também que um “truque relativamente simples” pode produzir a ilusão de que o computador está a pensar (Strachey, 1954). Doze anos mais tarde, entre 1964 e 1966, Joseph Weizenbaum programou o primeiro sistema artificial de conversação (*chatbot*), designado ELIZA, um sistema desenvolvido, nas palavras do próprio Weizenbaum, para “demonstrar a superficialidade da comunicação entre humanos e máquinas” (Epstein, 2001).

A sugestão de inteligência resulta, portanto, da sugestão de domínio da linguagem. Há muitas décadas que agentes de conversação artificiais são capazes de passar o Teste de Turing, mas nem por isso podem ser considerados inteligentes, pelo menos no sentido estabelecido por Ada Lovelace, a matemática que afirmou, ainda na primeira metade do século XIX, que uma máquina analítica (o futuro computador) só poderia ser considerada inteligente se fosse capaz de produzir um resultado para o

qual não tivesse sido programada. Passados duzentos anos, um sistema computacional, por complexo que seja, permanece determinístico: executa as premissas programadas, independentemente das camadas de computação e complexidade. Se as premissas forem estabelecidas com o objetivo de gerar texto numa lógica mimética da produção (ou da criatividade, ou da inteligência) humana, o resultado poderá ser mais ou menos aproximado desse referente, mas não deixará de ser uma mera imitação. Ter o humano como referente parece ser um requisito para a automação da produção simbólica. O mesmo não se aplica a outros domínios do uso de IA, em que os objetivos não passam pela imitação, mas antes pela descoberta de novas relações: pela capacidade de revelar informação de outro modo não legível à escala da cognição humana.

A mimese será talvez o uso menos criativo que podemos imaginar para um sistema computacional com a complexidade e o poder de processamento como o que encontramos no Chat GPT. Usar um tal poder computacional numa lógica não mimética, mas que permitisse explorar as possibilidades de outras lógicas, radicalmente diferentes do humano: talvez aí tivéssemos acesso a uma inteligência verdadeiramente *artificial*, que nos surpreendesse com algo novo para aprender, ao invés de uma inteligência “humana” *automatizada* e, logo, mais rápida, mais produtiva, mais previsível. Em causa está a distinção mais profunda e radical entre (a geração de linguagem por) um ser humano e (por) um agente algorítmico: o primeiro gera linguagem de um modo não determinado pelas circunstâncias; o segundo gera linguagem de um modo totalmente predeterminado. Afinal, um algoritmo programado para imitar humanos é o funcionário perfeito.

REFERÊNCIAS

EPSTEIN, J.; KLINKENBERG, W. D. From Eliza to Internet: A brief history of computerized assessment. *Computers in Human Behavior*, v. 17, n. 3, p. 295–314, 2001.

MUMFORD, Lewis. Authoritarian and Democratic Technics. *Technology and Culture*, v. 5, n. 1, p. 1–8, 1964.

PESSOA, Fernando; Soares, Bernardo. *Livro do Desasocego*. Ed. Jerónimo Pizarro. Lisboa: Imprensa Nacional-Casa da Moeda, 2011.

PESSOA, Fernando. *Poesias*. Nota explicativa de João Gaspar Simões e Luiz de Montalvor. Lisboa: Edições Ática, 1995.

PESSOA, Fernando. *Páginas Íntimas e de Auto-Interpretação*. Textos estabelecidos e prefaciados por Georg Rudolf Lind e Jacinto do Prado Coelho. Lisboa: Edições Ática, 1966.

PORTELA, Manuel. Writing under constraint of the regime of computation. TABBI, Joseph (Ed.). *The Bloomsbury Handbook of Electronic Literature*. Londres: Bloomsbury Academic, 2018.

SOARES, Bernardo; Guedes, Vicente. *Livro do Desassossego*. Editado por Teresa Sobral Cunha. Lisboa: Relógio d'Água, 4 ed., 2008.

STRACHEY, Christopher. "The "Thinking" Machine". *Encounter*, v. 3, n. 4, p. 25-31, out. 1954.

NOTAS DE AUTORIA

Ana Marques Silva (ana.marques.silva@gmail.com): É investigadora integrada no Centro de Literatura Portuguesa, e Gestora de Ciência e Tecnologia nesta unidade e no Centro de Estudos Clássicos e Humanísticos da Universidade de Coimbra. Em 2018, doutorou-se em Materialidades da Literatura pela Faculdade de Letras da Universidade de Coimbra, com uma tese dedicada às poéticas algorítmicas. É membro do MATLIT Lab – Laboratório de Humanidades, e do projeto "Arquivo Digital Colaborativo do Livro do Desassossego", onde atualmente estuda a receção crítica da obra de Bernardo Soares e a sua representação computacional.

Como citar este artigo de acordo com as normas da revista?

SILVA, Ana Marques. Escrita por algoritmos: pode um avião voar?. *Texto Digital*, Florianópolis, v. 20, n. 2, p. 86-99, 2024.

Contribuição de autoria

Não se aplica.

Financiamento

Fundação para a Ciência e a Tecnologia (FCT), Portugal.

Consentimento de uso de imagem

Não se aplica.

Aprovação de comitê de ética em pesquisa

Não se aplica.

Licença de uso

Este artigo está licenciado sob a Licença Creative Commons CC-BY. Com essa licença você pode compartilhar, adaptar, criar para qualquer fim, desde que atribua a autoria da obra.

Histórico

Recebido em: 2 out. 2024

Aprovado em: 4 out. 2024