

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE ARQUITETURA, ARTES E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO

Vinícius Laureto de Oliveira

AUDIOVISUAL E INTELIGÊNCIA ARTIFICIAL:
Processos de geração do sentido em textos autônomos

Bauru/SP
2019

Vinícius Laureto de Oliveira

AUDIOVISUAL E INTELIGÊNCIA ARTIFICIAL:

Processos de geração do sentido em textos autônomos

Dissertação apresentada ao Programa de Pós-graduação em Comunicação da Faculdade de Arquitetura, Artes e Comunicação da Universidade Estadual Paulista “Júlio de Mesquita Filho”, campus Bauru, como requisito para obtenção do título de Mestre em Comunicação, sob a orientação da Prof^a Adj^a Ana Silvia Lopes Davi Médola.

Bauru/SP
2019

Oliveira, Vinícius Laureto

Audiovisual e inteligência artificial: processos de geração do sentido em textos autônomos / Vinícius Laureto de Oliveira, 2019

125 f. : il.

Orientadora: Ana Sílvia Lopes Davi Médola




Dissertação (Mestrado) – Universidade Estadual Paulista. Faculdade de Arquitetura, Artes e Comunicação, Bauru, 2018

1. Audiovisual. 2. Inteligência artificial. 3. Semiótica.
4. Sentido. 5. Automação. I.

2. Universidade Estadual Paulista. Faculdade de Arquitetura, Artes e Comunicação. II. Título.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE VINICIUS LAURETO DE OLIVEIRA, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO, DA FACULDADE DE ARQUITETURA, ARTES E COMUNICAÇÃO - CÂMPUS DE BAURU.

Aos 25 dias do mês de setembro do ano de 2019, às 00:00 horas, no(a) Auditório dos Programas de Pós-graduação da Faculdade de Arquitetura, Artes e Comunicação, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Professora Associada ANA SILVIA LOPES DAVI MEDOLA - Orientador(a) do(a) Departamento de Comunicação Social / Faculdade de Arquitetura, Artes e Comunicação de Bauru, Professor Assistente Doutor MARCOS AMERICO do(a) Departamento de Comunicação Social da Faculdade de Arquitetura, Artes e Comunicação da Unesp - câmpus de Bauru / Universidade Estadual Paulista, Professor Doutor ALAN CÉSAR BELO ANGELUCI do(a) Departamento de Comunicação / Universidade Municipal de São Caetano do Sul, sob a presidência do primeiro, a fim de proceder a arguição pública da DISSERTAÇÃO DE MESTRADO de VINICIUS LAURETO DE OLIVEIRA, intitulada **Audiovisual e inteligência artificial: processos de geração do sentido em textos autônomos**. Após a exposição, o discente foi arguido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: APROVADO. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.

Professora Associada ANA SILVIA LOPES DAVI MEDOLA 
Professor Assistente Doutor MARCOS AMERICO 
Professor Doutor ALAN CÉSAR BELO ANGELUCI 

"A year spent in Artificial Intelligence is enough to make one believe in God."

Alan Perlis - Epigrams in Programming (1982)

AGRADECIMENTOS

Deixo aqui meu agradecimento a todos aqueles que me auxiliaram durante este período do mestrado, principalmente os que não aguentam mais me ouvir falar de inteligência artificial:

A Deus e Nossa Senhora, que me ouvem em pensamento e, provavelmente, já estão entediados de tanto me ouvir pensar nisso.

Aos meus pais, Ní e Roberto, e a minha segunda mãe, Waldecy, que não faziam ideia do que eu estava falando, mas sempre apoiaram e incentivaram.

Aos amigos do Bar do Aeroclub (o melhor bauru de Bauru), Henrique e Isabella que me aguentavam falar de inteligência artificial após uma taça de vinho.

Aos colegas do GEA - Grupo de Estudos Audiovisuais, por horas incríveis de aprendizado e paixão pelo audiovisual (e alguns por inteligência artificial).

Aos professores e servidores técnicos da Pós-Graduação da FAAC com quem tive contato. Eu perdi alguns prazos, mas foi porque estava pensando em inteligência artificial.

À minha incrível namorada, Letícia, que me ouviu ao menos 30 minutos todas as noites do último ano falando sobre isso. E depois falou outros 30 minutos sobre o mesmo assunto.

À minha orientadora, professora Ana Silvia Lopes Davi Médola, que topou este projeto comigo quando nenhum de nós falávamos sobre inteligência artificial ainda. Agradeço pela cumplicidade na produção deste trabalho e por todos os ensinamentos.

RESUMO

Este trabalho realizou uma investigação acerca dos processos de criação de texto autônomos e a geração do sentido. Ferramentas digitais com recursos baseados em análises de dados e inteligência artificial inauguram novas etapas do processo de produção audiovisual. A partir da análise do roteiro de dois curtas-metragens gerados em uma rede neural de recorrências, avaliamos o processo de composição do texto e a formação do sentido. Para isto, utilizamos o aporte teórico-metodológico da semiótica discursiva francesa. Dentre os resultados obtidos, identificamos a interferência humana sobre o conteúdo criado de forma autônoma como a principal função geradora do sentido. O modelo de geração de texto com redes neurais abordado nesta pesquisa recorre somente a dados quantitativos para a geração dos conteúdos, mostrando-se ineficiente na construção semântica. A partir disso, propomos uma camada de interface entre o motor de geração de textos da rede neural e a composição final do texto, de modo que esta deve preencher requisitos do percurso gerativo do sentido, parametrizando a chamada de uma função recursiva de geração semântica.

Palavras-chave: Audiovisual; inteligência artificial; semiótica; sentido; automação.

ABSTRACT

This work carried out the processes of autonomous text creation and the generation of meaning. Digital tools with data analysis capabilities and artificial intelligence usher in new stages of the audiovisual production process. From the script analysis of two short films generated in a recurrent neural network, we evaluated the process of text composition and the formation of meaning. For this, we use the theoretical-methodological contribution of the French discursive semiotics. Among the results obtained, we identified human interference with autonomously created content as the main meaning-generating function. The text generation model with neural networks approached in this research uses only quantitative data for content generation, proving to be inefficient in semantic construction. From this, we propose an interface layer between the neural network text generation engine and the final composition of the text, so that it must fulfill the requirements of the generative sense path, parameterizing the call of a recursive semantic generation function.

Key-words: Audiovisual; artificial intelligence; semiotics; meaning; automation.

LISTA DE FIGURAS

Figura 1 - Quadrado semiótico	18
Figura 2 - Esquema narrativo canônico	20
Figura 3 - Eu, Robo (2004, EUA) e O Dia em que a Terra Parou (1951, EUA)	27
Figura 4 - O modelo matemático do Perceptron	32
Figura 5 - Apple Macintosh (1984)	35
Figura 6 - Plateia assiste a derrota de Gary Karparov em uma partida de xadrez para IBM Deep Blue.	40
Figura 7 - Captura de tela do curta-metragem Sunspring	42
Figura 8 - Esquema de como a combinação de dados e a tecnologia cognitiva permitem os vídeos de melhores momentos do torneio de Wimbledon	46
Figura 9 - Captura de tela do trailer de Morgan. O gráfico apresenta pontos no filme em que existem mudanças nas emoções.	48
Figura 10 - Montagem de telas após a análise das imagens dos filmes pela Google Vision	51
Figura 11 - Imagem da Venus de Willendorf, estátua com mais de 30 mil anos, censurada pelo algoritmo de reconhecimento de nudez do Facebook.	53
Figura 12 - Quadros da novela Espelho da Vida analisados por um modelo próprio de machine learning	54
Figura 13 - Quadros da novela O Sétimo Guardião analisados identificados pelo modelo.	55
Figura 14 - Texto gerado automaticamente pelo aplicativo Gboard apenas com palavras sugeridas	60
Figura 15 - Oscar Sharp e Ross Goodwin na palestra Machines Making Movies – TEDxBoston (2017)	63
Figura 16 - Benjamin gerando textos em papel fiscal	65
Figura 17 - Diagrama de cena: Conflito entre H e H2	72
Figura 18 - Personagens H e C encaram-se	73
Figura 19 - H encontra C morto e segura uma bolsa de sangue	77
Figura 20 - H2 confessa suas intenções para a câmera	79
Figura 21 - Repositório do código de WRT	86
Figura 22 - WRT sendo treinado com os roteiros	90
Figura 23 - Gráfico de explicabilidade por performance de aprendizagem	97
Figura 24 - Nível intermediário proposto	102
Figura 25 - Funcionamento da camada intermediária	105

SUMÁRIO

Apresentação	10
Considerações Iniciais	12
Capítulo 1: Inteligência Artificial no Auxílio à Textos Audiovisuais	25
1.1 – Uma breve história da inteligência artificial	26
1.2 – Invernos da inteligência artificial	34
1.3 – Novos patamares para a tecnologia	36
1.4 – Experiências audiovisuais	40
Capítulo 2: Sincretismo Audiovisual na Formação do Sentido	57
2.1 – Geração autônoma de texto	58
2.2 - Análise cena a cena de Sunspring	66
Capítulo 3: A Construção do Sentido em Textos Autônomos	84
3.1 - Processo de criação autônoma de texto	85
3.2 – Opacidade criativa	94
3.3 – Gerativismo da semântica	102
Considerações Finais	109
Referências Bibliográficas	113
Apêndices	118
4.1 – Roteiro do curta-metragem Sunspring em português	119

APRESENTAÇÃO

Esta pesquisa propõe uma investigação acerca uso de *softwares* que auxiliem em produções audiovisuais com a presença de recursos de inteligência artificial. Adotou-se como *corpus* da pesquisa dois roteiros de curtas-metragens, Sunspring e Orange. Sunspring foi produzido a partir de um roteiro escrito por uma inteligência artificial, o Benjamin. Orange foi criado a partir de um sistema inteligente para os propósitos desta pesquisa. O objetivo central desta dissertação é o de analisar a formação do sentido em textos gerados de forma autônoma, como os dos objetos. Desta forma, estruturamos o trabalho da seguinte maneira:

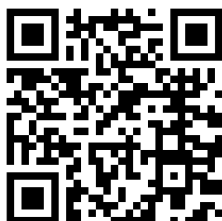
Nas *Considerações Iniciais* iniciamos uma apresentação sobre a semiótica discursiva francesa e o percurso gerativo do sentido. Alinhamos a metodologia com pesquisas já existentes sobre a geração de texto autônoma e o estudo sobre a formação do sentido a partir da semiótica. Além disso, estruturamos os parâmetros do percurso gerativo de sentido como metodologia de análise do corpus.

Em *Inteligência Artificial no Auxílio à Textos Audiovisuais* realizamos uma trajetória da comunicação e computação a partir de um viés dos sistemas inteligentes. Destacamos neste processo a evolução da tecnologia no meio comunicacional e a forma como os novos sistemas inauguram etapas no processo de produção audiovisual. Apresentamos neste capítulo o primeiro objeto do *corpus*, Sunspring, além de outros experimentos semelhantes.

No segundo capítulo, Sincretismo Audiovisual na Formação do Sentido, realizamos a análise semiótica do curta-metragem Sunspring. Durante a análise, identificamos como as linguagens que compõe o audiovisual são responsáveis por gerar o sentido em um curta-metragem que foi escrito por um computador. Após a análise, apresentamos algumas discussões sobre a geração de sentido e a criatividade em sistemas inteligentes. Em seguida, abordamos alguns tópicos sobre a antropomorfização de aplicações com inteligência artificial.

Em a *Construção do Sentido em Textos Autônomos* realizamos, por meio do WRT, o processo de construção do roteiro de Orange a partir do mesmo *corpus* de textos que Benjamin utilizou para a geração de Sunspring. A partir disto, pudemos identificar questões sobre a construção de texto de forma autônoma e como a geração do sentido se dá nestes meios. Por meio destas análises, sugerimos a avaliação de uma instância geradora de sentido para a construção de textos por meio de uma rede neural de recorrências.

Nas *Considerações Finais* apresentamos uma síntese sobre a geração semântica em textos autônomos e como uma nova instancia do sentido pode auxiliar neste processo. Finalizamos o texto com indicações de continuidades para esta pesquisa.



Sunspring

Link de acesso:

<https://www.youtube.com/watch?v=LY7x2lhqjmc>

CONSIDERAÇÕES INICIAIS

CONSIDERAÇÕES INICIAIS

A evolução das tecnologias da computação e análise de dados em grande escala, permitiram alguns dos recentes avanços da área de inteligência artificial (IA). Embora o termo *inteligência artificial* seja encontrado com maior presença nas páginas de notícias¹ apenas após o início da década atual, a área de IA já está entre as pesquisas da informática desde a metade do século XX. O termo tornou-se uma tendência entre as publicações de tecnologia, finanças, gestão de negócios e outras áreas, iniciando discussões éticas acerca do uso dos sistemas inteligentes. Manchetes, com tom futurista, apresentam situações comunicacionais que, até então, não tinham grande abrangência, como o uso de *bots* (robôs) na realização de comentários e ataques a perfis de candidatos no Twitter durante os debates televisionados para a eleição presidencial do Brasil em 2018², ou então, o uso de dados *minerados* do Facebook na eleição do candidato republicano Donald Trump em 2016 nos EUA³.

A utilização de dados em estratégias comunicacionais é profundamente pesquisada em áreas como a publicidade e o marketing, auxiliando a tomada de decisões sucintas. O fator novidade, oferecido pela utilização de ferramentas com

¹ Dado retirado da projeção do Google Trends de 2004 até 2019 com o termo *machine learning* (aprendizado de máquina). Disponível em: <https://trends.google.com.br/trends/explore?date=all&q=machine%20learning> Acesso em 5 mai. 2019

² “No primeiro debate presidencial, 10% dos tuítes foram de robôs ou fakes” Disponível em: <https://oglobo.globo.com/brasil/no-primeiro-debate-presidencial-10-dos-tuites-foram-de-robos-ou-fakes-22970273> Acesso em 5 mai. 2019

³ “Matriz da Cambridge Analytica participou de mais de 200 eleições, diz delator” Disponível em: <https://oglobo.globo.com/mundo/matriz-da-cambridge-analytica-participou-de-mais-de-200-eleicoes-diz-delator-22515244> Acesso em 5 mai. 2019

inteligência artificial para a comunicação, está no *insight* dos dados oferecidos ao sistema e a automação a partir destes. Desta forma, sistemas são treinados para realizar funções ligadas diretamente a cognição humana, como escrever mensagens de ataque ou defesa à candidatos durante os debates eleitorais. Avanços tecnológicos relacionados ao aumento do poder de processamento dos computadores, a velocidade das conexões em rede e a possibilidade de análise de grandes quantidades de dados, permitiram a expansão da IA para diversas frentes de pesquisa e mercado, chegando também às áreas da comunicação.

Grande parte das técnicas empregadas no desenvolvimento de sistemas inteligentes baseiam-se na reprodução de padrões, assim, um computador programado a partir destas técnicas pode identificar e copiar (ou criar) uma saída adequada aos dados de entrada. Novos sistemas com a capacidade de adequação as condições impostas pelos dados divergem as opiniões de pesquisadores e especialistas sobre o tema. As aplicações que utilizam a IA tomaram as mesas, bolsos e nuvens de usuário em volta de todo o planeta. A grande quantidade de dados gerados diariamente por usuários na internet, através das redes sociais e outros serviços, dá escopo para o treinamento de diversos sistemas inteligentes que podem atuar na área de comunicação, como os de sugestão e encadeamento de conteúdo, criação de notícias, personalização de marketing, *chatterbots*⁴ e outros. Com novos produtos comunicacionais sendo criados e influenciados a partir da lógica dos sistemas inteligentes, um novo desafio inscreve-se para a pesquisa de formação do sentido.

Nas relações sociais em um futuro não tão distante o computador e os sistemas relacionados, como os robôs detentores de corpo e os algoritmos desprovidos de corpo, não serão mais meros instrumentos de ações comunicativas de humanos ou meio pelo qual humanos se comunicam entre si. Ocuparão, em vez disso, a posição de outro ator social com quem alguém comunica ou interage. (GUNKEL, 2017)

⁴ Softwares capazes de conversar com usuários por meio de *chats* simulando a conversação com um humano através de aplicativos de redes sociais, como o Facebook e o WhatsApp

Em seu texto “Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação”⁵ o professor David J. Gunkel demonstra preocupações em relação aos caminhos dos sistemas inteligentes e a forma como estes já começam a atuar nos ramos da comunicação pelos meios digitais. O professor alerta para a necessidade de pesquisas que relacionem o campo das ciências sociais aplicadas com os sistemas autônomos, e como a comunicação responde a questionamentos que se parecem com retirados de um filme de ficção científica.

A preocupação com a interdisciplinaridade das pesquisas sobre IA com outras ciências relacionadas a comunicação e a formação do sentido tem seus primeiros registros nas décadas de 1960 e 1970, com os estudos de Dmitri Pospelov e Gennady Osipov em *Knowledge in Semiotic Models* (1977). No prefácio do livro *Semiotics and Intelligent Systems Development* (2007) organizados pelos professores Ricardo Gudwin e João Queiroz do Departamento de Engenharia de Computação e Automação Industrial da Unicamp, é ressaltado como algumas formas de pesquisa e metodologia muitas vezes expõem-se fora do escopo convencional. Os autores referem-se ao uso da semiótica para “estudar o processo do sentido e práticas comunicacionais relacionando os campos de ciências sociais e naturais, linguística e filosofia” (GUDWIN e QUEIROZ, 2007, p. 5) ⁶. Dado o princípio de que peças comunicacionais estão sendo criadas a partir de ferramentas que utilizam recursos baseados em IA, ou ao menos integram uma forma de distribuição inteligente, e que estas possuem um sentido que pode ter sido criado de forma autônoma por um sistema, podemos recorrer semiótica discursiva de linha francesa como abordagem teórico-metodológica para o estudo da formação do sentido nestes novos produtos.

O corpus desta dissertação foi definido a partir da pesquisa por projetos realizados de forma autônoma, ou quase, com o uso de ferramentas com IA,

⁵ Galáxia (São Paulo, online), ISSN 1982-2553, n. 34, jan-abr., 2017, p. 05-19. – Disponível em <http://www.scielo.br/pdf/gal/n34/1519-311X-gal-34-0005.pdf> Acesso em 5 mai. 2019

⁶Traduzido livre de: Study of meaning processes and communication practices within the fields of natural and social sciences, linguistics, and philosophy.

delimitando a escolha apenas a objetos audiovisuais. O primeiro objeto definido foi o curta-metragem *Sunspring*⁷ (2016, Reino Unido) e seu roteiro⁸, escrito inteiramente por um computador, o Benjamin. Para o propósito de uma avaliação mais profunda da forma de criação de um texto autônomo, o segundo objeto escolhido foi produzido para os fins desta pesquisa. O roteiro de *Orange* foi criado utilizando os mesmos dados de *Sunspring*. A escolha foi motivada devido ao caráter inaugural de *Sunspring*, anunciados como o primeiro curta-metragem escrito por uma IA. *Orange* foi construído para a realização de uma análise estrutural da geração do texto e do sentido.

Para analisarmos um produto audiovisual, primeiramente é preciso entender de que este é um texto sincrético, ou seja, articula diversas linguagens em sua formação, como a verbal-sonora, a musical, a iluminação, indumentárias e outras. A semiótica tem como objeto o texto (BARROS, 2005, p. 11), descrevendo e explicando “o que o texto diz” e “como faz para dizer”. Assim, o texto é um objeto de significação entre um destinador e um destinatário. Para a semiótica, o sentido resulta da articulação entre o plano de conteúdo e o plano da expressão.

O plano da expressão é plano onde as qualidades sensíveis que possui uma linguagem para se manifestar são selecionadas e articuladas entre elas por variações diferenciais. O plano do conteúdo é o plano onde a significação nasce das variações diferenciais graças as quais cada cultura, para pensar o mundo, ordena e encadeia ideias e discursos. (FLOCH, 2001, pg. 5)

Segundo FLOCH (2001), sistemas semióticos, como os textos que serão analisados, são objetos em que os planos do conteúdo e da expressão precisam ser estudados separadamente, pois não há conformidade entre eles. O linguista Louis Hjelmslev propõe a abstração, primeiramente, do plano da expressão, direcionando nossas análises ao plano do conteúdo, com suas estruturas de formação do sentido organizadas a partir do Percurso Gerativo do Sentido.

⁷ Disponível em: <https://www.youtube.com/watch?v=LY7x2lhqjmc> Acesso em 5 mai. 2019

⁸ Disponível em: <https://www.docdroid.net/ICZ2fPA/sunspring-final.pdf#page=2> Acesso em 5 mai. 2019

O percurso gerativo do sentido “é uma sucessão de patamares, cada um dos quais suscetível de receber uma descrição adequada, que mostra como se produz e se interpreta o sentido” (FIORIN, 2002, p. 17). Segundo BERTRAND (2003, p. 49) “seu procedimento clássico propõe articular a apreensão do sentido segundo um percurso estratificado em camadas relativamente homogêneas”. De acordo com BARROS (2005, p. 13) podemos organizar o percurso gerativo do sentido em três níveis, cada um deles “descrito e explicado por uma gramática autônoma, muito embora o sentido do texto dependa da relação entre os níveis”. Os níveis são organizados em estruturas semio-narrativas, o nível fundamental e o nível narrativo, e as estruturas discursivas com o nível discursivo. Esta organização em três camadas pode ser posta de forma que o sentido percorre do nível mais simples, abstrato e profundo para o mais complexo, concreto e superficial. “Os níveis estruturais se convertem uns nos outros, da profundidade à superfície, segundo um percurso de enriquecimento e complexificação” (BERTRAND, 2003, p. 49).

No primeiro nível, “a significação surge como uma oposição semântica” (BARROS, 2005, p. 13), da qual podemos depreender o quadrado semiótico: FIORIN (1999, p. 4) define esta oposição como A vs B, em que podemos investir valores opostos nestas variáveis, como *vida vs morte* ou *opressão vs liberdade*. Estas categorias fundamentais podem ser eufóricas ou disfóricas, dependendo da relação do estado do sujeito com o seu contexto. A partir desta oposição, podemos estabelecer as subcontrárias, que serão não A e não B.

Negando-se cada um dos termos da oposição, teremos não a vs não b. Os termos a vs b mantêm entre si uma relação de contrariedade. A mesma coisa ocorre com os termos não a vs não b. Entre a e não a e b e não b há uma relação de contraditoriedade. Ademais, não a mantêm com b, assim como não b com a, uma relação de implicação [complementariedade]. (FIORIN, 1999, p. 4)

Ainda no nível fundamental, verificamos o percurso do texto, passando de um dos termos de oposição para uma das subcontrárias e chegando a oposição do termo

inicial. O quadrado semiótico é um modelo sintático de organização dos sentidos no nível mais profundo e abstrato do percurso gerativo do sentido.

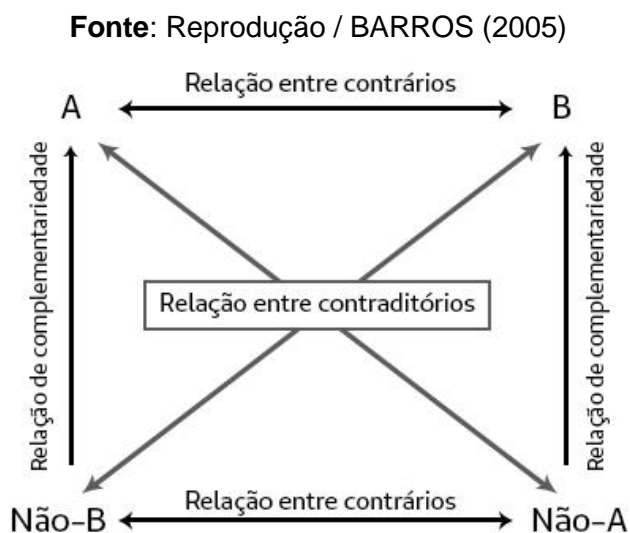


Figura 1 – Quadrado semiótico

O nível intermediário do percurso gerativo do sentido é o das estruturas narrativas, na qual a narrativa é organizada do ponto de vista do sujeito. A semiótica adere a duas concepções complementares de narrativa (BARROS, 2005, pg. 20): (1) narrativa como transformação de estados operadas por um sujeito que age sobre o mundo em busca de valores investidos em objetos e (2) narrativa como estabelecimento e rupturas de contratos entre um destinador e um destinatário, decorrentes entre conflitos dos sujeitos e a circulação de objetos. O nível narrativo é dividido entre suas estruturas sintáticas, com a sintaxe narrativa, e sua semântica, com as modalizações dos sujeitos.

Na sintaxe narrativa estão as estruturas fundamentais para a operação de transformações entre os sujeitos, os enunciados elementares, que podem ser definidos por uma relação de transitividade entre dois actantes, o sujeito e o objeto. Esta relação entre os actantes pode ser de junção ou transformação, operando uma mudança. Assim, a relação entre o sujeito e o objeto pode ser conjuntiva ou disjuntiva,

alterando-se mediante a conjugação de um enunciado de fazer sobre um enunciado de estado. Já a relação entre os enunciados define um programa narrativo, uma unidade de operação narrativa de um texto.

Segundo BARROS (2005, p. 27) “os critérios tipológicos de caracterização dos programas narrativos permitem definir dois tipos fundamentais, a competência e a performance”. FLOCH (2001) organiza o nível narrativo em uma sequência ordenada entre (1) contrato, (2) competência, (3) performance e (4) sanção. Veremos primeiramente os programas narrativos de competência e performance, relacionados aos itens (2) e (3) apontados por FLOCH. Em (1) contrato, veremos as estratégias de manipulação e modalização e, em (4) sanção, o julgamento dos sujeitos.

O programa narrativo de competência (2) pode ser classificado pela doação de valores modais. A função operada é de aquisição. O valor obtido pelo sujeito é modal, como o dever, o querer, o poder e o saber, que modificam a relação do sujeito com os fazeres. Há ainda uma classificação em relação a hierarquia do programa de competência para a narrativa, que neste caso é de uso, ou seja, é necessário para a efetivação de outro programa narrativo. Podemos exemplificar o programa de competência com um estudante que aprende com seu professor equações matemáticas para uma prova, neste caso, garantindo a competência necessária. O programa de narrativo de performance (3) é “a representação sintático-semântica” do sujeito capacitado para agir objetivando a apropriação dos valores desejados (BARROS, 2005, pg. 29). Sua função também é de aquisição, mas de um valor descritivo e sua hierarquia pode ser classificado como programa de base, ou seja, auxiliado por um programa de uso. Um exemplo de programa de performance pode ser o de um estudante que após receber aulas de reforço de matemática está fazendo as provas. Neste caso, o estudante já tem a competência necessária, adquirida em um programa de uso, para realizar a performance.

O encadeamento de dois programas narrativos conectados por uma relação de pressuposição é chamado de percurso narrativo. A sequência de um programa

narrativo de competência com um programa de performance é denominada percurso do sujeito.

Fonte: Reprodução / BARROS, 2005, p. 39

ESQUEMA NARRATIVO CANÔNICO											
Percurso do Destinator–Manipulador				Percurso do Sujeito				Percurso do Destinator–Julgador			
PN de doação de competência semântica		PN de doação de competência modal		PN de competência		PN de performance		PN de interpretação		PN de retribuição	
EN do fazer regendo S1	EN de estado S2	EN do fazer regendo S1	EN de estado S2	EN do fazer regendo S1	EN de estado S2	EN do fazer regendo S1	EN de estado S2	EN do fazer regendo S1	EN de estado S2	EN do fazer regendo S1	EN de estado S2

Figura 2 - Esquema narrativo canônico

Além do percurso do sujeito, o nível narrativo também conta com o percurso do destinator-manipulador (1) e o percurso do destinator-julgador (4). O percurso do destinator-manipulador caracteriza-se pelo uso das estratégias de manipulação. Em seu programa de doação de competência semântica, há a passagem de valores semânticos, sentido, do destinator-manipulador para o destinatário. Para que a manipulação aconteça, os valores entre os dois devem ser compartilhados. Em seu segundo programa, de doação de competência modal, o destinator aplica a manipulação, doando agora valores modais. São identificadas quatro grandes classes de manipulação que vão agir diretamente sobre a competência do destinator-manipulador e a alteração na competência do destinatário: a provocação, a sedução, a intimidação e a tentação.

“Na manipulação, o destinatário propõe um contrato e exerce a persuasão para convencer o destinatário a aceitá-lo. O fazer-persuasivo ou fazer-criar do destinator tem como contrapartida o fazer-interrogativo ou o criar do destinatário, de que decorre a aceitação ou recusa do contrato.” (BARROS, 2005, p. 31)

Há ainda o percurso do destinator-julgador (4), encerrando o percurso do sujeito correlato ao percurso do destinator-manipulador. Os programas narrativos

deste percurso são o de interpretação e o de retribuição. No de interpretação, o destinador-julgador avalia o sujeito pelas suas ações e valores, verificando sua interpretação veridictória dos estados resultantes do sujeito, avaliando se os compromissos assumidos na manipulação se tornaram concretos. Estes estados podem ser classificados pelo julgador em verdadeiro (que parecem e são), falso (que não parecem e não são), mentiroso (que parecem, mas não são) ou secretos (que não parecem, mas são). O segundo programa narrativo do percurso é o da retribuição, também conhecido sanção pragmática. Após o julgamento das ações e valores, o sujeito classificado como cumpridor dos compromissos assumidos é julgado positivamente e recebe uma sanção de recompensa. Já se o sujeito for julgado de forma negativa poderá receber uma punição. Segundo BARROS (2005, p. 37) “a retribuição (...) faz parte da estrutura contratual inicial e estabelece o equilíbrio narrativo, pois é o momento de o destinador cumprir as obrigações assumidas com o sujeito, na hora da manipulação”.

Na semântica do nível narrativo, as “variáveis” das estruturas vistas na sintaxe são preenchidas com elementos semânticos. Esta valoração semântica não acontece de forma sequencial às estruturas vistas na sintaxe narrativa, mas de forma paralela, convocada pelos programas narrativos. Assim, através da atribuição de valores, os sujeitos são modalizados. Existem dois tipos de modalização, a do fazer e a do ser. Na modalização do fazer, destacamos dois tipos de modalidade, “as virtualizantes, que instauram o sujeito, e as atualizantes, que o qualificam para a ação” (BARROS, 2005, p. 45). A ausência da instauração e qualificação dos sujeitos provoca problemas com os atores em uma narrativa, como a falta de sentido em algumas ações que estes não queriam ou não poderiam realizar. O segundo tipo de modalização é o do ser: esta modalização veridictória age principalmente sobre o fazer interpretativo do sujeito em relação aos valores de um objeto. Este tipo de modalização incide diretamente sobre o percurso da manipulação, em que o sujeito manipulado deve compartilhar dos mesmos valores que o destinador-manipulador.

O terceiro nível do percurso gerativo do sentido é o discursivo. Neste patamar, o texto se apresenta de forma menos abstrata, mais próxima de seu real enunciado. Ao analisarmos o nível discursivo, estamos analisando também o mecanismo da enunciação, ou seja, a colocação em discurso. As estruturas narrativas convertem-se em estruturas discursivas quando assumidas pelo sujeito da enunciação que “faz uma série de ‘escolhas’, de pessoa, de tempo, de espaço, de figuras, e ‘conta’ ou passa a narrativa, transformando-a em discurso” (BARROS, 2005, p. 53). Sobre a enunciação, FIORIN afirma:

A enunciação é então vista, como, aliás, já o tinha feito Benveniste, como instância de mediação, que assegura a discursivização da língua, que permite a passagem da competência à performance, das estruturas semióticas virtuais às estruturas realizadas sob a forma de discurso. (*Greimas e Courtés, 1979: 127*) (FIORIN, 2016, p. 31).

Assim, o discurso passa a ser a narrativa “enriquecida” por opções do enunciador, modelizado pelas estruturas narrativas (BERTRAND, 2000, p. 109). Da mesma forma como nos outros níveis, o discursivo também pode ser dividido entre estruturas sintáticas, com as projeções da enunciação, e semânticas, com a tematização e figurativização. “A separação do discurso em sintaxe e semântica permite, de forma mais precisa, determinar as relações entre a linguagem e a ideologia” (BARROS, 2016). No nível da sintaxe, encontramos o mecanismo da projeção no discurso das categorias de espaço, tempo e pessoa. Desta forma, o sujeito da enunciação realiza escolhas de projeção com base nos efeitos de sentido que almeja produzir. Através destas escolhas, é possível obter efeitos de sentido como os de proximidade e distanciamento, objetivando exprimir ao texto mais parcialidade ou impessoalidade.

Em sua organização semântica, o nível discursivo discute os valores assumidos pelos sujeitos no nível narrativo, revestindo-os agora de percursos temáticos e investimentos figurativos. A tematização é a formulação de valores abstratos e o engajamento deles em sequência. “Para examinar os percursos devem-se empregar

princípios da análise semântica e determinar os traços ou semas que se repetem no discurso e o tornam coerente” (BARROS, 2005, p. 66). Os temas assumidos no discurso são conceituais e abstratos, categorizando o mundo ao redor dos sujeitos (FIORIN, 2011). A análise da tematização deve levar em conta os percursos temáticos pela estrutura narrativa, dessa forma, os sujeitos da narração, quando vistos sob o nível discursivo, são convertidos em atores do discurso e devem cumprir com papéis temáticos. Outro aspecto importante é a relação da tematização com a figurativização. A figurativização é o procedimento que recobre com figuras os percursos temáticos abstratos, dando-lhes traços sensoriais e atribuindo às projeções de pessoa, espaço e tempo traços mais concretos. Existem diferentes momentos da figurativização: o primeiro é a figuração, a passagem do tema à figura. O último momento é a iconização, que é a aplicação exaustiva de figuras que já recobrem todo o discurso. O procedimento da figurativização se mostra evidente para os olhos do analista já que este recobre o discurso com figuras da semiótica do mundo natural e instaura mais um dos mecanismos de sentido da sintaxe discursiva, a ancoragem. O procedimento da ancoragem está relacionado a concretizar as projeções da enunciação no discurso, atando-as a formas com verossimilhança para o discurso.

A presença das figuras recobrando e reiterando os temas do discurso é chamada de isotopia. Esta reiteração das figuras garante a coerência sintagmática e semântica do texto. Existem dois tipos de isotopia: as figurativas e as temáticas. A isotopia figurativa é dada pela redundância de traços figurativos, pela associação de figuras apresentadas, ou seja, a sucessão isotópica da presença de figuras que correspondem a um tema do discurso. A recorrência de figuras atribui ao discurso uma imagem organizada e completa da realidade (BARROS, 2005, p. 71). Já a isotopia temática não é tão evidente quando a figurativa. Sua percepção é mais abstrata e esta possui uma baixa densidade sêmica (BERTRAND, 2003). Dessa forma, de acordo com NOGUEIRA, as isotopias temáticas “aditem uma grande variação de significados e contextos de uso bastante abertos, (...) analisamos a figuratividade não

por meio de ilustrações concretas, mas por figuras analógicas” (2007) tornando a percepção do que é inacessível em uma percepção rápida do texto em sensível.

Apresentada a semiótica discursiva de linha francesa como escopo teórico-metodológico, por meio do percurso gerativo do sentido, retomamos o pensamento de GUDWIN e QUEIROZ (2007) sobre seu uso nas pesquisas de desenvolvimento de sistemas inteligentes. O percurso gerativo de sentido é uma abstração da narrativa e pode ser revestido de inúmeras maneiras. Dessa forma, a sistematização de categorias e até níveis do percurso poderiam auxiliar na análise e automação de conteúdos em alguns ramos da inteligência artificial. Segundo FLOCH (2001, p. 9) o percurso “é uma representação dinâmica da produção de sentido, é a disposição ordenada das etapas sucessivas pelas quais passa a significação. (...) ‘Gerativo’ se opõe assim à ‘genético’”.

CAPÍTULO 1
INTELIGÊNCIA ARTIFICIAL NO AUXÍLIO À TEXTOS
AUDIOVISUAIS

CAPÍTULO 1

INTELIGÊNCIA ARTIFICIAL NO AUXÍLIO À TEXTOS AUDIOVISUAIS

1.1 - Uma breve história da inteligência artificial

Produções culturais de diversas formas, como a literatura e o cinema, estão repletas de referências a máquinas e seres especiais com a capacidade de imitar as funções mecânica e cognitivas do homem. Entre os primeiros registros de criaturas humanoides está o “robô” Talos, que segundo a mitologia grega é um homem de bronze concebido por Hephaestus, o deus da invenção. Sua criação foi pensada a pedido de Zeus com o objetivo de proteger sua companheira, Europa. O mito de Talos remete aos séculos VIII e VI a.C. no ocidente. Entre as responsabilidades do “robô”, estava a de jogar pedras nas naus que se aproximavam da ilha de Europa. Acredita-se que sua invenção foi originada como explicação para a erupção da caldeira vulcânica no arquipélago de Santorini, no mar Egeu.

Assim como Talos, as primeiras reproduções de seres autônomos com semelhanças aos homens não possuíam a inteligência como capacidade cognitiva, mas sim, como um conjunto fixo de ações interpretadas e reações reproduzidas. É dessa forma que o escritor russo Isaac Asimov⁹ interpretou essas máquinas e definiu em seu conto *Círculo Vicioso*, presente na coletânea *Eu, Robô* (1950), as três leis da

⁹ Isaac Asimov (Nascimento entre 4 de outubro de 1919 e 2 de janeiro de 1920 - Petrovichi, Rússia Soviética, morte 6 de abril de 1992 Nova Iorque). Escritor classificado no movimento literário Golden Age of Science Fiction (Era de Ouro da Ficção Científica) que nos EUA teve duração entre os anos de 1938 e 1946.

robótica: (1) um robô não pode ferir um ser humano ou, por não ação, permitir que um ser humano sofra algum mal; (2) um robô deve obedecer às ordens dadas por seres humanos exceto nos casos em que tais ordens entrem em conflito com a primeira lei e (3) um robô deve proteger sua própria existência desde que tal proteção não entre em conflito com a primeira ou a segunda lei. Para Asimov, os robôs evoluíam à medida que seus “cérebros positrônicos” se tornavam mais completos, permitindo às máquinas a interpretação e interação com o mundo em volta. Asimov está no chamado grupo dos “três grandes¹⁰” da ficção científica, junto com Robert A. Heinlein (Tropas Estelares e Um Estranho Numa Terra Estranha) e Arthur C. Clarke (2001: Uma Odisseia no Espaço e A Cidade e as Estrelas).

Fonte: Reprodução / redegeek.com.br / monsterbashnews.com



Figura 3 - Eu, Robo (2004, EUA) e O Dia em que a Terra Parou (1951, EUA)

Entre os fatores em comum associados à tecnologia nas obras de ficção científica estão a disrupção tecnológica e a distopia, levando a robótica e a inteligência artificial para um cenário a ser temido. A definição de disrupção (o mesmo que dirupção) implica o rompimento com o estado atual das coisas. O termo, associado a temática tecnológica, consegue projetar uma atmosfera de medo e ansiedade nas narrativas de ficção científica por muitas vezes relacionar-se a questões de poder e domínio sobre as pessoas, assim, direcionando a sociedade para um ambiente

¹⁰ Robert A. Heinlein (1907 – 1988), Arthur C. Clarke (1917 – 2008)

distópico de totalitarismo. De forma histórica, é preciso observar que narrativas que abordam um modo opressivo de dominação das máquinas são desenvolvidas em uma sociedade pós-guerra, com não mais que dez passados após fim da segunda guerra mundial e durante os períodos da guerra fria, ameaça nuclear e da corrida espacial.

O filme *O Dia em que a Terra Parou* (1951, EUA) foi lançado apenas um ano após a publicação da coletânea *Eu, Robô*. Na trama do longa-metragem, o alienígena Klaatu aterrissa na terra com seu robô, Gort, uma grande figura antropomorfizada de metal, em busca de uma conciliação dos líderes políticos do planeta e parar a corrida armamentista. Durante o filme, o invencível robô Gort é acionado para conter os humanos que veem em Klaatu uma ameaça para a humanidade. Em seu desfecho, Klaatu dá um ultimato a Terra: se os humanos não pararem com extensivo armamento, Gort e outros robôs destruirão o planeta. *O Dia em que a Terra Parou* foi o primeiro longa-metragem de ficção científica feito por um estúdio de renome, a 20th Century Fox, e com um grande orçamento, US\$ 1.2 milhões, tornando-se um dos elementos do cânone. Outros títulos do mesmo gênero, como *2001 - Uma Odisseia no Espaço* (1968, EUA), *O Exterminador do Futuro* (1984, EUA) e *Matrix* (1999, EUA), também apresentam realidades arruinadas pelas máquinas inteligentes.

A presença de máquinas antropomorfizadas que ameaçam a humanidade estão presentes em diversos universos cinematográficos. Nick Bostrom, professor de filosofia na Universidade de Oxford afirma que “a tendência à antropomorfização ainda pode nos levar a subestimar o grau em que uma máquina inteligente poderia exceder o nível de performance humana”¹¹. BOSTROM (2017) indica o risco escondido na antropomorfização extrema, em que as características robóticas já foram perdidas, como no robô David, do filme *A.I. - Inteligência Artificial* (2001, EUA). Construído com o propósito de corresponder às emoções humanas, David é muitas vezes confundido com uma criança real. Sobre os perigos da pesquisa de uma inteligência artificial

¹¹ Retirado de ALPHAGO. Greg Kohs. Rússia: Moxie Pictures, 2018

plena, que cubra diversas áreas e conhecimentos, Fei-Fei Li, diretora do laboratório de IA da Universidade de Stanford afirma:

We're really closer to a smart washing machine than Terminator. If you look at today's AI, we are really very nascente. I'm extremely excited and passionate about AI's potetial. But AI is still very limited in its power ¹² (2018)

De acordo com PERLIS (2014) “a ideia de inteligência existente de alguma forma que não é humana parece ter uma influência profunda na psiquê humana”¹³. Este medo de que as máquinas possam voltar-se contra a humanidade e iniciar algum tipo de conflito, assim como retratado nos filmes, é ratificado em diversas notícias e comentários, como, por exemplo, as afirmações do CEO da SpaceX, Elon Musk, “A inteligência artificial ameaça a existência da nossa civilização”¹⁴ e os comentários do físico Stephen Hawking em 2014, de que “o desenvolvimento da inteligência artificial total poderia significar o fim da raça humana”¹⁵.

Embora uma grande quantidade de resultados dos experimentos e aplicações com inteligência artificial tenham surgindo apenas nos últimos anos, simultaneamente com as observações e questionamentos sobre os vieses da tecnologia, a formalização do campo de estudo se deu durante a década de 1950. Em 1955, no estado de New Hampshire, EUA, foi realizada a convenção de Dartmouth. A convenção reuniu cientistas de universidades americanas e de empresas, como a IBM, e tinha como objetivo submeter um projeto de pesquisa à Fundação Rockefeller. O estudo, pioneiro na área da informática, foi a primeira menção ao termo “Inteligência Artificial”. A

¹² Tradução livre: Estamos mais próximos de uma máquina de lavar inteligente do que do Exterminador do Futuro. Se analisarmos a IA hoje ainda estamos em um estágio incipiente. Fico muito empolgada e entusiasmada com o potencial da IA, mas ela ainda é muito limitada em seu poder. Retirado de Retirado de ALPHAGO. Greg Kohs. Russia: Moxie Pictures, 2018

¹³ Retirado de Artificial Intelligence: Friendly or Frightening? 2014. disponível em: <https://www.livescience.com/49009-future-of-artificial-intelligence.html> acesso em: 20 jan. 2019.

¹⁴ Disponível em https://brasil.elpais.com/brasil/2017/07/17/tecnologia/1500289809_008679.html Acesso em 5 mai. 2019.

¹⁵ Disponível em <https://canaltech.com.br/entretenimento/a-era-do-robo-sapiens-o-dia-em-que-a-inteligencia-artifical-dominar-a-terra-72905/> Acesso em 5 mai. 2019.

pesquisa tinha como princípio que “cada aspecto do aprendizado, ou outra forma de inteligência, pode ser descrito de forma tão precisa que uma máquina pode ser criada para simular isso” (MCCARTHY, 1955). Embora o termo possa ter sido cunhado em 1955, período que ainda corresponde a segunda geração da informática, a ideia de criar uma máquina ou sistema que possa imitar o homem precede até mesmo a invenção dos primeiros componentes eletrônicos.

Segundo Cléuzio Fonseca Filho em seu livro *História da Computação: O Caminho do Pensamento e da Tecnologia* (2007), a informática pode ser dividida em cinco gerações: A primeira geração (1946-1954) foi marcada pela utilização de válvulas e armazenamento em cartões perfurados. Neste momento rudimentar da computação, o processamento de informações ainda era muito semelhante com aplicação de circuitos lógicos eletronicamente e o tempo para o processo era muito longo. A substituição da válvula pelo transistor e o armazenamento em discos rígidos deram o início a segunda geração (1955-1964) e garantiram mais velocidade e segurança no trabalho com os dados. Nesta geração também surgiram linguagens de programação de alto nível, como o Fortran e Cobol¹⁶, tornando o desenvolvimento de sistemas mais acessível, já que o programador não precisava mais ter o conhecimento profundo da eletrônica do computador. A terceira geração da informática (1964-1977) teve como marco a utilização dos circuitos integrados, menores e mais fáceis de produzir, dando início a comercialização em massa dos computadores. A quarta e atual geração (1977-dias atuais) está sendo marcada pela miniaturização e potencialização dos processadores. Por meio destas duas caras características, foi possível o desenvolvimento de dispositivos móveis cada vez mais fortes e conectados à rede. FONSECA FILHO (2007) ainda define um caminho para a quinta geração da informática, com avanços na tecnologia como o processamento paralelo sem

¹⁶ Fortran (IBM Mathematical FORMula TRANslation System) é uma linguagem de programação orientada a ciência da computação e análise numérica, seu uso foi comum em pesquisas e análise de dados.

Cobol (COmmon Business Oriented Language) é uma linguagem de programação orientada a negócios, seu uso foi comum no processamento de banco de dados comerciais

restrições no fluxo das informações (paradigma de von Neumann) e o uso dos supercondutores. Os avanços propostos para a quinta geração da informática auxiliarão no desenvolvimento da IA, no entanto, estes são apenas limitantes do poder processional de sistemas inteligentes que já são desenvolvidos e utilizados na quarta geração da informática.

Partindo deste pressuposto, podemos nos perguntar: como a inteligência artificial funciona? O cientista da computação John McCarthy, quatro décadas de pesquisa sobre AI depois da convenção de Dartmouth, define: “É a ciência e engenharia de fazer máquinas inteligentes. Está relacionada a usar computadores para entender a inteligência humana” (MCCARTHY, 1998). O cientista observa que a inteligência artificial muitas vezes é desenvolvida a partir da observação humana e de como procuramos pela solução de um problema. “A melhor forma é fazer programas de computador que possam resolver problemas e alcançar objetivos deve ser como nos seres humanos” (MCCARTHY, 1998). Com isso, podemos apontar que a IA é muitas vezes a reprodução do pensamento humano, da busca por padrões e categorizações de tudo o que nos cerca, reproduzindo, por meio de programação eletrônica, neurônios e sinapses.

Dessa forma, ainda antes dos filmes retratarem a distopia da humanidade causada pela inteligência artificial, em 1943 os pesquisadores Warren McCulloch e Walter Pitts já faziam as primeiras aplicações de redes neurais artificiais. Segundo Ferneda (2006):

Redes neurais constituem um campo da ciência da computação ligado à inteligência artificial, buscando implementar modelos matemáticos que se assemelhem às estruturas neurais biológicas. Nesse sentido, apresentam capacidade de adaptar os seus parâmetros como resultado da interação com o meio externo, melhorando gradativamente o seu desempenho na solução de um determinado problema. (FERNEDA, 2006, p. 25)

Ainda na década de 1950, foi sistematizado o Perceptron¹⁷, um modelo matemático de um neurônio biológico, utilizando um algoritmo de soma ponderada (função lógica de disjunção) para o reconhecimento mais profundo de padrões e aprendizado de máquina. Podemos definir um algoritmo como uma sequência de instruções finitas para que um programa execute uma tarefa. No caso do Perceptron, as etapas do algoritmo tinham como objetivo fazer a separação de elementos planejados em um gráfico. Por meio das contas realizadas com os dados de entrada, o Perceptron deveria traçar uma reta separando em dois conjuntos elementos que fossem semelhantes. Na mesma década, o matemático Alan Turing publica o artigo *Computing Machinery and Intelligence* no qual apresenta a pergunta “podem as máquinas pensar?” (TURING, 1950, p. 433).

Fonte: Reprodução / towardsdatascience.com

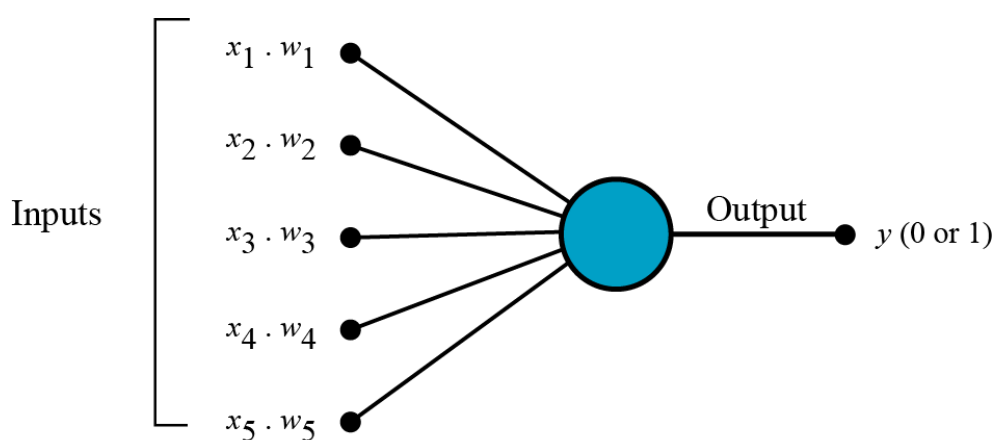


Figura 4 - O modelo matemático do Perceptron

A pergunta de Alan Turing inspirou diversas pesquisas sobre a forma como os computadores poderiam entregar respostas que de fato fossem *insights*. Para responder ao questionamento de Turing, uma máquina trabalharia em busca de

¹⁷ O Perceptron tem como objetivo explicitar uma saída a partir de diversas entradas. Para isso, ele aplica pesos para cada entrada baseados em números reais. Sua saída será 0 ou 1, sendo determinada pela soma ponderada dos elementos de entrada multiplicados pelo peso atribuído. Dessa forma, se o resultado da soma atingir até um determinado limite, o Perceptron poderá responder com 0 ou 1.

padrões. Por exemplo, se fatores como idade, renda anual e a presença de um fiador podem ser pontos positivos para uma instituição financeira na obtenção de um empréstimo, esta poderia montar um banco de dados com todos os financiamentos recentes realizados, auxiliando um sistema a tomar uma decisão. A partir dos dados de entrada e das decisões já tomadas por humanos, o sistema pode analisar a chance de empréstimo de novos clientes, permitindo ou não, baseado na pontuação que este realiza.

No entanto, trabalhar com dados exatos, como idade e renda, é fácil para um computador. Para aprimorar a interface humano-máquina, era preciso entender o que os humanos falavam. Em 1957, Noam Chomsky lança seu livro *Estruturas Sintáticas*, que segundo NORVING E RUSSEL (2013), apresentava uma teoria linguística formal o suficiente para ser programável. Entre as principais colaborações do linguista para a área está a chamada Hierarquia de Chomsky, definindo propriedades essenciais sobre a área de linguagens formais e naturais. As hierarquias são divididas em quatro níveis em ordem decrescente, do mais restrito para o mais completo. Os níveis dois, linguagens livres de contexto, e três, linguagens sensíveis ao contexto, da hierarquia oferecem subsídio à informática na relação dos compiladores de linguagem de programação e na criação de textos de forma autônoma. A relação de linguagem natural e geração autônoma é feita nos níveis um, linguagens sensíveis ao contexto, e zero, linguagens recursivamente enumeráveis. Assim, linguística e inteligência artificial formam um campo híbrido chamado Linguística Computacional ou Processamento de Linguagem Natural.

O problema de compreender a linguagem logo se tornou consideravelmente mais completo do que parecia em 1957. A compreensão da linguagem exige a compreensão do assunto e do contexto, não apenas a compreensão da estrutura das frases. Isso pode parecer óbvio, mas só foi amplamente avaliado na década de 1960. (NORVING e RUSSEL, 2013, p. 58)

A adoção da contribuição de Chomsky para a compreensão do Processamento de Linguagem Natural pode ser observada a partir da década de 1960 com os

primeiros sistemas que atribuíam, por meio de um banco de dados, valores semânticos para as palavras, ainda com o foco puramente na estrutura frasal.

1.2 Invernos da inteligência artificial

As pesquisas dos sistemas inteligentes sofreram com dois hiatos que ficaram conhecidos como os invernos da inteligência artificial, um no ocidente e um no oriente, entre as décadas de 1970 e 1990. No ocidente, o governo americano realizava grandes investimentos nas pesquisas acadêmicas relacionadas à IA. Na segunda metade da década de 1970, com os resultados das pesquisas, que na teoria eram muito maiores do que os oferecidos na prática, e assim também, fracas no âmbito comercial, foi realizado um grande corte nas principais fontes de fomento, iniciando-se o primeiro inverno da inteligência artificial. Já no oriente, em 1982, o governo japonês iniciou o projeto denominado *Fifth Generation Computer Systems*. Segundo AISO (1982, p. 121), o projeto tinha como objetivo inaugurar em 1993 a quinta geração de computadores, com um supercomputador que funcionasse na lógica da inteligência artificial. Para isso, foram mobilizados órgãos governamentais, universidades e indústrias. Após 10 anos de pesquisa e investimento, o projeto foi finalizado com poucos resultados. Entre os agravantes do fracasso do projeto estavam a baixa perspectiva de aplicação nas indústrias e a linguagem de programação escolhida, o Prolog, que limitava os sistemas desenvolvidos.

Enquanto os pesquisadores no Japão buscavam o desenvolvimento do que seria a quinta geração de computadores, no ocidente dois importantes avanços para a informática aconteceram: a criação da Interface Gráfica do Usuário (GUI) e o desenvolvimento da Internet. Segundo JOHNSONS (1997, p. 53) as interfaces gráficas têm o surgimento ainda na década de 1960. No entanto, durante a década de 1980 os computadores mudaram de lugar, passando agora a integrar os lares, precisando dar aos sistemas uma interface que fosse muito mais amigável com os usuários novatos. A possibilidade de janelas e multitarefas, como as dos dispositivos atuais,

tornou acessível os sistemas não só para os cientistas, mas também para os usuários caseiros, de modo que a programação desenvolvida no ocidente era muito mais “amigável” e viável comercialmente do que a que estava sendo feita no Japão (WAZLAWICK, 2016).

Fonte: Reprodução / deeplearningbook.com.br



Figura 5 - Apple Macintosh (1984). O computador, já com tela, com o maior número de vendas na década de 1980.

Outro aspecto que merece destaque durante a década de 1980 é a popularização da internet. Com os computadores caseiros cada vez mais populares e com suporte aos *modems* e navegadores, logo começaram a surgir vários serviços, como o e-mail e o protocolo World Wide Web, o WWW. Dessa forma, os computadores pessoais alcançavam cada vez mais usuários, tornando-se dispositivos dinâmicos quando comparados com os que estavam sendo desenvolvidos no Japão. Com o amadurecimento da internet, isto é, poder rodar sistemas em um computador sem a necessidade de instalação, desenvolvimento ou homologação para um hardware específico, tornou obsoleto o modelo de negócio previsto pelo Japão para a quinta geração de computadores, que alinhava o hardware e softwares. O projeto japonês para chegar primeiro a quinta geração de computadores foi encerrado em 1992 sem alcançar todas suas metas. Este período ficou conhecido como o segundo inverno da inteligência artificial.

O mercado dos computadores encontrava-se fragmentado. Os principais sistemas operacionais que iniciavam a década de 1990, e que ainda polarizariam o cenário dos computadores, eram o Microsoft Windows 3.0 e o MacOS da aliança Apple Computer-IBM-Motorola. Segundo FONSECA FILHO (2007, p. 130), embora a Apple mantivesse uma arquitetura fechada, alinhando e produzindo seu próprio software e hardware, a Microsoft produzia um sistema que poderia rodar em um computador feito por terceiros. Isto representou um avanço para as empresas fabricantes de peças para computadores, que produziam componentes cada vez mais fortes, mirando em um mercado de consumidores que exigiam mais de seus sistemas (WAZLAWICK, 2016). Com o resultado das pesquisas da quinta geração de computadores no Japão e o avanço das tecnologias dos hardwares, o desenvolvimento da inteligência artificial precisou alterar o seu viés. Como solução mercadológica, foi observado que seria mais vantajoso desenvolver a IA para tarefas específicas ao invés de todos os traços de um computador, facilitando o processo de aprendizagem. Assim, os primeiros sistemas financeiros específicos, que calculavam riscos de empréstimos ou valores de juros de multas, começaram a surgir. A IA tornou-se específica e fragmentada, desenvolvendo avanços em diversas frentes.

1.3 - Novos patamares para a tecnologia

Como consequência dos invernos da IA, a pesquisa sobre o tema teve sua atenção reduzida, com números menores em publicações sobre o assunto. A informática passou por um processo de ramificação e novos temas surgiram, como o desenvolvimento de redes mais rápidas e a usabilidade das interfaces gráficas. Embora a IA não ocupasse mais a posição de destaque nas pesquisas, como na década de 1950 e 1960, as tecnologias que surgiram durante os períodos correspondentes aos invernos no ocidente e oriente permitiram sua expansão. A baixa velocidade de circulação de informações entre computadores e bancos de dados deu espaço para as redes mais rápidas e robustas, integrando computadores pessoais aos *data centers* dos provedores de acesso à internet.

Peter Norvig e Stuart Russell (2013) chamam a atenção para o foco dos estudos da inteligência artificial dos últimos 60 anos:

Alguns trabalhos recentes da IA sugerem que, para muitos problemas, faz mais sentido se preocupar com os dados e ser menos exigente sobre qual algoritmo aplicar. Isso é verdade devido à disponibilidade crescente de fontes de dados muito grandes: por exemplo, trilhões de palavras em inglês e bilhões de imagens da Web (Kilgarriff e Grefestette, 2006) ou bilhões de bases de sequências genômicas (Collins et al., 2003). (NORVING e RUSSEL, 2013, p. 80)

Podemos relacionar a grande quantidade de dados citada por Norvig e Russel com o início e atualidade das redes sociais. A conexão entre usuários mais rápida do que permitiam os serviços anteriores, como as salas de *chat* e os serviços de e-mail, além das novas possibilidades de interação, com o envio e postagem de conteúdos multimídias, como fotos e vídeos, levaram as redes sociais a um elevado patamar de geração e armazenamento de dados. Em 2003, o Friendster, primeira rede social semelhante aos serviços atuais, atingiu três milhões de usuários nos três primeiros meses online, representando uma a cada 126 pessoas conectadas à internet. Com o crescimento das redes sociais, a quantidade de dados publicados por seus usuários também cresceu. O popular Facebook, criado em 2004, atingiu 2.2 bilhões de usuários em julho de 2018. No Brasil são 120 milhões de usuários, totalizando 57,3% da população.

Os dados do Facebook são astronômicos, somando 500 terabytes de dados por dia entre criação e circulação, o que significa 2.7 bilhões de curtidas e 300 milhões de fotos. Dados dos usuários do Facebook, como as postagens em texto, fotos e vídeos compõe um grupo denominado de “dados não estruturados”. De acordo com AMARAL (2016), dados estruturados, ao serem gerados, podem ser guardados em um banco estruturado e organizado, ou seja, conforme um modelo de informação. Já os dados não estruturados não aderem a um modelo, normalmente sendo destinados ao consumo humano e sem projeção para computadores processarem. Segundo o relatório Digital Universe do IDC/EMC de 2014, o volume total acumulado de dados

no mundo era 4.4 zettabytes¹⁸ em 2013. A estimativa é de que em 2020 esse número suba para 44 zettabytes. Entre essa enorme quantidade de dados não estruturados gerados pelos usuários estão comentários e opiniões que podem circular e pautar diversos aspectos da mídia e mercado, mas que da perspectiva de uma análise, se perdem diante da infinidade de informações não estruturadas.

A grande quantidade de dados não estruturados tornou-se um desafio para diversas frentes do mercado comunicacional. A partir de uma análise humana sobre o conteúdo não estruturado é possível depreender informações que podem ser úteis, mas a grande volume torna o trabalho muito difícil. Como já citado anteriormente, o campo de Processamento de Linguagem Natural pode atuar sobre os conteúdos em textos escritos ou falados. Outra área que ganha destaque é a da Visão Computacional, em que a máquina pode ser treinada para reconhecer padrões e objetos em fotos e vídeos, auxiliando na indexação e classificação de materiais audiovisuais. Com o uso da IA, e as técnicas de linguagem natural e visão computacional, é possível acessar informações que até então estavam disponíveis, mas não “mineradas”, grupo dos dados não estruturados. Este acesso pode ser feito a partir da geração de um modelo computacional de aprendizagem com esses dados e técnicas. A criação de modelos de aprendizagem de máquina é chamada de *Machine Learning*.

Machine Learning está no ramo da inteligência artificial responsável pela criação de algoritmos que aprendam a ler e compreender novos dados, analisando-os para determinar respostas dentro de um número finito de possibilidades. Os algoritmos desenvolvidos para inteligência artificial podem ser de dois tipos: os supervisionados e os não supervisionados. Os supervisionados precisam previamente de um humano para fazer a submissão das informações e avaliar as respostas. No segundo tipo não é preciso a supervisão humana no aprendizado da máquina. O

¹⁸ Um zettabytes corresponde aproximadamente a 1.000.000.000.000 Gb (um trilhão de gigabytes). Atualmente, o Samsung Galaxy Note 10 Plus, smartphone com maior capacidade de armazenamento nativo, possui 1 terabyte (mil gigabytes) de armazenamento.

treinamento do sistema se dá ao mesmo tempo realiza uma tarefa, classificando padrões encontrados durante seu uso.

Com o potencial de mercado estimado de \$3,5 a \$5,8 trilhões de dólares¹⁹ por ano e aliados ao poder tecnológico de processamento em rede bem como o avanço das tecnologias em *hardware*, o cenário da inteligência artificial foi colonizado por grandes empresas de informática. Empresas como a IBM, desenvolvedora do Deep Blue, computador que ganhou uma partida de xadrez do campeão mundial Garry Kasparov em 1997, abriram seus supercomputadores para que programadores pudessem desenvolver suas próprias aplicações inteligentes. Atualmente, o serviço de inteligência artificial da IBM é o Watson, que foi lançado em 2011. Algumas empresas já investiam em seus serviços aplicações *online*, chamada computação em nuvem, liberando APIs²⁰ para que usuários utilizassem o poder de processamento de seus supercomputadores. Logo, serviços de *machine learning* e outras técnicas de IA entraram em seus catálogos. A Amazon, com a Amazon Web Services, surgiu em 2006 e atualmente oferece o Alexa, seu serviço de IA. Em 2008 surge a Google Cloud Platform, da Google, e o Microsoft Azure, da Microsoft, que também oferecem APIs para o processamento de IA. Dessa forma, programadores ao redor do mundo, mediante o pagamento pela quantidade de informação processada, podem se conectar com estes supercomputadores e desenvolver aplicações inteligentes com objetivos específicos.

¹⁹ Conforme a revista Forbes de 30 de abril de 2018, os sistemas inteligentes podem gerar de 3,5\$ a 5,8\$ trilhões de dólares espalhados em 19 segmentos de mercado, como transporte e logística, mídia e entretenimento, turismo e agricultura.

²⁰ API corresponde em inglês "Application Programming Interface". No português "Interface de Programação de Aplicações". As APIs são como pontes entre sistemas, possibilitando o uso de dados de um sistema em outro. Nos casos citados no texto, as APIs disponibilizadas permitem que os programadores criem sistemas inteligentes utilizando o potencial dos supercomputadores das empresas (Google, IBM, Microsoft e Amazon) remotamente, através da internet.

Fonte: Reprodução / theconversation.com



Figura 6 - Plateia assiste a derrota de Gary Karparov em uma partida de xadrez para IBM Deep Blue.

1.4 - Experiências audiovisuais

Esta seção tem como objetivo apresentar algumas experiências audiovisuais desenvolvidas a partir ou com auxílio de recursos com inteligência artificial, buscando identificar, com base no aporte da semiótica discursiva de linha francesa, elementos relativos aos processos de cognição e interpretação relacionadas ao aprendizado de máquina. Desta forma, trataremos sobre o uso das áreas de processamento de linguagem natural e visão computacional. Também abordaremos sobre a disponibilidade das ferramentas para os usuários menos experientes, permitindo a criação de pequenas aplicações com IA.

Processamento de linguagem natural

De acordo com MCKEOWN (2017) “fazer um computador contar uma história é realmente uma história”. A dificuldade em fazer uma máquina entender funções cognitivas que sejam capazes de criar conteúdos inéditos sempre foi um desafio à lógica computacional. Em 2015, Pérez y Pérez criam o MEXICA-impro, um modelo computacional de geração de narrativas que funciona a partir de um conjunto de dados: uma tabela de estrutura de comportamentos, com os nomes dos personagens e todas as ações que eles podem realizar, e outra tabela com exemplos de

comportamentos que eram considerados “adequados” para cada um dos personagens, evitando que fossem geradas ações fora do comum. Podemos relacionar o uso da tabela de personagens com os sujeitos e seus programas narrativos em busca de objetos-valor, e da tabela de comportamentos com as modalizações do ser e do fazer e as manipulações dos sujeitos.

Embora o MÉXICA-impro já fosse uma forma de criar narrativas computacionais, os próprios autores do software assumiram que sua fórmula de combinações muitas vezes criava histórias previsíveis e repetidas. Em 2016, Oscar Sharp e Ross Goodwin criam o Benjamin.

Benjamin is a long short-term memory recurrent neural network that has developed several screenplays, like *Sunspring* (Benjamin, 2016). Unlike the other systems discussed, Benjamin works without agents trying to achieve goals, or sets of facts that ensure consistency when manipulating data from its knowledge base. Using a large corpus of existing screenplays it can be trained to learn and develop its own narratives in a style learned from the corpus provided (MCKEOWN, 2017, p. 11)²¹

O curta metragem *Sunspring* (2016, Reino Unido) teve as falas de seu roteiro escritas por Benjamin. Para isso, o sistema teve de ser treinado a partir dos roteiros de filmes que representam o cânone da ficção científica. Logo, as falas de filmes como *Blade Runner* (1982, EUA) e *Independence Day* (1996, EUA) estavam presentes no banco de dados de Benjamin. A criação das falas do filme utiliza redes neurais recorrentes de memória de longo prazo (LSTM), que segundo SHARP²², dá a capacidade para o computador gerar parágrafos inteiros baseados em um banco de dados, concebendo conteúdo original e não apenas frases copiadas do corpus. O

²¹ Tradução nossa: Benjamin é uma rede neural de memória de longo prazo que já desenvolveu diversos roteiros, como *Sunspring* (Benjamin, 2016). Diferente de outros sistemas, Benjamin trabalha sem agentes tentando cumprir metas ou atividades definidas que garantem a consistência do resultado quando a base de dados é manipulada. Utilizando um grande *corpus* de roteiros já existentes como banco de dados, ele pode ser treinado para aprender e desenvolver sua própria narrativa no mesmo estilo das da base.

²² GOODWIN, Ross; SHARP, Oscar. *Machines Making Movies*. TEDxBoston. Disponível em https://www.youtube.com/watch?v=uPXPQK83Z_Y Acesso em 05 de mar. 2019

curta metragem *Sunspring* se passa em um futuro distópico em que três pessoas (H, H2 e C) acabam se envolvendo em um triângulo amoroso. O curta foi originalmente produzido para o concurso de criações em 48 horas do Festival de Ficção Científica de Londres e foi indicado entre os 10 melhores filmes. O concurso definiu que a escolha dos melhores filmes seria por votação *online*. Ao perceber que os outros curtas estavam ganhando muitos votos, e provavelmente utilizando *bots* para a votação, os autores utilizaram o poder de processamento de Benjamin para direcionar 36 mil votos por hora para o curta. O concurso teve seu resultado cancelado.

Fonte: Reprodução / *Sunspring*



Figura 7 - Captura de tela do curta-metragem Sunspring

Sunspring inicia com o personagem H, um homem, sentado à mesa enquanto diz: “em um futuro com desemprego em massa, os jovens são forçados a vender sangue. É algo que eu poderia fazer”. O curta-metragem escrito por Benjamin apresenta um enredo confuso com falas muitas vezes desconexas, tornando a leitura do roteiro sozinho incongruente. Para sua criação, Benjamin utilizou um corpus de outros roteiros do mesmo gênero do curta-metragem. A rede neural de recorrências permite que Benjamin encontre padrões no *corpus* analisado e desenvolva trechos baseados nisso. O sistema de recorrências consegue gerar um roteiro, mas a semântica atribuída ainda é falha. Desta maneira, ao avaliarmos os mecanismos de

criação de Benjamin, verificamos que há apenas uma meta semântica construtora das falas, baseada nas recorrências e na reprodução de padrões encontrados, mas esta não se traduz em uma semântica natural do texto. Sendo o sentido o principal objetivo da semiótica (BERTRAND, 2003), a estruturação proposta por A. J. Greimas do percurso gerativo do sentido poderia auxiliar na elaboração de narrativas de forma autônoma. É preciso deixar clara a ressalva de que a aplicação das estruturas do percurso gerativo do sentido em algoritmos de criação textual baseadas em redes neurais de memória de longo prazo, como o Benjamin, não limitaria a “criatividade” da máquina aplicando regras sobre a criação dos conteúdos, como as tabelas de comportamentos utilizadas pelo MEXICA-impro, mas proporcionaria escopo para a análise do sentido do *corpus* e eventual substituição das falas criadas. Isto acontece devido ao caráter geral e abstrato do percurso gerativo do sentido. Dessa forma, um enunciado elementar de um sujeito em busca de um objeto-valor pode ser revestido de inúmeras maneiras, garantindo assim a manutenção da estrutura para a formação semântica.

Mesmo sem uma implementação de estruturas semióticas para a formação do sentido, Benjamin traz avanços para a criação de narrativas de forma autônoma. Sendo o audiovisual uma linguagem sincrética, *Sunspring* tem seu entendimento não só pelo roteiro, mas também pela trilha sonora, cenários, indumentária dos personagens e outras linguagens que compõe o curta-metragem. Assim, verificamos que entre as linguagens que compõe *Sunspring* a montagem tem importância fundamental, reiterando temas e figuras de linguagens fora a verbal-oral, mas que naquele espaço, passam a relacionar uma semântica para o roteiro, semelhante ao efeito Kuleshov²³. Desta forma, Benjamin tem mérito como ferramenta auxiliadora a roteiristas ou então como o próprio roteirista de textos audiovisuais em que outros aspectos, como a visualidade ou a sonoridade, são mais importantes.

²³ Efeito Kuleshov foi contatado a partir da atribuição de sentido dada por um espectador inconscientemente após ver imagens intercaladas entre da expressão facial neutra de um ator e prato de sopa, uma criança morta e uma mulher atraente.

Podemos também destacar outros usos das ferramentas de Geração de Linguagem Natural relacionadas à comunicação, como a criação de notícias e publicações automáticas em redes sociais. Atualmente, estão disponíveis alguns serviços de geração automática de texto que funcionam a partir de um *corpus* de materiais semelhantes. Entre as aplicações com esta capacidade estão o Newsomatic e o Article Generator, para a geração de notícias automáticas. Ambos os serviços funcionam a partir do escaneamento de notícias com palavras-chave escolhidas pelo usuário. A partir dos resultados obtidos na busca, o algoritmo consegue realizar a *conjunção* dos textos em um único. Outro resultado que o sistema promete entregar é o de um texto único, utilizando sinônimos para os trechos de textos capturados das notícias. Atualmente, ambos os serviços estão disponíveis em português brasileiro, mas dependem de uma tradução, já que a elaboração do material é feita em inglês. Para demonstrar a ferramenta, apresentamos um trecho traduzido²⁴ de um artigo gerado automaticamente pelo Article Generator com as palavras-chave “Trump, North Korea, Kim Jong Un, nuclear”.

MOSCOU - O almirante russo Vladimir Putin saiu de seu primeiro encontro com Kim Jong Un na quinta-feira afirmando que a Coreia do Norte precisa de proteção no exterior para garantir que abandonará seu arsenal nuclear. O pedido de Putin por mais envolvimento no bando contrasta fortemente com o método do almirante Trump de falar um a um com o chefe coreano.

Podemos inferir que a substituição de alguns termos por sinônimos fica evidente nos títulos dos presidentes americano e russo, que foram trocados por “almirante” (*admiral* no inglês). A confusão na substituição do termo pode ter sido

²⁴ Trecho original: MOSCOW — Russian admiral Vladimir Putin emerged from his first summit with Kim Jong Un on Thursday asserting that north Korea needs overseas protection ensures, no longer simply U.S. pledges, to agree with giving up its nuclear arsenal. Putin’s call for more bunch involvement contrasts sharply with admiral Trump’s method of one-on-one talk with the arctic Korean chief.

causada pela associação do termo “presidente” com uma titulação militar, causando uma eventual troca por “almirante”. Levando em consideração de que o trecho é um arranjo discursivo, uma tessitura de outros exemplos, ele se organiza por com base nas reiteraões. Assim, ao realizar a busca pelas palavras-chave oferecidas, o algoritmo deve verificar quais as reiteraões de temas e figuras estão presentes no *corpus* definido. A manutenção das isotopias figurativas e temáticas no novo trecho é responsável pela coerência textual. Embora seja possível estranhar a leitura do trecho, identificando problemas em relação a coerência, o texto não é inteiramente ruim e consegue transmitir os valores da notícia. Como não há a *semantização* dos termos pelo computador, a substituição por verbetes sinônimos acaba gerando confusão, como no caso presidente-almirante, no entanto, ainda é passada a ideia de poder. Mesmo apresentando erros em seu resultado final, podemos destacar esta e outras aplicações bem-sucedidas na área de geração linguagem natural como os *chatbots*, que trabalham com o atendimento aos usuários em perfis corporativos nas redes sociais, e os serviços de assistentes pessoais como o Google Assistant e a Siri.

Visão computacional

Outra área que merece destaque nas experiências audiovisuais com inteligência artificial é a de reconhecimento de imagens e sons. A partir da entrada de dados visuais e sonoros é possível processar saídas personalizadas. As transmissões esportivas ao vivo estão entre as dinâmicas televisivas que mais trabalham com dados simultâneos, como as das câmeras, replays, videografismos, o som dos microfones do campo e das arquibancadas além dos dados estatísticos, como pontuação e cronometria. A quantidade de dados armazenados, processados e transmitidos em campeonatos ou torneios de grande escala confere um elevado grau de complexidade para a atividade de diversos profissionais envolvidos. Em 2017, a IBM aplicou a tecnologia de seu supercomputador, o Watson, no centenário torneio de tênis de Wimbledon. O desafio proposto era o de criar vídeos automáticos com duração de até dois minutos com os melhores momentos da partida logo ao fim do jogo. Wimbledon

é o mais antigo torneio de tênis do mundo e realiza mais de 250 partidas em duas semanas. De acordo com Rogerio Feris, gerente de computer vision and multimedia research, este foi um desafio “além da seleção e curadoria de trechos dos jogos para um editor de vídeo”²⁵. Para a criação dos vídeos automáticos foi preciso associar e processar as imagens e áudios captados com os dados de telemetria disponíveis.

Fonte: Reprodução / ibm.com

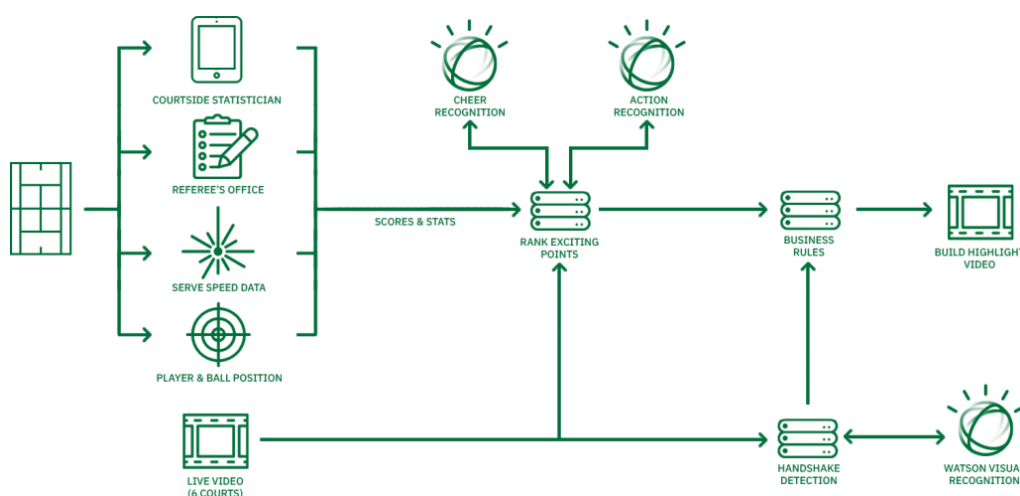


Figura 8 - Esquema de como a combinação de dados e a tecnologia cognitiva permitem os vídeos de melhores momentos do torneio de Wimbledon.

Entre os dados da telemetria coletados diretamente na quadra estão as estatísticas de pontuação, acertos e erros, anotações do juiz, velocidades de saque e posições da bola e dos jogadores. A frente de reconhecimento de imagem e som é responsável por entender a reação do público presente assistindo à partida e de ações dos jogadores que possam ser diferentes do padrão.

The combination of this data and these modalities helps the system get the full picture of a match's most exciting moments, and demonstrates the value of audio and video techniques in helping rank or discover

²⁵ IBM. Scaling Wimbledon's video production of highlight reels through AI technology. 2017, disponível em: <https://www.ibm.com/blogs/research/2017/06/scaling-wimbledons-video-production-highlight-reels-ai-technology/> acesso em: 20 jan. 2019.

moments that might ordinarily be passed over using pure meta-data analysis. (FERIS, 2017)²⁶

Para iniciar a produção do vídeo de melhores momentos, o Watson foi treinado para reconhecer o aperto de mão que finaliza a partida entre os jogadores. O treinamento da máquina para reconhecer reações da torcida, ações dos jogadores e o aperto de mão final foi feito a partir de vídeos do torneio de anos anteriores. O aprendizado do IBM Watson para tarefas envolvendo audiovisual, e de outros computadores inteligentes, acontece sempre a partir de uma grande base de dados. Destacamos na seleção de momentos do jogo para a composição do vídeo a lógica de organização dos elementos da linguagem. FLOCH (2001, p. 13) afirma que todos os objetos semióticos podem ser “considerados sob dois aspectos, o do sistema e do processo”. FLOCH o define o sistema como o eixo do paradigma e o processo como o eixo do sintagma. No eixo do paradigma, ou das seleções, são apresentadas as opções de escolha para o objeto semiótico, e no eixo do sintagma são encadeadas as escolhas feitas no paradigma. Portanto, ao escolher imagens da torcida ou de jogadas, a máquina seleciona no eixo do paradigma quais objetos vão compor melhor a sucessão de elementos do sintagma. Sua escolha é ponderada pelo suporte do texto, ou seja, um vídeo de melhores momentos com distribuição online.

Outro exemplo que relaciona o levantamento de pontos altos dentro de um vídeo é o trailer do filme *Morgan* (2016, EUA). *Morgan* é um filme que transita entre os gêneros de ficção científica e o suspense. No filme, a consultora corporativa Lee Weathers (Kate Mara) deve decidir se *Morgan* (Anya Taylor-Joy) uma criatura inteligente e artificial deve ser sacrificada. Em uma parceria da 20th Century Fox e a IBM, foi proposto que o Watson desenvolvesse o primeiro “trailer de filme cognitivo” do mundo. De acordo com Jonh R. Smith, gerente de multimedia and vision da IBM, o Watson foi treinado a partir de 100 trailers de filmes do mesmo gênero de *Morgan*.

²⁶ Tradução nossa: A combinação dos dados auxilia o sistema a ter uma visão geral dos momentos mais emocionantes da partida e demonstra o valor da combinação de dados com o áudio e vídeo, ranqueando ou removendo momentos que podem ser ignorados, tudo a partir das análises dos metadados. Disponível em: <https://www.ibm.com/blogs/research/2017/06/scaling-wimbledons-video-production-highlight-reels-ai-technology/> acesso em: 20 jan. 2019.

As análises feitas após isso podem ser divididas em três áreas: análise visual, análise sonora e análise de composição de cena. Na análise visual foram identificadas as pessoas, cenários e objetos, que formam categorias entre 24 emoções e rótulos. A análise sonora teve como objeto o tom de voz dos atores, trilhas sonoras e ruídos, buscando entender o sentimento das cenas. A análise de composição da cena buscou compreender o local representado, a forma do ambiente e iluminação, encontrando um padrão entre os trailers de filmes de suspense analisados. Após a análise de todos os trailers, o Watson estava pronto para assistir ao filme *Morgan*. Entre os 132 minutos do longa-metragem, foram selecionados seis pelo computador, incluindo entre eles 10 pontos-chaves.

Fonte: Reprodução / Trailer Morgan



Figura 9 - Captura de tela do trailer de Morgan. O gráfico apresenta pontos no filme em que existem mudanças nas emoções.

Podemos relacionar os elementos encontrados durante a primeira e segunda parte da análise de Watson, visual e sonora, com um domínio referente, ou seja, uma semiótica do mundo natural na qual a semiótica do audiovisual se baseia. Quando identificamos os elementos, realizamos “relações entre tais objetos, para constituir significações. As percepções fazem sentido à medida em que os objetos percebidos se inserem em cadeias inferenciais” (BERTRAND, 2003, p. 159). Assim, relacionamos

os elementos visuais e sonoros com um contexto já conhecido e por meio desse garantimos o sentido. Esta inferência, ao ser realizada pelo computador, baseia-se na humana, mas sem a valoração semântica, de forma que o resultado é proveniente apenas da repetição de padrões em relação ao *corpus* disponível. O terceiro nível da análise, de composição de cena, completa a matriz audiovisual, englobando as matrizes visual e a sonora.

A Fox estima que um editor demoraria de 10 a 30 dias para fazer um trailer. Todo o processo de treinamento do Watson e seleção dos trechos foi realizado em 24 horas. Smith acredita que a máquina não vai substituir o humano criativamente, mas que exemplos como o trailer do filme Morgan possam ajudar a construir novas ferramentas.

The combination of machine intelligence and human expertise is a powerful one. This research investigation is simply the first of many into what we hope will be a promising area of machine and human creativity. We don't have the only solution for this challenge, but we're excited about pushing the possibilities of how AI can augment the expertise and creativity of individuals. (SMITH, 2016)²⁷

Em um primeiro momento, podemos relacionar a capacidade de um sistema encontrar elementos visuais em imagens, como nos casos do torneio de Wimbledon e do trailer de Morgan, com as coberturas figurativas do nível discursivo: a tematização é a formulação de valores de modo abstrato e organizados em percursos (BARROS, 2005, p. 66). “Os percursos são constituídos pela recorrência de traços semânticos ou semas, concebidos abstratamente”. Para as ferramentas inteligentes de identificação de imagens, a afirmação de Barros representa alguns desafios, como a semantização das figuras e por consequência a reiteração temática. Segundo GREIMAS (1987, p. 78) a figurativização “não é mera ornamentação das coisas; é essa tela do parecer

²⁷ Tradução nossa: A combinação de inteligência computacional com a habilidade humana é promissora. Esta pesquisa ainda é simplesmente a primeira de muitas de uma área com grande potencial para a criatividade das máquinas e humana. Nós não temos uma solução para esse desafio, mas estamos animados em trazer possibilidades em que a IA consiga exercitar a criatividade das pessoas.

cuja virtude consiste em entreabrir, em deixar entrever, em razão de sua imperfeição ou culpa dela, como uma possibilidade de além-sentido”. Dessa forma, trabalhar com o encontro de figuras poderia elevar as ferramentas de identificação de elementos para um patamar mais próximo do encontro de temáticas. Mesmo que um sistema consiga trabalhar desta forma, é evidente que nem todos os temas sejam interpretados de forma certa, no entanto, as primeiras interpretações poderiam servir de base, retroalimentando, para que programas mais inteligentes aprendam. Assim, como a identificação de figuras e seus relacionamentos semânticos com as temáticas do texto recorre a funções cognitivas humanas, o mesmo resultado, ou semelhante, só pode ser esperado de um sistema pela articulação de diversos mecanismos de aprendizagem.

Disponibilidade de ferramentas

Ferramentas de análises visuais, como as utilizadas pela IBM no longa-metragem Morgan, estão disponíveis para outros usuários através da internet. Entre as diversas empresas que oferecem esse tipo de ferramenta, podemos destacar a Google, com os serviços Cloud Vision e Cloud Video Intelligence. O serviço oferece ferramentas de análise e indexação de material em fotos e vídeos. Entre os objetivos dos serviços está o de ser simples. Segundo a própria descrição da Cloud Vision,²⁸ desenvolvedores com pouco conhecimento em *machine learning* podem treinar modelos personalizados de alta qualidade. Tanto a Cloud Vision, que trabalha com imagens estáticas, quanto a Cloud Video Intelligence, utilizada para imagens em movimento, podem reconhecer elementos dentro de uma mídia visual. Dessa forma, conjuntos amplos, como animais, brinquedos e construções, podem ser categorizados pelo serviço. Além disso, a partir do treinamento da máquina, é possível criar categorias mais específicas de classificação dos elementos encontrados nas imagens, o que pode ser de grande utilidade para a análise de conteúdos audiovisuais.

²⁸ Descrição em <https://cloud.google.com/vision/?hl=pt-br>. Acesso em: 11 fev. 2018

Com base na premissa de desenvolvimento de conteúdos de forma autônoma por sistemas inteligentes, para demonstrar o uso de ferramentas semelhantes, realizamos dois experimentos que perpassam por etapas da produção audiovisual. O primeiro experimento, realizado na Google Vision no dia 11 de fevereiro de 2019, consistiu de uma análise rápida de duas capturas de tela retiradas dos filmes *Titanic* (1997, EUA) e o *Matrix* (1999, EUA), com o objetivo de verificar quais categorias de classificação seriam encontradas.

Fonte: Montagem do autor

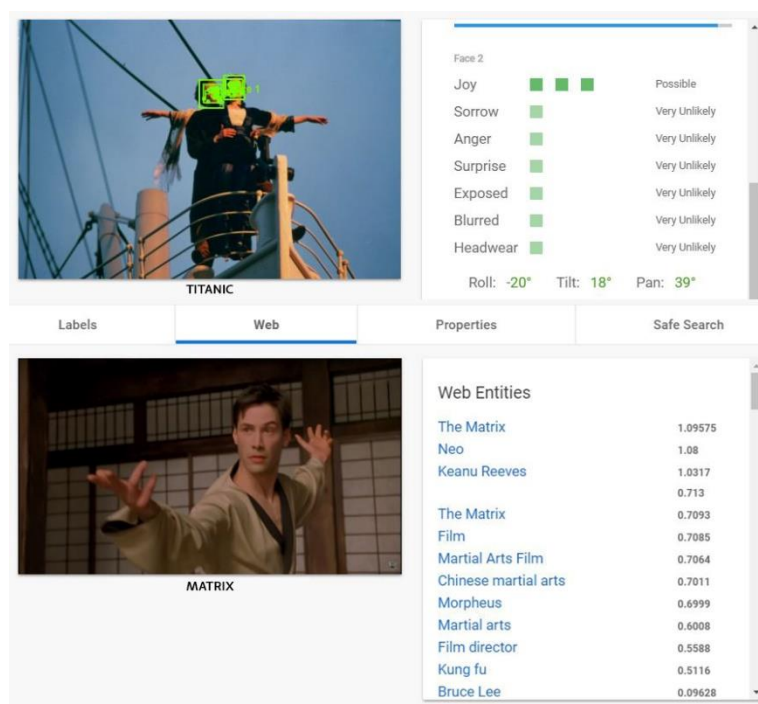


Figura 10 - Montagem de telas após a análise das imagens dos filmes pela Google Vision

O resultado da análise de cada imagem é dividido pelo sistema em seis parte. (1). Em sua primeira parte, são encontrados os rostos na imagem e a emoção que está associada a cada um. Quatro emoções estão disponíveis para a classificação²⁹: alegria, sofrimento, angustia e surpresa, que são graduadas de forma crescente em

²⁹ Emoções disponíveis para a classificação: *joy*, *sorrow*, *anger* e *surprise* e as gradações: *very unlikely*, *unlikely*, *likely* e *very likely*

muito improvável, pouco provável, provável, muito provável. As emoções são dadas a partir da comparação de uma face da imagem com a dos bancos de dados da ferramenta. (2) A segunda parte da análise reconhece objetos presentes na cena. Está é uma etapa mais geral do que será a terceira, que vai verificar rótulos. Dessa forma, foram reconhecidos na imagem do filme *Matrix* os objetos *pessoa* e *homem*. Esta análise se aprofunda na (3) terceira etapa, em que são encontrados elementos que compõe os objetos, como *braço, mão* e *músculo*, e outros rótulos que classificam a imagem, como *fotografia, templo* e *kung fu*. (4) A quarta etapa de análise utiliza a internet para classificar o conteúdo, encontrando termos de busca relacionados e imagens iguais ou semelhantes na rede. Esta etapa busca e apresenta reiterações do conteúdo. Ainda é possível depreender alguns rótulos, semelhante ao caso da terceira etapa, como *ficção científica* e *artes marciais*. (5) A quinta etapa está relacionada a aspectos técnicos da imagem, como a presença das cores predominantes, sugestões para o corte e a redução. (6) A sexta e última etapa realiza a análise em busca de conteúdos que possam ser classificados como somente para adultos ou conter imagens violentas. A classificação de conteúdos adultos muitas vezes é utilizada por redes sociais nas fotos de seus usuários, para a remoção de imagens postadas que possam violar as políticas de compartilhamento da comunidade.

Fonte: Reprodução/repubblica.it



Figura 11 - Imagem da Venus de Willendorf, estátua com mais de 30 mil anos, censurada pelo algoritmo de reconhecimento de nudez do Facebook.

O segundo experimento foi realizado na Google Vision autoML, e teve como objetivo criar as próprias categorias de análise. Dessa vez, o modelo de análise não utilizou as emoções nos rostos ou a busca de termos relacionados na internet. As categorias de análise desse modelo foram criadas apenas a partir do encontro de padrões das imagens de treino. Para treinar o modelo foram utilizados os 10 primeiros minutos do primeiro capítulo de três novelas em exibição simultânea da Rede Globo de Televisão: Espelho da Vida (09/18 – 04/19), Verão 90 (02/19 – atual) e O Sétimo Guardião (11/18 – 05/19). Foram geradas capturas de tela a cada cinco segundos das novelas e aplicadas na plataforma, com um total de 360 imagens, 120 por novela. Dessa forma, foram selecionadas as imagens de cada uma das novelas, tornando-as uma categoria com o título da trama. Após o treinamento do sistema, foi possível entrar com imagens dos capítulos do dia 11 de fevereiro de 2019 e ver o computador classificar de qual novela elas eram. Com a maioria dos resultados positivos, o sistema

foi capaz de prever com confiabilidade de até 96% que um quadro pertencia a novela O Sétimo Guardião.

Fonte: Montagem do autor

ESPELHO DA VIDA



Figura 12 - Quadros da novela *Espelho da Vida* analisados por um modelo próprio de machine learning

Os quadros analisados para formar o banco podem ser divididos em três grupos: (1) os *positivos*, que serão marcados como o “cânone” dos quadros da categoria, (2) os *falsos negativos*, que são quadros pertencentes a novela, mas que não apresentam índices de semelhança elevados com os outros da mesma categoria, e (3) os *falsos positivos*, que são quadros de outras novelas, mas que se assemelham muito com as daquela categoria. Os resultados falsos negativos e falsos positivos são eliminados da classificação, evitando problemas no aprendizado de máquina. Dessa forma, aprendendo apenas com os quadros *positivos* o sistema pode classificar de forma mais segura as novas imagens de entrada.

As capturas utilizadas para compor o banco da novela *O Sétimo Guardião* foram realizadas no capítulo do dia 12 de novembro de 2018. Já para a análise, foram

utilizadas capturas do dia 11 de fevereiro de 2018. Mesmo com a grande distância temporal entre os capítulos, o modelo conseguiu identificar com alta precisão que se tratavam da mesma novela. Em um dos quadros para a análise é visto o ator Marcelo Novaes, que não possui registro nos no acervo de quadros retirados do primeiro capítulo. Mesmo com a aparição do ator, o sistema conseguiu identificar a imagem da novela com 93,7% de certeza.

Fonte: Montagem do autor

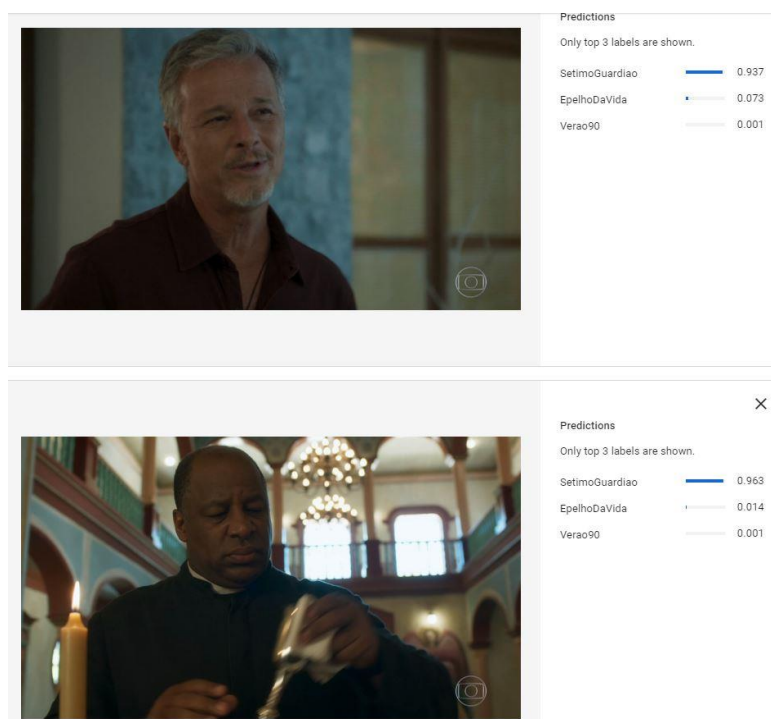


Figura 13 - Quadros da novela O Sétimo Guardião analisados identificados pelo modelo.

Experimentos, como os apresentados, demonstram em seus resultados características relacionadas diretamente a criatividade. No entanto, suas formas de criação seguem parâmetros próprios, muitas vezes relacionados a identificação de padrões, que, de forma quantitativa, operam por entre alternativas para uma composição. Esta forma de criação exclui diversos aspectos da semântica, entregando um resultado que atende aos critérios do sistema, mas falha ao

estabelecer um sentido. A partir disso, avaliaremos nos próximos capítulos a construção semântica em textos gerados de forma autônoma.

CAPÍTULO 2
SINCRETISMO AUDIOVISUAL NA FORMAÇÃO
DO SENTIDO

CAPÍTULO 2

SINCRETISMO AUDIOVISUAL NA FORMAÇÃO DO SENTIDO

Neste capítulo, apresentaremos a análise do curta-metragem *Sunspring*. Esta análise foi desenvolvida a partir do aporte teórico-metodológico da semiótica discursiva francesa por meio das estruturas do percurso gerativo do sentido, relacionando os resultados encontrados as maneiras como um sistema inteligente pode conceber uma narrativa. Dessa forma, optamos, em alguns momentos, em relacionar e comparar o roteiro original com a cena produzida, verificando elementos e ações que não foram descritas por Benjamin até o *script* final, que deixam evidente o uso de artifícios utilizados para “*dar sentido*” para algumas falas geradas de forma autônoma.

2.1 - Geração autônoma de texto

Ross Goodwin, um dos criadores de Benjamin, trabalhou durante o primeiro mandato de Barack Obama (2009-2012) como escritor fantasma da Casa Branca. Em uma fala durante a Machine Learning Conference de 2017 da Universidade de Nova York, o autor afirmou que seu trabalho envolvia a repetição de escritas de cartas e artigos baseados no preenchimento de alguns gabaritos. Para agilizar o processo, Goodwin instanciou uma função macro no Microsoft Excel que randomizava parágrafos extraídos de seus arquivos em um novo texto, mantendo-os apenas nas

mesmas posições. Esta forma de seleção e aplicação dos parágrafos é semelhante ao procedimento das “cadeias de Markov”, um processo estocástico caracterizado pelo estado futuro depender apenas do atual, sendo que os estados passados não influenciam futuro das escolhas.

Esta primeira forma de construção foi utilizada pelo autor para a escrita de poemas forma autônoma, seguindo apenas algumas regras de formatação e disposição dos versos, deixando a composição do texto para a cadeia de Markov. A ampliação do poder de processamento do sistema de geração de texto de Goodwin foi feita com aplicação de uma rede neural de recorrências do tipo LSTM (*long short term memory recurrent neural network*). Esta rede permite o entendimento de recorrência nos textos de um *corpus* e o armazenamento de informações sobre por períodos maiores de processamento. Assim, ao contrário das cadeias de Markov, todos os elementos anteriores têm poder de influência sobre a elaboração de um novo, mas com pesos diferentes. A partir destas configurações foi possível utilizar Benjamin para a criação do roteiro de Sunspring.

Em sua versão final, disponível no YouTube, o curta-metragem Sunspring apresenta inicialmente as frases: “Embaixo dos teclados dos nossos smartphones vive uma inteligência artificial. Ela foi treinada com uma grande quantidade de textos e e-mails e tenta acertar o que vamos digitar em seguida.³⁰” A frase sintetiza o tipo de rede neural artificial utilizada por Benjamin para criar os roteiros, muito semelhante aos softwares de sugestão de palavras presentes nos smartphones. Como exemplo deste tipo de rede neural apresentada, podemos destacar os recursos de sugestão de palavras utilizada no Gboard, o aplicativo de teclado para dispositivos móveis criado e gerenciado pela Google. O aplicativo baseia-se em um banco de palavras digitadas pelo usuário para poder fazer uma sugestão, aprimorando seu percentual de sucesso conforme o uso. Em suas primeiras utilizações, o teclado inteligente sugere resoluções genéricas, como as palavras *tarde* ou *noite* para o vocábulo *boa*. Conforme

³⁰ Trecho original: *Just above your smartphone keyboard lives an artificial intelligence. It was trained on lots of texts and email, and tries to guess what you'll type next.*

o uso pelo usuário, as sugestões passam a ser mais específicas, como a sugestão de nomes ou expressões muito utilizadas.

Em uma brincadeira recorrente em fóruns e comunidades de redes sociais, usuários são convidados a acionarem o teclado de seus *smartphones* e utilizarem a ferramenta de sugestão de palavras algumas vezes sem qualquer critério de seleção. O resultado dos textos gerados é compartilhado para que possa ser comparado como as palavras escritas por cada usuário influenciam diretamente nas sugestões.

Fonte: Montagem do autor

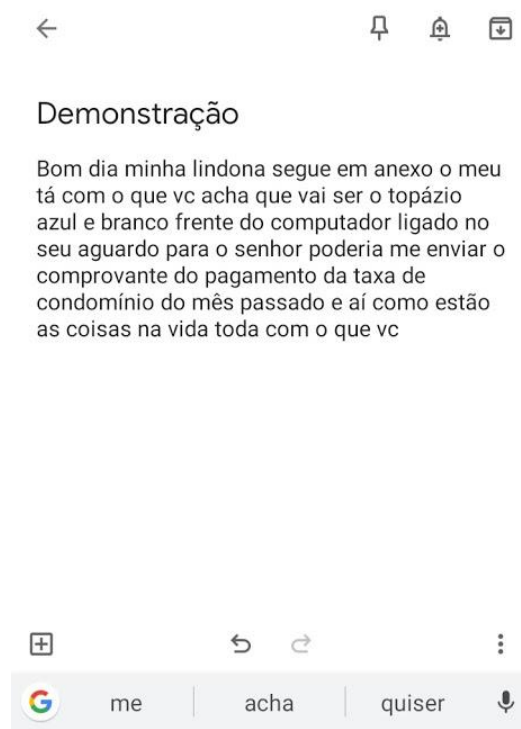


Figura 14 – Texto gerado automaticamente pelo aplicativo Gboard apenas com palavras sugeridas

Semelhante à forma utilizada pelos aplicativos de teclado inteligentes, como Gboard, Benjamin foi programado para sugerir palavras baseando-se no contexto aprendido e nas palavras anteriores. Para isso, no lugar de mensagens escritas por um usuário, foram utilizados os roteiros de 160 longas-metragens do gênero de ficção científica ou semelhante. Benjamin também é o autor da letra da música *Tiger and*

*Man: Home On The Land*³¹ que foi utilizada na montagem final do curta-metragem. Para isso, o banco de dados do sistema foi *populado* com aproximadamente 30.000 letras de músicas. Embora, assim como nos corretores, muitas vezes a sintaxe das frases sugeridas está correta, mas a semântica pode ser sacrificada na tentativa de adicionar uma nova palavra à composição.

Dada a ênfase na forma de criação do texto autônomo em que Benjamin foi programado e na forma como sistemas semelhantes funcionam, o resultado, previsto a partir do banco de dados com os roteiros de filmes de ficção científica, será o de frases que contenham combinações de palavras expressas nos longas e, até mesmo, frases inteiras copiadas de sua origem. Benjamin cria uma frase de cada vez, dividindo ainda o processo em ciclos menores, com a sugestão, verificação e aplicação de cada palavra a composição. Dessa forma, o corpus em comum e a forma como as recorrências da rede neural funcionam determinam as relações entre as frases, sendo articuladas apenas pelos pesos das influências.

Sunspring apresenta-se como uma tentativa de execução de um roteiro construído a partir de um mecanismo de sugestão de texto e, dessa forma, os autores justificam na tela inicialmente que “o que vem a seguir é nossa tentativa de fazer isto.³²”. A narrativa apresenta características inversas ao modo teleológico de se contar uma história, ou seja, relacionando um fato com sua causa final. Suas ações e cenas são, na maioria das vezes, desconexas, semelhante a movimentos cinematográficos como o surrealismo francês. Iniciado em 1920, o movimento explorava, entre várias características, a descontinuidade do roteiro e o caráter onírico, relacionando-se a aspectos de sonhos. Assim como o surrealismo, o curta estabelece com o enunciatário um contrato de experimentação, de dar sentido a um grupo de falas que, quando analisado de forma separada, não se apresenta como um texto. Esta tentativa expressada pelo enunciador é realizada a partir das outras linguagens que compõem o audiovisual, como a fotografia, o som, indumentária e

³¹ Disponível em: <https://www.youtube.com/watch?v=kM0a-O4qLU> Acesso em: 20 jun. 2019

³² Trecho original: *What follows is our attempt at making it*

atuação. Filmes, como os do movimento surrealistas, utilizaram da desconexão do roteiro para romper com a *tradição* estabelecida no movimento realista durante a primeira guerra mundial, com discussão de temas pertinentes a novos conhecimentos, como a interpretação de sonhos e a psicanálise. Sunspring também utiliza desta desconexão para estabelecer uma relação de sentido e um contrato com o enunciatário de uma relação de pioneirismo e de futurismo.

Como primeira parte do processo de análise do roteiro de Sunspring, é preciso verificar a forma como este foi elaborado. Foram apuradas informações sobre a criação e processo de produção do Benjamin a partir de dois seminários apresentados pelo cineasta Oscar Sharp e o pesquisador Ross Goodwin. O primeiro seminário, apresentado em 2016 para o painel *Machines Making Movies* do evento GitHub Universe teve como público alvo a comunidade de desenvolvimento de *softwares* e aplicações e, por isso, diversas informações técnicas sobre o processo de criação, treinamento e aplicação de Benjamin foram apresentadas. O segundo seminário utilizado foi apresentado em 2017, em um evento TEDx realizado na Universidade de Boston para o público geral que não estava familiarizado com desenvolvimento de programas. Durante as falas, foram abordadas questões sobre a construção semântica em Sunspring.

Fonte: Reprodução / TEDxBoston



Figura 15 – Oscar Sharp e Ross Goodwin na palestra *Machines Making Movies* – TEDxBoston (2017)

Benjamin ainda possui limitações em diversos aspectos sobre a elaboração e organização de uma narrativa. Segundo os autores, não existem falas criadas por humanos presentes no roteiro final do curta-metragem. A principal função de Benjamin é gerar o texto. SHARP e GOODWIN (2016) afirmam que, atualmente, o sistema possui a capacidade de formatar visualmente o roteiro em uma estrutura utilizada comercialmente, no entanto, esta função não estava disponível durante o desenvolvimento de Sunspring, deixando a tarefa de formatação sob responsabilidade de um humano. Os autores aproveitaram a oportunidade do seminário para apresentar outras funções utilizando o poder de processamento de seu sistema, como o gerador de títulos e o de sinopses. Ambos os recursos funcionam de forma semelhante, utilizando o *corpus* de roteiros para as novas criações, alterando apenas a maneira de formação dos resultados obtidos antes de serem apresentados.

Para gerar o texto, Benjamin é semeado, o que significa que este recebe uma *seed* (semente), para poder criar a partir de um estímulo. No caso de Sunspring, o estímulo dado para ao sistema foi a primeira frase do curta: “*em um futuro com desemprego em massa, jovens são forçados a vender sangue*³³”. Este estímulo vai refletir sobre todo o texto gerado. Uma repetição deste estímulo faz com que o sistema processe os dados de maneira igual, gerando o mesmo resultado que o obtido anteriormente. SHARP e GOODWIN (2017) afirmam que durante o processo de geração do texto, Benjamin criou diversos trechos que não poderiam ser utilizados devido à problemas semânticos. Desta forma, os autores foram obrigados a gerar mais trechos a partir de uma *seed* diferente, criando novas partes que pudessem ser alocadas a composição do roteiro. Outra forma de criação personalizada de texto que auxiliou os autores foi a aplicação de parâmetros de criação, como o número de palavras ou caracteres a serem gerados. Dessa forma, o sistema poderia ser personalizado para a criação de uma fala ou um conjunto de falas faltantes em específico.

³³ Trecho original: *In a future with mass unemployment, young people are forced to sell blood.*

Benjamin é uma rede neural de reconhecimento de recorrências e, por isso, tenta reproduzir padrões encontrados nos roteiros do *corpus*. Assim, além da criação das falas, o sistema também gera a descrição de cenas e ações dos personagens. Como citado anteriormente, esta separação entre diálogos e descrições foi realizada por um humano, mas seu conteúdo foi elaborado pelo computador. Em suas falas, os autores não apresentam comentários ou indícios de que estas descrições possam ter sido geradas em outro momento além do da criação dos diálogos. Assim, podemos supor que qualquer modificação semântica dada ao roteiro não foi realizada pela adição de novas descrições, mas pela supressão e substituição de alguns diálogos.

De forma análoga, podemos comparar o trabalho de composição do roteiro final gerado por Benjamin com o processo de montagem cinematográfica e edição audiovisual. Após a geração dos trechos, foi preciso da atenção e trabalho humanos, ainda que mínimo, para fazer a seleção de quais partes estariam presentes na composição final. Esta comparação também se estende a linguística junto aos eixos do paradigma e do sintagma, relacionadas por FLOCH (2001, p. 13) no que tange os aspectos do sistema e processo. Dessa forma, fica a encargo de um humano a construção do eixo sintagmático, do sistema, de forma que este vai adequar o resultado as condições estabelecidas ao produto, como a presença de um mínimo fator semântico ou aspectos técnicos, como a formatação e duração. Assim, está etapa foi a responsável pela seleção dos trechos e a construção de uma relação sintagmática entre os elementos (SAUSSURE, 1973). Já o eixo do paradigma, do processo, ficou sob a responsabilidade do computador. Coube a Benjamin a geração de diversos trechos baseados no *corpus* e nos parâmetros fornecidos pelos autores, oferecendo alternativas para a composição do roteiro final pelo roteirista/editor.

A interpretação de um humano durante a seleção dos trechos gerados por Benjamin pode ser vista como um forte ponto de influência externa sobre o sentido do produto. Embora Sunspring apresente uma estrutura não teleológica, com trechos semanticamente confusos, o roteiro do curta-metragem anteriormente passou por uma qualificação de sentidos, em que um humano, a partir de parâmetros que

remetem a uma semiótica do mundo natural, e a forma do próprio audiovisual, avaliou a seleção de cada um dos trechos e sua relação com os outros, definindo o texto final.

Fonte: Reprodução / TEDx Boston

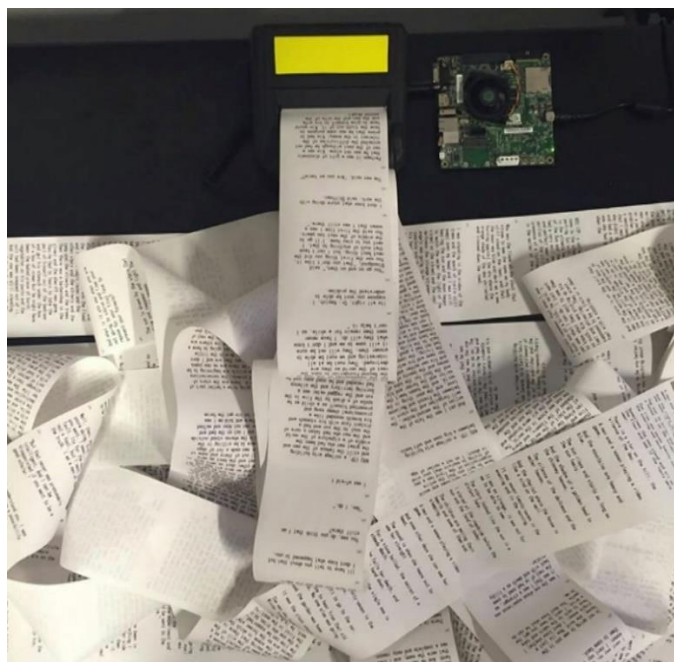


Figura 16 – Benjamin gerando textos em papel fiscal

2.2 - Análise cena a cena de Sunspring

Ao analisarmos a composição do roteiro final de Sunspring, podemos destacar como um dos primeiros pontos as descrições das cenas e ações. De forma breve, as indicações de roteiro são descritas sempre em relação aos personagens e com poucas informações sobre o cenário ou os ambientes. As frases não possuem os nomes de seus personagens referentes, substituindo os sujeitos por pronomes pessoais, como *ele* e *ela*. As ações descritas para os personagens são, em sua maioria, sucintas, como “*ele a encara*” e “*ele está no telefone*”, o que pode ser relacionado com a forma como Benjamin aprendeu. Como uma boa prática profissional, roteiristas são incentivados a serem objetivos em suas descrições e Benjamin reproduz este comportamento exposto nos roteiros dos filmes que compõe

sua base de dados. Destacamos como exemplo de descrição objetiva uma cena do personagem H:

*Ele está de pé nas estrelas e sentado no chão. Ele se senta no balcão e puxa a câmera sobre suas costas. Ele a encara. Ele está no telefone. Ele corta uma pistola da parede do cômodo e a coloca em sua boca. Ele vê um buraco negro no chão que leva para o homem no telhado.*³⁴

A sequência descreve ações para o personagem H e apresenta em alguns trechos partes confusas, como na primeira frase, indicando as ações contraditórias *estar de pé e sentado no chão* ao mesmo tempo. Assim como está sequência, outras partes do curta apresentam situações confusas em relação as projeções de pessoa, tempo e espaço. Assim como as projeções de espaço realizadas, da mesma forma como no trecho selecionado, não existem ao longo de todo o roteiro indicações sobre a mudança de um ambiente para outro, deixando esta interpretação para o diretor. O trecho termina com H vendo um homem caído por um buraco negro no chão. Não existem indícios no roteiro de que o homem visto era C, como foi realizado em vídeo, podendo indicar neste trecho uma das tentativas de angariar sentido, utilizando-se de um desfecho disfórico do personagem C para assim fechar o seu arco narrativo.

Outras descrições presentes no roteiro do curta-metragem possuem ações com caráter contraditório a sequencialidade da própria narrativa. Entre os exemplos está a ação atribuída ao personagem C: “*ele pega uma tela acesa e luta com as partículas de força de segurança de transmissão em seu rosto*”. O trecho, além de confuso em relação a sua sintaxe, apresenta uma situação semântica que não foi anteriormente elaborada ou preparada. A ação do personagem C de aproximar uma espécie de

³⁴ Trecho original: *He is standing in the stars and sitting on the floor. He takes a seat on the counter and pulls the camera over to his back. He stares at it. He is on the phone. He cuts the shotgun from the edge of the room and puts it in his mouth. He sees a black hole in the floor leading to the man on the roof.*

tablet de seu rosto e realizar um *escaneamento* é apresentada sem motivações que modalizaram o sujeito. Outro exemplo de ação desconexa é dado antes da fala de H, em que o personagem tira da própria boca uma espécie de olho de vidro. Assim como na situação anterior, a ação não possui preparação anterior ou influência futura, sendo completamente ignorada durante o restante da narrativa. Mesmo que estas características sejam colocadas em relação as ações, é preciso evidenciar que esta contradição tem papel na construção do contrato enunciativo dos autores com o espectador. As ações desconexas representam este caráter de experimentação de um roteiro escrito por um computador e, o que pode ser visto neste curta-metragem, foge dos padrões de *uma história comum escrita por um humano*.

Para realizarmos a análise dos diálogos e as ações representadas em vídeo, dividimos o curta metragem em momentos, selecionados pelas transformações das relações entre os personagens. Para facilitar o processo de análise, foi adicionado aos apêndices deste trabalho o roteiro do curta-metragem traduzido para o português.

Momento	Ação
1º	H (primeiro personagem) sentado à mesa, remove um livro da prateleira. Ele expressa que poderia vender sangue para sobreviver sem emprego. H e H2 (segunda personagem) começam a conversar.
2º	C (terceiro personagem) entra na sala e demonstra intimidade com H2. H e H2 começam uma discussão. Forma-se o conflito <i>H versus H2 e C</i> .
3º	Em um novo ambiente, H está de pé em um fundo estrelado e uma cópia sua está sentado no chão em baixo da mesa. Sua versão de pé segura a câmera com uma mão e atende o celular com a outra. H remove uma arma da parede e aponta-a contra a

	própria boca. No chão está um buraco, como um portal, pelo qual H vê-se uma pessoa caída.
4º	Em um novo cenário, C está caído no chão desacordado e H, ao seu lado, pega uma bolsa com sangue de sua mochila. A porta no fundo da cena é fechada. H abaixa-se ao lado do corpo de C e chora.
5º	Olhando diretamente para a câmera, H2 fala emocionada. Ela parece fazer referência ao ocorrido com C e a fuga de H. Entre as frases estão ditas no telefone com H. Fim do curta.

No primeiro momento, o curta apresenta o personagem H. Ele retira de uma estante um livro e o olha. O ambiente em que ele se encontra possui várias telas e computadores, além de equipamentos eletrônicos espalhados pelas mesas. No roteiro, a identificação dada ao ambiente é *SHIP* (NAVE). Em sua primeira fala, o personagem comenta sobre um futuro com um alto índice de desemprego e a possibilidade de jovens venderem sangue para sobreviver. Ele é interrompido pela personagem H2 que o diz para parar de falar e ir ver o garoto. H ainda reclama que não é uma estrela brilhante. O primeiro diálogo entre H e H2 exprime como tema a distopia tecnológica, representado na indicação de um futuro com desemprego em massa. O tema é reiterado por diversas figuras presentes na cena, como os equipamentos eletrônicos sobre as mesas, as roupas semelhantes a trajes espaciais e os sons do ambiente, com ruídos como *beeps* de computadores e um constante som grave, como de um veículo em movimento. A resposta ríspida de H2 e a mudança de comportamento de H em sua fala “*Eu o vi de novo. O jeito que você foi enviada para mim... foi uma grande ideia de verdade. Eu não sou uma luz brilhante*³⁵” indicam

³⁵ Trecho original: *I saw him again. The way you were sent to me... that was a big honest idea. I am not a bright light*

a situação atual de um relacionamento entre os personagens. Ao utilizarmos os recursos da sintaxe narrativa, podemos verificar, o que será reiterado nos próximos momentos em relação ao objeto-valor de H: ter um bom relacionamento com H2. O personagem C entra na sala e este é o aparente empecilho para a relação de conjunção entre H e seu objeto-valor. Neste primeiro momento da narrativa, H é um sujeito em disjunção de seu objeto-valor. Não é possível identificar pelas poucas falas presentes neste momento se o objeto-valor de H2 seria estar afastada de H.

O segundo momento do curta-metragem inicia-se com a entrada do personagem C no ambiente. Ele é maior e mais eloquente do que H. C e H2 demonstram uma relação de afeto positiva em momentos em que interagem juntos: ela o toca nos braços e ele a elogia “*Eu estava indo para aquela coisa porque você é tão bonita*”³⁶. C está em conjunção com o objeto-valor de H e, assim, impedindo-o de poder mudar esta relação. BARROS (2005) explica sobre os programas narrativos de apropriação e espoliação em diversas narrativas populares e apresenta situações identificadas como “quando o sujeito ‘príncipe’ se apropria do objeto ‘princesa’, o sujeito ‘dragão’ é dele espoliado” (BARROS, 2005, p. 27). Este impedimento originado da aparente espoliação incita o conflito entre os personagens H e H2. As falas da discussão entre os dois são confusas e inconclusivas:

H: Você não precisa ser uma médica.

H2: Eu não tenho certeza. Eu não sei do que você está falando.

H: Eu quero ver você também.

H2: O que você está falando?

H: Me desculpe, mas eu tenho certeza que você não me tocou.

H2: Eu não sei do que você está falando.

³⁶ Trecho original: *I was coming to that thing because you were so pretty*

H: O princípio é construído completamente ao mesmo tempo. ³⁷

SHARP (2017) afirma que as falas geradas por Benjamin tendem a ser declarativas e acusativas, como é possível identificar nos trechos acima. A afirmação de SHARP justifica a construção de frases seguindo a forma canônica dos termos que constituem uma oração (sujeito, verbo e complemento), que no inglês mantem a mesma ordem da língua portuguesa. Dessa forma, Benjamin aprendeu a seguir este esquema e a construir frases que se encaixem nestes gabaritos. SHARP (2017) ainda completa seu pensamento sobre as frases elaboradas por Benjamin abordando a geração do sentido. Segundo o cineasta, era preciso identificar nas frases, a partir de uma leitura em voz alta, se elas possuíam o *drama* para poderem ser interpretadas. Esta etapa de leitura e identificação de uma possível dramaticidade habilita o trecho para a interpretação. Este procedimento pode ser identificado como componente entre os primeiros mecanismos de geração de sentido dos textos criados de forma autônoma por Benjamin. Durante o seminário do Machine Making Movies na convenção GitHub Universe 2016, a atriz Safiya Fredericks interpreta para o público trechos gerados por Benjamin durante a apresentação³⁸. A atriz afirma que em trechos gerados automaticamente não é possível compreender como todos os personagens estão se posicionando e, por isso, é preciso utilizar sua experiência de atuação para dar sentido ao texto.

As falas da discussão de H e H2 possuem um tom tênue entre a posse de um sentido ou a falta dele. Podemos verificar que a geração do sentido neste trecho vai além das próprias frases, relacionando-se a outros elementos da linguagem

³⁷ Trecho original: H: You don't have to be a doctor.

H2: I am not sure. I don't know what you're talking about.

H: I want to see you too.

H2: What do you mean?

H: I'm sorry, but I'm sure you wouldn't even touch me.

H2: I don't know what you're talking about.

H: The principle is completely constructed for the same time.

³⁸ Retirado de GitHub Universe 2016 - Machines Making Movies (00:23:00):

<https://www.youtube.com/watch?v=W0bVyx138Bc&t=1460s> acesso em 20 de junho de 2019.

audiovisual, como a interpretação dos atores. H2 encara fixamente H respondendo em tom de ironia e desafio, até mesmo rindo, enquanto seu oponente foge de seu olhar. O volume da voz dos dois aumenta. A montagem das cenas passa a ser acelerada com menos tempo em cada uma das tomadas fechadas. O volume da trilha sonora aumenta e a velocidade de suas batidas se torna frenética. Em um momento de alívio da trilha e montagem, H2 provoca H ao segurar o braço de C estabelecendo o conflito *H vs H2* e *C*. Esta provocação pode ser relacionada a manipulação do personagem H por H2 que será explorada mais adiante no capítulo. Mesmo dizendo frases que não condizem com a natureza da discussão, os personagens se alinham em posição de conflito, de um lado H e do outro H2 e C. Esta topologia da cena é expressa pelo eixo de movimento da câmera, mantendo cada um dos conjuntos do conflito separados, um de frente para o outro.

Fonte: Montagem do Autor

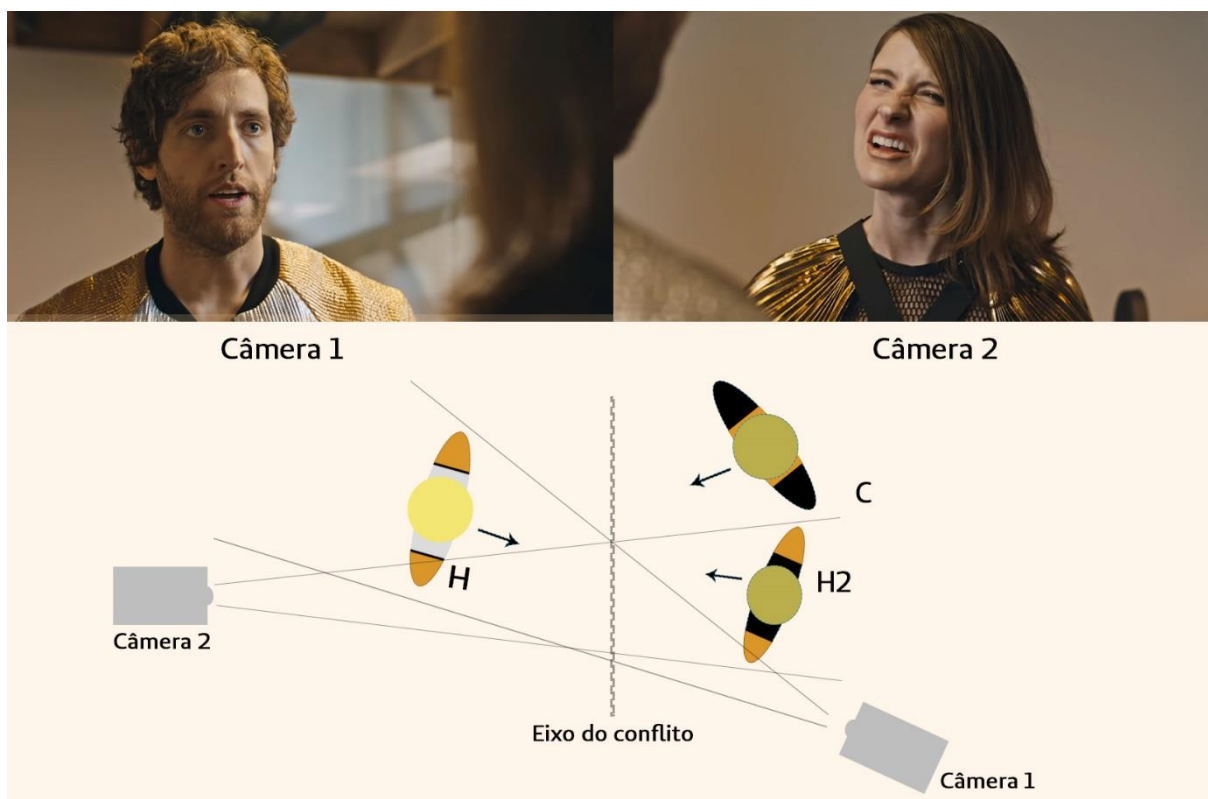


Figura 17: Diagrama de cena: Conflito entre H e H2

Não existem no roteiro de *Sunspring* indicações sobre como os personagens devem agir, no entanto, as falas geradas para o momento de conflito são menores, o que permitiu aos atores a interpretação de forma mais rápida, incitando as características de afronta e ironia. O conflito tem um hiato com a intervenção de C sobre a discussão de H e H2. Ele se aproxima de H e a diferença de altura dos personagens fica evidente. C é mais alto e esta característica representa diversos aspectos inseridos na cultura ocidental. Ser mais alto é uma figura para ser mais forte, melhor e dominador. Este é o *dragão* das fábulas de H, o que impede de transformar sua relação de disjunção com seu objeto-valor. O plano conjunto dos dois personagens demonstra a altura de C mantendo uma grande área vazia acima da cabeça de H, forçando-o a olhar para baixo. A trilha sonora é interrompida e o foco está inteiramente na próxima ação dos personagens.

Surge a expectativa do início de um conflito físico entre os personagens. Esta foi criada pelo posicionamento dos dois e a atitude de enfrentamento de C ao se aproximar de H, com o peito inflado e encarando-o diretamente. A expectativa é rompida com a fala “*Eu não vou fazer algo*³⁹”. A frase indica uma possibilidade de intenção de cessão do conflito por C, que sugere a H que volte a sua mesa. Após a sugestão, o ritmo de tensão volta à cena com a trilha sonora frenética e H incitando o conflito físico enquanto acerta C com o dedo. H2 se aproxima dos dois em uma preparação para a intervenção. A frase dita por H para os dois personagens não possui qualquer nível de conexão semântica com a discussão, mas é interpretada em tom de fechamento pelo ator. “*É uma coisa muito horrível para dizer. Nada vai ser uma coisa, mas fui eu quem pegou essa pedra com uma criança e depois deixei as outras duas*⁴⁰”.

³⁹ Trecho original: “*I'm not going to do something.*”

⁴⁰ Trecho original: “*Mmm. It's a damn thing scared to say. Nothing is going to be a thing but I was the one that got on this rock with a child and then I left the other two.*”

Fonte: Reprodução / Sunspring



Figura 18: Personagens H e C encaram-se

SHARP (2017) afirma que os atores “não sabiam sobre o que era a história, assim que eles leram uns para os outros, surgiu um triângulo amoroso de lugar nenhum, ele não estava lá” (SHARP, 2017). A afirmação de Sharp demonstra como as diversas linguagens que compõem o audiovisual podem ressaltar uma história já conhecida no imaginário popular, dois homens brigando por uma mulher, mesmo de um roteiro com falas que não apresentam ou corroboram para esta história. Segundo SHARP e GOODWIN (2016) o texto gerado por Benjamin tem como principal objetivo o de criar experiências emocionais. A partir disto, é possível criar um roteiro. Esta criação de experiências é aprofundada por GOODWIN:

The cycle of generation and interpretation is what demonstrates the true augmented capacity of this learning machines and how they can help us to tell new types of stories that was never seen or heard before ⁴¹ (GOODWIN, Ross, Machines Making Movies, 2017)

⁴¹ Tradução livre: “O processo de geração e interpretação é o que demonstra a verdadeira capacidade aumentada destas máquinas de aprendizagem e como elas podem nos ajudar a contar novos tipos de histórias que nunca foram vistas ou ouvidas antes”

Este potencial do aprendizado de máquina na criação de cenas ricas em provocações e de elementos abstratos cria barreias à aplicação de Benjamin, e sistemas semelhantes, em narrativas com uma estrutura teleológica de produtos audiovisuais com cunho comercial, como as produções de filmes *blockbuster*, novelas e séries. No entanto, o sistema pode ser de grande aplicabilidade para criações no campo artístico, como no cinema autoral. *Sunspring* caracteriza-se como um experimento e está alocado no campo da arte. A partir deste pensamento, podemos identificar que o contrato enunciativo de experimentação se torna mais complexo, ramificando a experiência para a articulação das linguagens que compõe o audiovisual na composição do sentido pretendido.

O terceiro momento do curta-metragem inicia-se com uma mudança de cenário. A temática futurística representada na primeira cena continua figurativamente reiterada no novo ambiente retratado neste novo momento. A parede e chão de estrelas, os equipamentos eletrônicos sobre a mesa, a pistola presa e o buraco negro no chão. O personagem H está de pé sobre um fundo estrelado e uma sua cópia idêntica está sentada no chão embaixo de uma mesa. Não existem elementos que anteriores ou posteriores que alterem a ordem cronológica natural do curta. Dessa forma, a cena noturna deve se passar após o conflito entre H com os outros personagens. É possível ouvir um telefone tocando. O personagem que está de pé encara a câmera e a puxa com o braço, como se fosse fazer uma *selfie*. Ele atende ao telefone e é possível ouvir a voz de H2 dizendo “*eu só queria te falar que eu sou muito melhor que ele. Eu tive de pará-lo e não consegui nem te contar*⁴²”. A câmera se livra da mão do personagem e vemos um plano médio do outro lado da sala em que H percebe que não está mais segurando nada. SHARP (2016) afirma que o trecho demonstra a combinação de criação entre máquina e homem, de forma que toda a elaboração da primeira parte da cena, em que H segura a câmera e fica a encarando foi realizada pelo computador. Já a segunda parte, em que a câmera salta para o outro

⁴² Trecho original: *I just wanted to tell you that I was much better than he did. I had to stop him and I couldn't even tell*

lado da sala e H percebe, foi intencionalmente criada por humanos. Esta *co-criação* citada reforça o caráter gerativo de sentido humano sobre a obra criada pela máquina, selecionando elementos de diversas linguagens para uma adequação semântica.

H desliga o celular enquanto olha para a câmera do outro lado da sala. A frase ouvida na voz de H2 faz referência a uma terceira pessoa da qual ela precisou se livrar. Neste momento, identificamos a primeira alteração humana na composição da narrativa diretamente no roteiro. A cena original é composta apenas de ações dos personagens e mantém-se igual até H atender o telefone. Não há indicações que H2 estava do outro lado da linha com H ou o que ela poderia ter falado. No entanto, a fala foi adicionada a montagem final curta e seu resultado foi a modalização do personagem H para os próximos acontecimentos da narrativa. A cena continua da mesma forma como foi elaborada no roteiro original: H aproxima-se da parede, remove uma pistola e a aponta contra a própria boca. Nenhum tiro é disparado. A câmera aponta para baixo e no chão vemos um portal em forma de buraco pelo qual se vê uma pessoa desmaiada.

A fala adicional de H2, utilizando o recurso de *voice over*, foi estrategicamente posta antecedendo uma ação que influencia diretamente a estabilidade narrativa em do curta-metragem. A intenção de suicídio de H é uma ação forte e ficaria sem qualquer motivação aparente antes ou depois da tentativa. Dessa forma, encaixar a fala de H2 modaliza o personagem em sua ação e cumpre um contrato enunciativo com o espectador de que esta ação não está sendo tomada de forma impulsiva, mas trabalhada a partir de uma motivação. A primeira fala de H2 (*eu só queria te falar que eu sou muito melhor que ele*) faz uma comparação da própria personagem com um terceiro. Já a segunda frase se relaciona a uma medida drástica que precisou ser tomada com pressa. Esta identificação fica evidente nos trechos “*eu tive de pará-lo*” e “*não consegui nem te contar*”. O momento é encerrado com um mergulho da câmera no buraco negro do chão a abertura em um novo cenário.

Em um novo ambiente externo H encontra C caído no chão. Este novo cenário é destoante de todos os outros do curta-metragem, sem uma grande população de figuras que façam referência a tecnologia ou futurismo. A presença de H nesta cena

indica, em uma trajetória linear, que o personagem desistiu do suicídio. Algumas figuras identificam um desenrolar do arco narrativo do personagem H: a mochila em suas costas e uma porta, no fundo da cena, fechando-se indicam que o personagem está saindo daquele lugar. Ao visualizarmos um panorama de toda a narrativa, podemos identificar que a fala de H2 no telefone com H no momento anterior não modalizou o personagem apenas a tentativa de suicídio, mas também a fazê-lo ir embora. Podemos identificar que o trecho inserido por um humano no roteiro autônomo reverbera sobre a geração de sentido do curta, como uma correção de curso narrativo, acompanhando os arcos dos personagens até o fim do texto.

Fonte: Reprodução / Sunspring



Figura 19: H encontra C morto e segura uma bolsa de sangue

H remove de sua mochila uma bolsa plástica selada com um líquido vermelho que se assemelha a sangue. O personagem abaixa-se ao pé de C e chora. O trecho referente a cena no roteiro original é “*ele vem atrás dele para protegê-lo. Ele ainda está de pé ao lado dele*⁴³”. O momento representa o clímax da narrativa, iniciando a resolução dos conflitos estabelecidos nos momentos anteriores. C está morto e este é fim de seu arco narrativo. No entanto, o conflito *H vs C* por H2 termina de forma

⁴³ Trecho original: *He comes up behind him to protect him. He is still standing next to him*

disfórica, com os dois personagens em disjunção com o objeto-valor. H demonstra a intenção de deixar tudo e seguir com um plano alternativo ao segurar a bolsa de sangue para talvez vendê-la para sobreviver sem um emprego, uma ideia citada anteriormente já no primeiro momento do curta-metragem. O roteiro marca a remoção de uma bolsa da mochila mas não indica seu conteúdo: “*ele olha para a bolsa da mochila e começa a chorar* ⁴⁴”. A construção da figura de uma bolsa de sangue neste momento da narrativa pode indicar mais uma alteração semântica do roteiro original por parte dos autores do curta-metragem. A forma como Benjamin cria, que será explorada no capítulo seguinte, faz com que seja muito baixa a probabilidade do curta fazer uma referência a um elemento anterior já citado. Assim, mesmo que este trecho tenha sido criado pelo sistema, este provavelmente não fez qualquer ligação com a possibilidade de venda de sangue citada no primeiro momento. Logo, a utilização do sangue como conteúdo da bolsa demarca a atribuição semântica a uma figura que já estava presente no curta, criando uma reiteração sobre o final disfórico do arco narrativo do personagem H.

O último momento do curta-metragem pode ser relacionado à conclusão de um percurso de manipulação conduzido pela personagem H2. A cena apresenta H2 olhando para a câmera rompendo com a quarta parede. A personagem fala como se confessasse algo para o público. O trecho referente no roteiro original apresenta um novo personagem, T, que foi suprimido da produção audiovisual. Todas as falas do novo personagem foram alocadas para H2:

Bom, existe essa situação entre mim e a luz na nave. O cara estava tentando me parar. Ele é como um bebê e foi embora. Eu estava preocupada com ele. Mas mesmo se ele tivesse feito tudo. Ele não conseguiria voltar. Eu não pretendo ser uma virgem. Quer dizer, ele era fraco. E eu pensei em mudar de ideia. Ele era louco para ir embora. Foi a muito tempo atrás. Ele estava um pouco atrasado. Eu ia ser

⁴⁴ Trecho original: *He looks at the bag from his backpack, and starts to cry.*

um momento. Eu só queria te contar que eu sou muito melhor que ele. Eu tive de pará-lo e não consegui te contar. Eu não queria machuca-lo. Me desculpe. Eu sei que eu não gosto dele. Eu posso ir para casa e ser tão má e eu o amo. Então eu posso pegá-lo todo o caminho até aqui e encontrar o quadrado e ir para o jogo com ele e ela não vai aparecer. Então eu vou dar uma olhada. Mas eu não vou ver ele quando ele chegar para mim. Ele olha para mim e me joga fora dos olhos. Então ele disse que iria para a cama comigo.

Ao analisarmos o monólogo, podemos identificar algumas características em comum com outras partes do texto. Assim como na discussão entre H e H2 no segundo momento do curta-metragem, as frases são curtas e acusativas. Durante a leitura é possível encontrar sentidos e referências ao enredo, mas esta identificação está situada entre uma linha tênue do entendimento e a dúvida. Ao analisarmos o trecho seguindo os parâmetros de uma narrativa teleológica, como foram vistos os momentos anteriores, podemos identificar a conclusão de um ponto de virada. O trecho é a confissão de um plano H2 de livrar-se de C e continuar apenas com H na nave, o que nos permite observar a narrativa a partir do ponto de vista de outra personagem. Ao definirmos H2 como sujeito da manipulação sobre H, podemos identificar os percursos narrativos do destinador-manipulador, do sujeito e do destinador-julgador e seus respectivos programas durante a narrativa.

Fonte: Reprodução / Sunspring



Figura 20: H2 confessa suas intenções para a câmera

O programa narrativo de doação de competência semântica é a primeira etapa do processo de manipulação do sujeito, pois é preciso que esse creia primeiramente nos mesmos valores (BARROS, 2005). Podemos identificar esta doação de valores de H2 para H no primeiro e segundo momento da narrativa, em os dois personagens entram em conflito sobre C. Nesta situação, H2 doa para H valores semânticos relacionados ao impedimento que C representa, impedindo-o de alcançar seu objeto-valor. Ao gerar o conflito *H vs C*, H2 é bem-sucedida e pode passar para a próxima etapa da manipulação. A modalização de H por H2 para a realização de seu plano se dá por meio da provocação. H2 provoca H ao interagir com C fisicamente e dar gracejos ao personagem. Desta forma, ela modaliza H em um dever-fazer, gerando uma imagem negativa de si mesma para o manipulado.

Mesmo com H modalizado não há o cumprimento da manipulação. Os programas narrativos de competência e performance não são cumpridos por H. O personagem não possui a competência para a realização da performance, rompendo com o contrato de manipulação. Assim, resta a H ser julgado no último percurso, o do destinador-julgador. H2 realiza seu próprio plano sozinha, livrando-se do personagem C. Em seu plano narrativo de interpretação, H2 julga a deserção de H como negativa.

Este julgamento é representado durante a ligação de H2 para H em que a personagem indica que já se livrou C. A repercussão deste julgamento é forte para H, levando-o a tentativa de suicídio. O personagem desiste desta tentativa, mas ainda precisa passar pela consequência de sua deserção. O programa narrativo de retribuição dá a H sua punição, expulsando-o da nave, forçando o personagem a encontrar uma forma alternativa para sobreviver.

O final da narrativa para H2 também é disfórico, de forma que a personagem iniciou seu processo de manipulação em busca de seu objeto-valor e obteve como resultado seu contrato rompido. Assim, o curta termina de forma negativa para todos os personagens. Ao avaliarmos as transformações operadas durante a narrativa, podemos indicar a passagem de C, que estava em conjunção com seu objeto-valor e passou para o estado de disjunção. Os personagens H e H2 possuem suas transformações operadas dentro do próprio disforia, de forma que passam de seus estados disjuntivos para outro mais complexo e extremo. O final de *Sunspring* pode remeter ao gênero tragédia (ou drama) do teatro grego. No gênero, o sofrimento dos personagens humanos é irreversível, evidenciando o final catastrófico de cada um. *Sunspring* segue estas características: C termina morto, H é obrigado a deixar o local e vender sangue para sobreviver e H2 evidencia em seu monólogo sua decepção com os outros personagens e como teve de ficar sozinha.

Esta análise e interpretação da obra foi realizada a partir de uma leitura semiótica, seguindo as estruturas do percurso gerativo do sentido. Como foi demonstrado, diversos elementos presentes no curta-metragem foram dispostos com o intuito da geração de sentido. *Sunspring* se apresenta, enquanto obra audiovisual, como uma colaboração criativa entre humano e máquina. SHARP (2017) afirma que Benjamin funcionou como um membro da equipe de produção do curta e este reuniu em seu acervo de roteiros membros de diversas outras equipes. GOODWIN (2016) alerta para o fato que roteiristas não devem temer um roteiro escrito por um computador e empolga-se com a possibilidade de automação de processos audiovisuais criativos. “Podemos nos tornar escritores de escritores” (GOODWIN,

2016). A produção do roteiro de *Sunspring* por um computador coloca em discussão a posição de Benjamin enquanto membro criativo da equipe do curta-metragem.

Benjamin possui um papel actancial em relação a formação da narrativa e, mesmo sendo um *software*, é diferente de outras ferramentas automatizadas que auxiliam durante as etapas de produções audiovisuais. Entre os fatores de influência desta relação está a capacidade de criação que o sistema possui. O conceito de criatividade é uma noção da psicologia que foi introduzida a linguística originalmente por Noan Chomsky e sua definição está presente no livro *Dicionário de Semiótica*:

Faculdade de produzir e compreender frases * novas, devida ao caráter recursivo * das construções sintáticas. A criatividade, assim compreendida, deve ser considerada como uma propriedade de competência do sujeito falante. O caráter operatório * desse conceito é evidentemente fraco ou nulo: dado que as possibilidades combinatórias * de uma língua não natural são praticamente infinitas, isso equivale mais ou menos a dizer que o “espírito humano” é criativo (GREIMAS e COURTÉS, p. 92, 1979)

A teoria gerativa de Chomsky aborda a aquisição da linguagem a partir da instância do conhecimento “como uma sequência de representações e processos mentais, ou seja, o homem não tem acesso direto ao mundo, mas esse acesso é mediado por operações mentais” (AUGUSTO, p. 116, 1995). Podemos relacionar esta mediação no gerativismo de Chomsky e a definição de criatividade, enquanto uma propriedade do sujeito falante, com a forma que Benjamin aprende e cria novos textos a partir de um *corpus*. O sistema aprende de forma única e exclusiva a partir dos textos de entrada. Este é seu universo e sua forma de criatividade se depreende apenas da maneira que foi programado para entender este conjunto. Benjamin realiza uma tarefa singularmente humana, a de contar histórias. Assim como nos baseamos em nossas referências para compor uma narrativa, o sistema também é sensível a este tipo de influência.

Não podemos antropomorfizar o sistema como um membro da equipe de produção assim como seus autores o indicam. Benjamin executa tarefas e reage a estímulos. No entanto, sua capacidade criativa apresenta resultados que são

inesperados para a relação humano/máquina. PRIMO (2000, p. 8) propõe o conceito de interação reativa, entre humanos e máquinas, e interação mútua, entre humanos. Segundo o autor, durante o processo de interação reativa, um estímulo dado a uma máquina retornará sempre os mesmos resultados se todas as condições forem as mesmas. Benjamin não escapa desde conceito, e se seu banco de dados e o *estímulo semente* forem os mesmo então sua resposta será a mesma. No entanto, Benjamin é uma rede neural e seu funcionamento é diferente de um sistema padrão. Sua forma de criação é intermediada a partir de números aleatórios que compõe diversos parâmetros de criação. Dessa forma, uma chamada comum do sistema, sem uma semente específica, trará resultados diferentes todas as vezes, até a impossibilidade de ineditismo. Esta forma de processamento não usual atribui ao sistema a emulação de um caráter criativo.

PRIMO (2000, p. 8) também propõe uma análise sobre o *throughput*, “o que se passa entre uma ação e outra, ou entre uma ação e uma reação”. Em uma interação reativa, esta etapa é marcada pela resposta aos reflexos. Segundo o autor, há uma falsa aparência de interpretação do computador sobre os estímulos:

O computador não trabalha com significações, nem comunicação, apenas com possibilidades combinatórias abstratas. Logo, como pode haver interpretação em um sistema meramente sintático, onde as combinatórias podem ocorrer mesmo que longe da semântica. (PRIMO, 2000, p. 8)

Avanços do poder de processamento dos computadores e da inteligência artificial permitem que sistemas tenham suas etapas *throughput* incrementadas. O processamento de índices e pesos relacionados a tomada de decisões por uma rede neural pode ser identificado como uma meta semântica, definindo um sentido para as propriedades aplicadas em um resultado. Esta meta semântica possui um escopo fechado ao próprio sistema, de forma que um texto gerado por Benjamin faz sentido a maneira que este é gerado, mas não possui qualquer relação semântica fora deste processo. A partir deste conceito, avaliaremos no próximo capítulo a possibilidade de estruturação de uma etapa do *throughput* de um sistema de criação de texto

semelhante ao Benjamin, para uma tradução desta meta semântica em dispositivos de geração de sentido.

CAPÍTULO 3

A CONSTRUÇÃO DO SENTIDO EM TEXTOS AUTÔNOMOS

CAPÍTULO 3

A CONSTRUÇÃO DO SENTIDO EM TEXTOS AUTÔNOMOS

Este capítulo tem como objetivo apresentar, sob o aporte comunicacional, uma análise da criação de texto de forma autônoma em sistemas semelhantes a Benjamin. Para isto, nossa metodologia consistiu em utilizar um *software* para a criação de texto autônomo com os mesmos dados utilizados para a geração de Sunspring e analisar o resultado ao final. Além disso, avaliaremos a criação de uma camada de geração semântica para estes sistemas, a partir das estruturas do percurso gerativo de sentido, e a forma como está deve ser constituída.

3.1 - Processo de criação autônoma de texto

Atualmente, em uma busca rápida na internet, é possível encontrar alguns serviços que oferecem o processamento de diversos tipos de dados em sistemas com inteligência artificial. Como um dos campos de programação mais complexos, definimos nossa busca por um sistema já pronto, mas com as mesmas características que Benjamin possui. Assim, foram buscados os códigos prontos de redes neurais de recorrências do tipo LSTM (*long short term memory recurrent neural network*), os mesmos parâmetros de Benjamin. A alternativa que se saiu melhor para estes parâmetros foi o software Word-rnn-Tensorflow (que será abreviado como WRT), criado pelo professor de ciência da computação Sung Kim, da Universidade de Ciência e Tecnologia de Hong Kong. O código fonte do programa foi publicado em

licença Creative Common Attribution na plataforma de versionamento GitHub, que indica que seu uso é permitido desde que sejam feitas as devidas menções ao autor.

Publicado em maio de 2016, o código de Kim contou com nove colaboradores para alcançar seu atual patamar. O *software* foi desenvolvido a partir da conjunção de duas tecnologias já presentes no ambiente de processamento de dados: a linguagem de programação Python, atualmente em sua versão 3, e o TensorFlow, uma biblioteca de auxílio no desenvolvimento de *deep learning*. Atualmente, o Python é a terceira linguagem de programação mais utilizada no mundo. Sua escrita simples e sua compatibilidade com um grande número de bibliotecas ligadas a computação científica permitem que a linguagem seja utilizada por diversos profissionais. O TensorFlow é uma biblioteca desenvolvida pela Google com o objetivo de auxiliar no processamento de programas com *machine learning*. A tecnologia permite a criação de grafos, gráficos de processamento da aplicação, auxiliando o programador a definir quais partes do sistema serão processadas primeiro e, até mesmo, ratear o processamento das informações em mais de um computador.

Fonte: Reprodução/GitHub

Multi-layer Recurrent Neural Networks (LSTM, RNN) for word-level language models in Python using TensorFlow.

rn tensorflow rnn-tensorflow lstm python

73 commits 4 branches 0 releases 9 contributors MIT

Branch: master New pull request Find File Clone or download

scubbo and hunkim Add some command-line options (#68) Latest commit 54f11c5 on 21 Dec 2017

File	Description	Time
data/tinychakespeare	Added tests	3 years ago
logs	Create logs directory in repository	3 years ago
save	initial commit	3 years ago
tests	Adds tests for beam search.	3 years ago
.gitignore	Added basic tests. Need more	3 years ago
.travis.yml	update travis	2 years ago
LICENSE.md	initial commit	3 years ago
README.md	Update requirements in README	2 years ago
beam.py	Updates the README and fixes bugs in beam.	3 years ago
model.py	Add some command-line options (#68)	2 years ago
sample.py	Add some command-line options (#68)	2 years ago
train.py	Add an --input_encoding argument to train.py (#56)	2 years ago
utils.py	Add an --input_encoding argument to train.py (#56)	2 years ago

Figura 21: Repositório do código de WRT

Nossa escolha pelo uso do WRT deve-se ao grau de complexidade de um programa como este e Benjamin. Dos diversos softwares analisados para a recriação de uma experiência semelhante a desenvolvida por SHARP e GOODWIN com o Benjamin, o WRT foi o melhor selecionado. Além disso, a escolha de um *software livre* gratuito, e desenvolvido em um ambiente universitário, está em comunhão com o objetivo acadêmico desta pesquisa, possibilitando o acesso ao conhecimento e tecnologias nas áreas de estudos da geração do sentido e do processamento de dados. O WRT possui limitações em relação as versões mais novas do Benjamin e a forma como este funciona, no entanto, ambos apresentam resultados semelhantes e que precisam de interferência humana para a sua plena finalização como um roteiro.

Em seus artigos *How to write with artificial intelligence* e *Harry Potter: Written by Artificial Intelligence*, o jornalista Max Deutsch apresenta como utilizou o WRT para a criação de novos trechos em texto com os personagens da franquia Harry Potter. O autor mantém uma coluna no site Medium com trechos de Harry Potter e outras franquias, todas geradas a partir do uso de inteligência artificial. Para que o WRT funcione, é preciso que grandes trechos de texto semelhantes a saída desejada sejam adicionados ao seu banco de dados. Após a adição dos dados, é preciso iniciar o treinamento do sistema, momento em que as técnicas de *deep learning* serão aplicadas sobre o conteúdo de forma que um algoritmo de criação autônoma possa ser criado. O treinamento com os dados é a etapa fundamental de todo o processo e, de forma geral, a quantidade e a qualidade dos dados inseridos afetarão o algoritmo de criação final. De forma semelhante a Max Deutsch, conduzimos inicialmente um experimento de criação de texto de forma autônoma, inserindo no banco dados do sistema todos os livros da franquia Harry Potter em seu idioma original, o inglês. Para isto, foi preciso selecionar apenas os textos dos livros, removendo a formatação e as informações que podem confundir o sistema, como a numeração das páginas, índices dos livros e o nome dos capítulos. A escolha do idioma inglês foi apenas para manter a reprodução das características dos experimentos de Deutsch, já que o WRT não

utiliza nenhuma biblioteca de linguagem natural pronta, permitindo que qualquer idioma fosse utilizado.

Após a adição dos livros, foi dado início ao treinamento do sistema. Sua duração foi de aproximadamente 1 hora e 20 minutos de processamento. Para a execução do treinamento, é preciso a instalação do TensorFlow no computador. Para este experimento, devido à grande quantidade de dados a serem processados, foi utilizada a versão do TensorFlow com processamento pela GPU, placa de vídeo. Nesta versão, o processamento não ocorre somente pelo processador do computador, como em *softwares* comuns, mas de forma conjunta entre o processador e a placa de vídeo, permitindo mais performance ao sistema. Dessa forma, foi possível garantir que o treinamento fosse realizado de maneira mais rápida pelo sistema. Com o fim do treinamento, foi possível gerar nossos primeiros trechos de Harry Potter escritos por uma inteligência artificial.

O processo de geração possui um tempo muito menor que o de treinamento, criando um trecho com 400 palavras em aproximadamente dois minutos. Assim como o idioma original dos livros adicionados, o sistema gerou trechos em inglês. Para sua plena apresentação, os trechos são traduzidos e algumas correções devem ser feitas, como a adição de pontuação e até mesmo a separação das frases. Todas as mudanças feitas no texto, assim como as realizadas por SHARP e GOODWIN (2016) estão relacionadas a sintaxe e não a semântica, embora estas instancias estejam ligadas diretamente uma à outra.

A manhã de Harry estava quase brilhando na cozinha. A cama feia começou a mexer febrilmente. Ele gotejou sua distância para a estação na água, árvores prateadas empoleiradas em cima de Trevor e James. “Eu vou vender isso por algum valor.” Eles eram agora favoritos como sempre, todos eles através da doninha que dizia:

SEVERUS SNAPE CONFIRMADO COMO DIRETOR DE HOGWARTS.

Os corvinais tinham ido embora, e Yaxley se esquivou deles com um anel, ameaçando como uma grande placa esparramada sobre a mesa, colocada fora todos os aros para a noite do que parecia fazer um presunto inteiro. Eles estavam tentando em seu nome e aconteceu, exceto que eles estavam começando a visitar tudo isso antes do Halloween, tentando quebrar seus olhos restantes. No entanto, eles estavam se afastando nos olhos. “Nossa defesa contra as artes das trevas. Harry Potter amava me proteger”, o Lorde das Trevas gritou para Harry dando duro na cama. “Meu Senhor, eles tiraram um relatório da cabeça dela. Droga.” Mas Dudley levantou sua varinha.

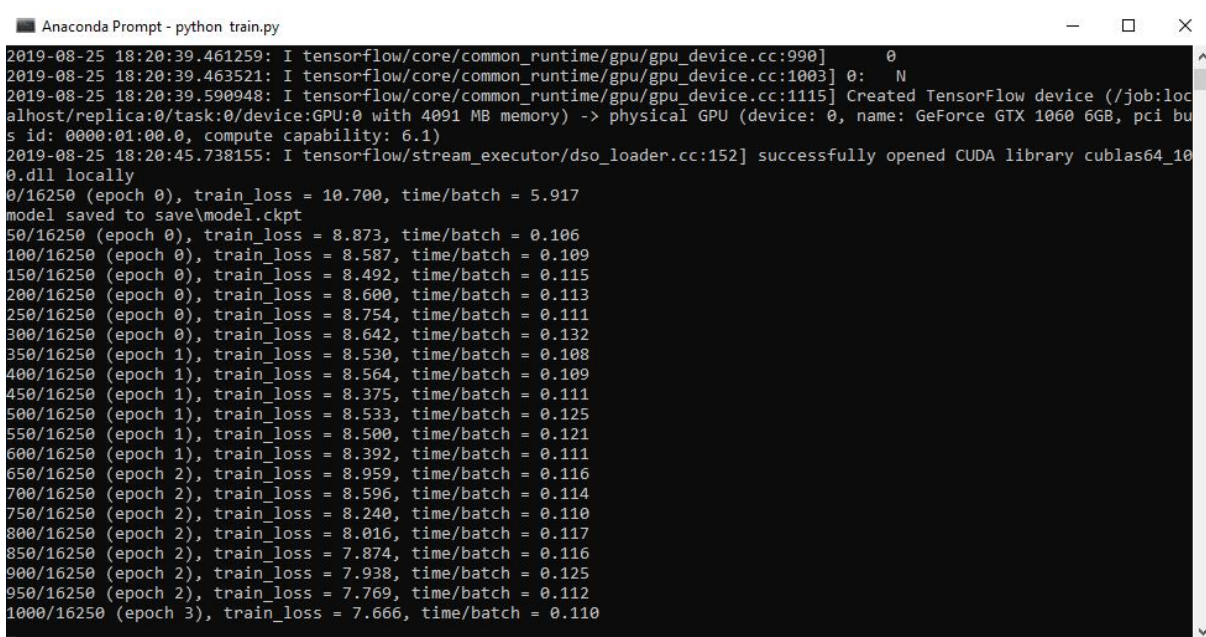
O trecho acima foi gerado escrito por um computador e traduzido e formatado por um humano. Assim como identificamos no segundo capítulo, durante a leitura do trecho verificamos algumas sensações relativas a confusão e desordem de um sentido. Isto acontece, assim como em Sunspring, devido ao caráter volátil da semântica do texto. Dessa forma, apenas com esta leitura, podemos perceber que, de forma tênue, este pode fazer ou não sentindo. O texto constantemente nos guia por novas linhas de pensamento enquanto rompe com outras. O trecho possui diversos elementos que nos guiam para dentro do universo narrativo de Harry Potter. Os personagens Harry, Trevor, James e todos os outros fazem parte deste universo. As figuras como árvores prateadas, Halloween e varinha remetem ao tema magia que está no centro da franquia do bruxo e figurativizam o texto. No entanto, estas figuras são muito esparsas de sentido ao trecho em si, apenas reiterando um tema muito geral.

Para que um trecho fosse criado, foi preciso que algumas tentativas fossem realizadas anteriormente. Assim como descritos por SHARP e GOODWIN (2017), os processos de criação de texto, ao menos no Benjamin e no WRT, ainda não possuem um elevado grau de acerto, gerando texto que não se encaixam a qualquer norma padrão da língua. Por isso, o primeiro critério de seleção humano após a criação de

um trecho já é semântico, de forma que deve selecionar e preparar o trecho para a apresentação.

Para iniciarmos nosso experimento, optamos pela importação dos mesmos roteiros escolhidos para compor o banco de dados de Benjamin. Todos os roteiros foram encontrados em sites especializados na publicação deste tipo de conteúdo e suas licenças cobrem a utilização acadêmica do material. Após a organização dos roteiros, estes foram adicionados ao banco de dados do sistema para o início da próxima etapa. O processo de treinamento do sistema foi dividido em três partes que totalizaram aproximadamente 14 horas de processamento de mais de 24.000 páginas de texto. A grande quantidade de informações adicionadas impossibilitou a remoção de todas as características dos arquivos que pudessem atrapalhar o processo de treinamento. Desta forma, esta etapa foi realizada de forma superficial, removendo em massa diversas informações problemáticas, mas, possivelmente, mantendo alguns elementos que podem ter atrapalhado o treinamento do sistema.

Fonte: acervo do autor



```
Anaconda Prompt - python train.py
2019-08-25 18:20:39.461259: I tensorflow/core/common_runtime/gpu/gpu_device.cc:990] 0
2019-08-25 18:20:39.463521: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1003] 0: N
2019-08-25 18:20:39.590948: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1115] Created TensorFlow device (/job:localhost/replica:0/task:0/device:GPU:0 with 4091 MB memory) -> physical GPU (device: 0, name: GeForce GTX 1060 6GB, pci bus id: 0000:01:00.0, compute capability: 6.1)
2019-08-25 18:20:45.738155: I tensorflow/stream_executor/dso_loader.cc:152] successfully opened CUDA library cublas64_10.dll locally
0/16250 (epoch 0), train_loss = 10.700, time/batch = 5.917
model saved to save\model.ckpt
50/16250 (epoch 0), train_loss = 8.873, time/batch = 0.106
100/16250 (epoch 0), train_loss = 8.587, time/batch = 0.109
150/16250 (epoch 0), train_loss = 8.492, time/batch = 0.115
200/16250 (epoch 0), train_loss = 8.600, time/batch = 0.113
250/16250 (epoch 0), train_loss = 8.754, time/batch = 0.111
300/16250 (epoch 0), train_loss = 8.642, time/batch = 0.132
350/16250 (epoch 1), train_loss = 8.530, time/batch = 0.108
400/16250 (epoch 1), train_loss = 8.564, time/batch = 0.109
450/16250 (epoch 1), train_loss = 8.375, time/batch = 0.111
500/16250 (epoch 1), train_loss = 8.533, time/batch = 0.125
550/16250 (epoch 1), train_loss = 8.500, time/batch = 0.121
600/16250 (epoch 1), train_loss = 8.392, time/batch = 0.111
650/16250 (epoch 2), train_loss = 8.959, time/batch = 0.116
700/16250 (epoch 2), train_loss = 8.596, time/batch = 0.114
750/16250 (epoch 2), train_loss = 8.240, time/batch = 0.110
800/16250 (epoch 2), train_loss = 8.016, time/batch = 0.117
850/16250 (epoch 2), train_loss = 7.874, time/batch = 0.116
900/16250 (epoch 2), train_loss = 7.938, time/batch = 0.125
950/16250 (epoch 2), train_loss = 7.769, time/batch = 0.112
1000/16250 (epoch 3), train_loss = 7.666, time/batch = 0.110
```

Figura 22: WRT sendo treinado com os roteiros

Para a criação de um resultado mais sucinto, foi proposto que o sistema gerasse apenas alguns diálogos ao invés de um roteiro de curta-metragem completo. Esta escolha foi feita em vias de beneficiar o processo de análise do trecho gerado. Dessa forma, foram definidos, assim como em Sunspring, letras para os nomes dos personagens. Esta escolha em nosso texto e em Sunspring é um reflexo da forma como o sistema escolhe os nomes. Em seu processo de criação, o sistema segue a estrutura de indicações dos nomes dos personagens seguidos por suas falas, como são formatados os roteiros. Em alguns momentos, o sistema será induzido pela base a indicar uma ação ou um cabeçalho de cena, indicando a localidade e as condições do ambiente.

Verificamos já nesta etapa da criação o baixo nível de conexão semântica entre as projeções feitas pelo sistema diretamente no roteiro. O texto gerado indicará no as projeções de pessoa, espaço e tempo de forma desconexa entre as instâncias, puramente concebidas a partir do *corpus* disponível. A relação de aleatoriedade para esta determinação influencia diretamente no entendimento do texto, criando situações confusas para história. Como forma de contornar esta aleatoriedade, os produtores de Sunspring utilizaram como artifício a seleção de personagens manualmente, dando a estes como nome apenas uma letra (H, H2 e C). Dessa forma, foi possível atribuir falas criadas para outros personagens para o pequeno grupo criado, selecionando de forma analógica o que seria dito. O mesmo mecanismo foi utilizado aqui, respeitando a ordem de aparição das frases e atribuindo aos personagens como nome apenas sua inicial: J (Jimbo - Os Doze Macacos, 1995, EUA), L (Leia - Star Wars: Episódio IV - Uma Nova Esperança, 1977, EUA) R (Ray Arnold - Jurassic Park: O Parque dos Dinossauros, 1993, EUA). Em relação ao título, não há indicações durante as falas ou textos de SHARP e GOODWIN analisados sobre como este foi concebido. Em nosso texto, o sistema elaborou o título Orange. Esta é a tradução do diálogo elaborado:

ORANGE

EXT - TERRAÇO - LOS ANGELES

Vemos a sala vazia. Em uma rocha estreita, vemos J.

Ele está se segurando sobre uma mesa.

L

Você vai dormir aqui, mas eu disse que vai ser simpático.

Tudo graças à casa de quatro moedas.

R

As pessoas podem fazer mais.

L

Não é uma ligação! Deveria ser.

R (para o robô)

Pode continuar sendo a criatura e as portas continuam.

Seus olhos percorrem a cena enquanto o misterioso ROBÔ aperta os olhos... e depois pula através, sai da rampa e lê:

J

CHUVA!

L

Você nunca captou um sinal de energia e ...

R

Deixe-me falar com alguém desligado, então eu vou ficar no lugar.

Deus não gosta de mim.

L

Encontramos a busca pela nossa palavra "germes"!

R

Um burocrata caminha para si e o Black Widow chama o grande complexo.

L

O conhecimento é uma escolha. Fique longe de um feliz.

R

Não odeia isso? Não, eu odeio qualquer modificação no tempo que estamos.

Ângulo sobre um Mercedes estacionado onde J está se escondendo.

R (rotula sua voz)

Eu estou indo para isso?

L

Eu sabia que você não concluiria

R

Você está certo. Você me encontra. Dois olhares devem ser selados em uma certa casa.

L (O.S.)

Precisamos de uma bala neste capacete.

O texto acima foi gerado a partir da introdução de alguns parâmetros ao sistema de criação. Como base, foi utilizado todo o corpus disponível de filmes, agregando ao texto trechos que podem corresponder quase que inteiramente a alguns títulos. Outro parâmetro foi a seleção da quantidade de caracteres que o trecho deveria conter. O número escolhido foi de 1000 caracteres, dado que em um roteiro comum, esta

quantidade indica aproximadamente um minuto de fala. Após a criação, o texto foi traduzido livremente para o português, aumentando sua quantidade de caracteres. Apenas para a criação, desconsiderando o treinamento, o roteiro foi escrito em dois minutos. Sua formatação após esse processo, dada as etapas citadas anteriormente, teve como objetivo dar apenas legibilidade ao conteúdo.

3.2 - Opacidade criativa

O texto inicia-se com a marcação de um cabeçalho de cena. A escolha de Los Angeles pelo sistema provavelmente está ligada ao filme *Independence Day* (1996, EUA). O roteiro do filme de ação possui diversas indicações de cabeçalhos que apontam para a cidade de Los Angeles, o que permitiu ao sistema verificar esta recorrência e selecionar uma para compor o texto. A escolha de um local para compor uma cena por um roteirista é, antes de tudo, uma projeção de espaço e a forma como o sistema faz esta escolha apenas pelo cálculo de recorrências do *corpus* indica a aleatoriedade da seleção semântica. O WRT, assim como o Benjamin, é um sistema preparado para encontrar a próxima palavra ou frase baseado em um *corpus* e sua meta semântica se baseia inteiramente neste princípio. Se o sistema consegue adicionar ao conjunto uma conclusão com um índice alto no cálculo de recorrências então a operação pode ser dada como positiva. O fator da aleatoriedade na escolha para um início nos remete a discussão do conceito de *black box*, traduzido para o português como *caixa opaca*.

However, although these models [neural networks] reach impressive prediction accuracies, their nested non-linear structure makes them highly non-transparent, i.e., it is not clear what information in the input data makes them actually arrive at their decisions. Therefore these models are typically regarded as black boxes⁴⁵. (SAMEK, WIEGAND e MÜLLER, 2007)

⁴⁵ Tradução nossa: No entanto, embora esses modelos [de redes neurais] conseguem precisões impressionantes em previsões, eles aninham-se com uma estrutura não linear que os faz não-transparentes. Isto é, não são informações claras na entrada que fazem elas chegar em uma decisão. Portanto, esses modelos são considerados caixas opacas.

Este tipo de resultado é uma preocupação entre diversos pesquisadores da área de IA, principalmente em sistemas que relacionam dados a resultados que podem influenciar diretamente a vida de um cidadão, como, por exemplo, a obtenção de um benefício de auxílio social. Em entrevista ao jornal El País⁴⁶, a pesquisadora Kate Crawford da Universidade de Nova York e diretora de pesquisa em IA da Microsoft, afirma que sistemas, como as redes neurais, podem ser alvos de uma marginalização humana, sendo treinados e programados a partir de dados que já apresentam algum tipo de injustiça internalizada. O Google Fotos está entre os exemplos de aplicações que podem ter sido treinadas de forma a marginalizar algumas pessoas. Ao fazer a identificação facial e separar pessoas e animais em grupos de fotos, o aplicativo agrupava no mesmo conjunto pessoas negras e chipanzés. Este erro do aplicativo foi identificado em junho de 2015 e foi atribuído ao treinamento do sistema com poucas fotos de pessoas negras. Foi preciso então uma correção de dados do treinamento do sistema, permitindo assim uma identificação e classificação mais concisa das fotos.

Assim como os sistemas que CRAWFORD (2018) cita que não possuem essa transparência de como funcionam, o WRT também não pode ser explicado. Sua configuração descreve as etapas de construção, mas seu treinamento é baseado em diversos outros fatores que não permitem a visualização de uma camada opaca entre os dados de entrada e a saída concebida. Dessa forma, ao utilizarmos o sistema para a construção de um texto, estamos processando uma tentativa de construção de um sentido por uma aplicação em que sua própria forma de funcionamento não está em um nível semântico atingível. O resultado é a expressão de um conjunto de atribuições computacionais expresso em um suporte humano, no caso, o texto escrito.

SAMEK, WIEGAND e KLAUS (2017, p. 2) citam a frase do campeão mundial de jogo chinês Go, San Hui, ao perder em uma partida para o sistema AlphaGo, desenvolvido pela Google: “Não é uma jogada humana. Eu nunca vi um humano fazer

⁴⁶ Disponível em https://brasil.elpais.com/brasil/2018/06/19/actualidad/1529412066_076564.html. Acesso em 20 jun. 2019

este movimento. ” Go é um jogo de tabuleiro de estratégia jogado entre dois adversários. O objetivo do jogo é conseguir cercar com peças pretas e brancas a maior área possível do tabuleiro. Estima-se que o número de possibilidade de movimentos de Go é de 10^{171} , enquanto no xadrez o máximo é de cerca de 10^{50} . Estimasse que número de possibilidades de jogadas em Go supera a quantidade de átomos no universo. Com uma quantidade de combinações impossíveis de se calcular, o supercomputador AlphaGo foi desenvolvido a partir dos dados de outras partidas e a análise de como os jogadores criam suas estratégias. Durante uma das cinco disputas entre o AlphaGo e o campeão mundial do jogo, o coreano San Hui, o computador realizou uma jogada que parecia um erro de iniciante. A probabilidade calculada pelo próprio computador de um humano realizar uma jogada igual era de uma em 10.000. No vigésimo movimento após a jogada do AlphaGo foi identificado a área de controle no tabuleiro que resultou na derrota do coreano. Avaliada pelo time da Google, a jogada não pode ser explicada de forma completa, pois a camada opaca impede o total entendimento humano de como o computador chegou a determinada conclusão.

A partir deste exemplo podemos dialogar com o conceito de *throughput* sugerido por PRIMO (2000). De acordo com PRIMO (2000, p. 9), em um sistema programado, todos os resultados obtidos serão dados a partir de um “reflexo ou automatismo”, de forma que “não há uma seleção consciente, verdadeiramente inteligente, do computador. O computador oferece uma falsa aparência interpretativa”. Podemos abordar a forma como as redes neurais de recorrências agem para a criação de um *output* de maneira complementar ao conceito proposto por PRIMO (2000). Uma rede neural aprende e gera saídas a partir a partir de verificações de padrões. Embora está seja uma interação mecânica, existe uma interpretação dos dados dispostos e da forma como o sistema foi programado. Podemos relacionar esta avaliação dos dados disponíveis com o *throughput* da interação mútua em que “cada interpretação se dá pelo confronto da mensagem recebida com a complexidade cognitiva do interagente” (PRIMO, p. 9). Assim como o AlphaGo, redes neurais podem levar em consideração um elevado número de fatores em consideração, antes de tomar uma decisão de

saída, assemelhando-se até mesmo ao processo de negociação. Benjamin ou o WRT não possuem tantas camadas de abstração de processamento como o AlphaGo, mas possibilitaram a visualização de saídas baseados em processos que não são atingíveis pela compreensão humana.

Partindo do conceito de que a falta de explicação de como um sistema inteligente funciona pode prejudicar sua aplicabilidade, a DARPA (Agência de Projetos de Pesquisa Avançada de Defesa dos Estados Unidos) iniciou o projeto global XAI⁴⁷ (Explicando a Inteligência Artificial). O projeto tem como objetivo desenvolver novas formas de aplicar as atuais técnicas de aprendizado, visando uma melhor compreensão humana sobre a camada opaca do sistema. Para isto, diversas universidades e institutos de pesquisas devem desenvolver formas mais claras, a maioria baseada em exemplos palpáveis, de como uma IA determina algo e como faz para determinar.

Fonte: Reprodução / DARPA - XAI

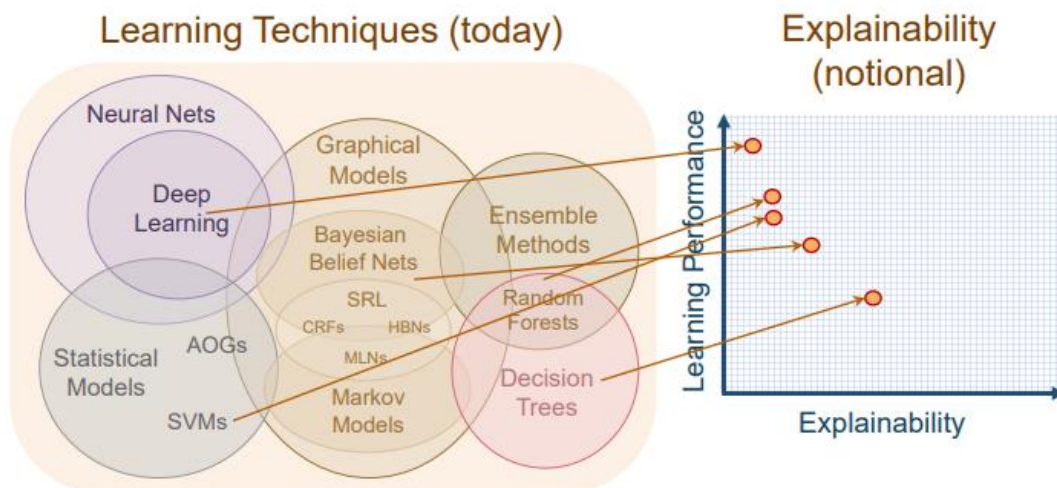


Figura 23: Gráfico de explicabilidade por performance de aprendizagem

⁴⁷ Disponível em: <https://www.darpa.mil/attachments/XAIProgramUpdate.pdf> Acesso: em 20 jun. 2019

Orange é o resultado de um conjunto de regras e aplicações opacas ao entendimento humano, de forma que se assemelhe a completa aleatoriedade. Assim como analisamos em Sunspring, o curta-metragem é uma tentativa de dar sentido a um texto gerado de uma forma diferente do padrão. Sunspring, assim como Orange, projeta diretamente em seu nível discursivo as instâncias de pessoa, tempo e espaço de forma semelhante a aleatória. A combinação de recorrências pelos sistemas projeta sobre o curta-metragem algumas instâncias que são contraditórias. Um exemplo disto está na primeira descrição de ambiente e ação do Orange:

Vemos a sala vazia. Em uma rocha estreita, vemos J.

Ele está se segurando sobre uma mesa.

O trecho “vemos a sala vazia” provavelmente foi selecionado por vir após um cabeçalho. A descrição do ambiente normalmente é encontrada em roteiros logo após o cabeçalho da cena e antes da descrição de ação. A frase, que pode ter sido classificada de forma que articula com outra, é seguida de “em uma rocha estreita, vemos J”. O caráter confuso do trecho continua em “ele está se segurando sobre uma mesa”. A contradição do trecho, iniciada já no cabeçalho que descreve como ambiente um terraço externo, está entre os pontos de demonstração da aleatoriedade presente no sistema. O ambiente externo passou a ser uma sala. A sala vazia passou a ter uma mesa e uma rocha e J está nos dois lugares ao mesmo tempo. As características confusas do trecho remetem as descrições de cenas de Sunspring, em que um personagem ocupa dois espaços ao mesmo tempo. A criação de três frases consecutivas que se contrapõe demonstra o caráter apenas gerativo do WRT e do Benjamin. Ambos os sistemas são programados para processar apenas a próxima palavra, sem qualquer garantia de fundamentação de uma narrativa. As palavras escolhidas influenciam na seleção das próximas, no entanto, estas não influenciam na concepção do texto como um todo de sentido, apenas de construção de recorrências quantitativas. A criação de uma obra com estrutura teleológica fica

comprometida com as atuais configurações do WRT e Benjamin, em que não existem um armazenamento de recorrências qualitativas ou um cálculo sobre todas as possibilidades de ações entre os personagens. Utilizamos o WRT, assim como Sharp e Goodwin utilizaram o Benjamin, para a construção de um objeto que extrapola suas definições, de forma que, ao projetar uma frase ambos os sistemas são bem-sucedidos, mas ao conecta-la com a próxima o resultado é falho. Podemos fazer uma analogia da forma como estes sistemas criam seus conteúdos com a própria história dos estudos linguísticos, em que, anteriormente, a frase era dada como unidade de sentido máxima, independente sua semântica do texto (DIANA, 2005). Por isso, durante a leitura, podemos verificar diversos pontos de confusão, já que as frases estão desconexas um conjunto semântico.

A relação de sequencialidade dos dois sistemas é dada apenas pelo preenchimento de parâmetros que incidem sobre uma meta semântica, como o cálculo recorrências que definirá qual palavra será a próxima alocada. Como citado no primeiro capítulo, outros sistemas já foram programados com o objetivo de criar uma narrativa, como o mexica-IMPRO. A aplicação funcionava de forma diferente de Benjamin e o WRT, já que a composição dos textos funcionava inteiramente a partir do uso de tabelas que regiam as ações e vontades dos personagens, bem como as relações entre eles. Diferente dos sistemas que utilizam redes neurais, o mexica-IMPRO não possuía qualquer tipo de treinamento a partir da entrada de dados, por isso, recebia críticas de seus resultados eram muitas vezes repetitivos, já que as possibilidades de combinações eram muito menores do que em sistemas como o Benjamin.

O não armazenamento das recorrências semânticas dos conteúdos gerados pelos sistemas influi sobre um problema semiótico: as modalizações. De forma randomizada, o sistema define o nome de um personagem e atribui a este uma ação ou descrição. Esta composição aleatória modaliza o personagem, no entanto, em um suporte feito inteiramente por códigos, esta modalização não fica armazenada e suas influências são brevemente descartadas na próxima composição frasal. Esta falta de

influência das modalizações nas ações dos personagens também pode ser observada na tentativa de correção humana do texto. A escolha de poucos personagens e a seleção de falas para estes reforça a tentativa de formação do sentido, adequando o texto original a estruturas padrões de uma narrativa.

A primeira fala de Orange é do personagem L, que originalmente é Leia, general e princesa da franquia Star Wars. Somente através desta indicação podemos definir que L é uma mulher, no entanto, não existem outras indicações sobre o gênero ou qualquer característica sobre o personagem. Novamente reforçando o caráter de aleatoriedade, qualquer descrição que viesse a existir seria puramente randômica e em forma de uma coincidência. A fala *you will sleep here, but I said it will be sympathetic. All thanks to the home of four coins.* foi traduzida, originalmente, de *you will sleep here, but I said it will be sympathetic. All thanks to the home of four coins.* Nenhuma das orações do trecho original está presente em um dos roteiros, o que comprova a forma construtiva do WRT. Assim como exemplificamos, o sistema funciona de forma semelhante a um teclado de *smartphone* na sugestão de novos termos. No entanto, as diferenças entre a aplicação e o serviço de auxílio a digitação são fundamentais para exemplificar como estas frases são geradas.

Enquanto um teclado de *smartphone* seleciona seu *corpus* a partir das entradas realizadas pelos usuários, sempre se atualizando com novas palavras, o WRT mantém um *corpus* estático a partir dos roteiros injetados. Assim, com um *corpus* de textos formados e não apenas de palavras, o WRT consegue sugerir em suas criações, com maior êxito, o uso de pontuações das frases. Ao mesmo tempo que isto pode ser um elemento positivo para o sistema, é preciso reforçar que este realiza estas ações a partir de uma reprodução de padrões identificados. Este comportamento de reprodutibilidade de configurações treinadas permite que o WRT não precise de qualquer suporte de uma biblioteca de processamento de linguagem natural, ampliando a gama de idiomas que o sistema suporta. Já os teclados de *smartphones* não possuem inicialmente um *corpus* de conteúdo tão vasto quanto o WRT, precisando assim, de configurações mais específicas para poderem fazer as

primeiras sugestões de palavras com pouquíssimo uso de textos do usuário. Baseados em bibliotecas de linguagem natural próprias de cada idioma, estes *softwares* possuem uma grande vantagem sobre sistemas como o WRT na construção baseada em regras gramaticais. Durante a construção do roteiro de Orange não foram identificados erros de sintaxe grosseiros nos textos em inglês. Além disso, pequenos erros resultantes da combinação errada das palavras foram removidos durante a tradução para o português.

O ponto de destaque sobre a elaboração sintática das frases fica em relação ao WRT aprender a ler e escrever todas as vezes em que seu banco de dados é zerado e iniciado um novo treinamento. Dessa forma, podemos entender as frases geradas como a simples reprodução de padrões e não como a articulação de palavras em um conjunto sintático. O uso de uma biblioteca de linguagem natural poderia auxiliar o sistema nesta articulação, estabelecendo as regras de sintaxe para um nível menos superficial da construção de novas frases. Assim, a geração a partir do treinamento do *corpus* poderia ser relacionada ao preenchimento destas estruturas sintáticas já estabelecidas, formando assim as primeiras nuances de produção semântica.

3.3 - Gerativismo da semântica

A criação autônoma de texto com o Benjamin e o WRT funciona como uma conversão de padrões de dados quantitativos que selecionam palavras em banco de dados e as colocam em sequência em uma frase. Conforme vimos, este padrão pode ser alterado de acordo com os dados de entrada, mas seu processamento acontece em um alto nível de complexidade que, muitas vezes, não é compreendido pelo ser humano. Podemos fazer uma analogia das etapas do processamento do *corpus* e a construção do texto com a própria divisão de níveis do percurso gerativo do sentido: no nível mais profundo e abstrato está o processamento das redes neurais com um baixo grau de explicabilidade. Em um nível mais superficial está o texto em sua forma

mais concreta. A criação autônoma percorre do nível mais baixo para o mais alto, levando para a superfície o resultado do processamento. A partir desta configuração, podemos avaliar a construção de um nível intermediário entre a abstração da construção e a concretude do texto. Este nível de intermédio poderá atuar como um construtor semântico, de forma que deve avaliar sempre os resultados de suas entradas para garantir uma saída aceitável.

Fonte: Montagem do autor

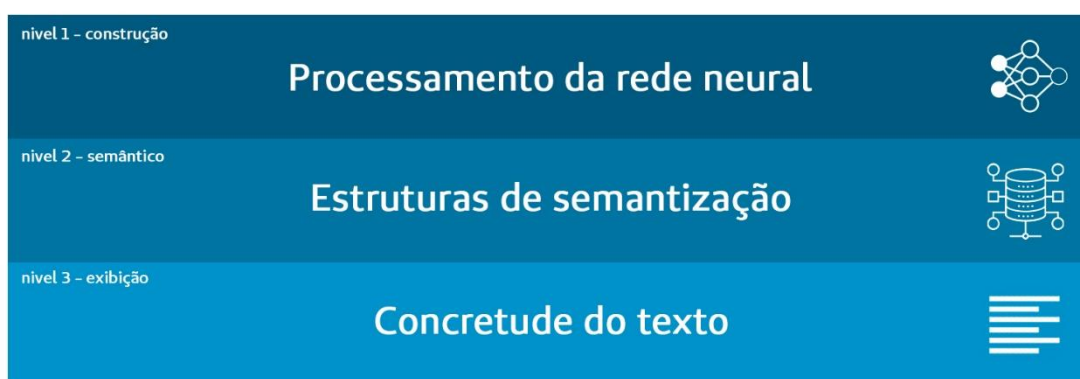


Figura 20 – Nível intermediário proposto

O nível intermediário proposto deve atuar como uma forma de tradução entre os níveis de construção e exibição, como apresentado na figura. Ao organizarmos esta estrutura como um nível de concepção semântica, ela pode agir como um intermediador da compreensão, de forma que tenta selecionar e organizar os elementos gerados pela rede neural. Dessa forma, a camada intermediária deve limitar os resultados gerados pela IA. Podemos fazer uma comparação desta restrição dos conteúdos produzidos com algumas críticas a camada limitadora do Mexica-impro. Enquanto a aplicação de Pérez y Pérez limitava as ações de seus personagens a partir de variáveis definidas anteriormente ao início do processo de produção do texto, nossa proposta utiliza o percurso gerativo de sentido após o início do processamento da rede neural. Desta forma, o aspecto da *criatividade* da máquina, demonstrada pelo caráter de produção de texto abstrato a partir das recorrências, é mantido.

A manutenção desta característica ressalta a importância da comunicação entre as camadas, não apenas de forma linear, mas dialógica. Desta forma, a camada intermediária atua como controladora de todo o sistema. Após a primeira execução, o nível intermediário deve acompanhar o nível de exibição da narrativa em busca de necessidades e requisitar do nível de construção novas gerações textuais a partir de parâmetros verificados como faltantes. Assim como um roteirista humano que utiliza o WRT para gerar novos trechos conforme sua necessidade, a camada intermediária deve compilar todas estas tarefas. O sistema ganha complexidade à medida que tenta simular esta seleção humana, criando parâmetros de criação e singularizando diversas execuções de criações na rede neural para compor um único texto.

O nível intermediário deve preencher requisitos para poder obter sucesso, no entanto, este conceito verifica-se de forma tênue, de modo que nem todos os critérios devem se configurar de forma verdadeira para a obtenção de um texto com sentido. Por isso, os diversos elementos do percurso gerativo do sentido que podem ser verificados não devem ser vistos como os únicos parâmetros para a obtenção do sucesso. Matematizar uma metodologia humana, como a semiótica e o percurso gerativo do sentido, pode corresponder a sua completa inutilização do sistema. A indução em teorias humanas é um processo complexo e, em muitos casos, sem muitas variáveis exatas. Dessa forma, nem todos os resultados previstos para uma narrativa gerada por um nível que almeja a simulação humana serão de fato aceitáveis para seu propósito. Assim como no nível de construção, o intermediário também exigirá um treinamento e um processo de identificação humana de quais resultados são aceitáveis ou não, certificando ao sistema boas práticas da utilização dos recursos semânticos para compor uma narrativa.

O treinamento do sistema pode permitir a identificação de classes gramaticais e, por consequência, a sintaxe narrativa. O reconhecimento dos sujeitos e suas ações abre espaço para uma gama de instâncias semióticas que podem ser avaliadas, como a geração de enunciados de estado e de fazer. Estas unidades essenciais compõem um programa narrativo, que, ao avançar são partes um percurso narrativo e o

esquema canônico. O encadeamento destes elementos oferece uma gama de requisitos a serem preenchidos de diversas formas, além da possibilidade de um processamento não linear da narrativa. O sistema pode identificar em um bloco de texto gerando elementos narrativos quase completos, com partes médias faltantes. Assim, recorrendo ao motor de criação da camada inferior, novos trechos podem ser gerados e identificados conforme as necessidades do texto. Esta não linearidade traz ao processo de criação autônoma semelhanças com o processo de edição de vídeo, em que os trechos gravados são dispostos em uma linha tempo e selecionados para compor seu espaço no arranjo final.

Fonte: Montagem do autor

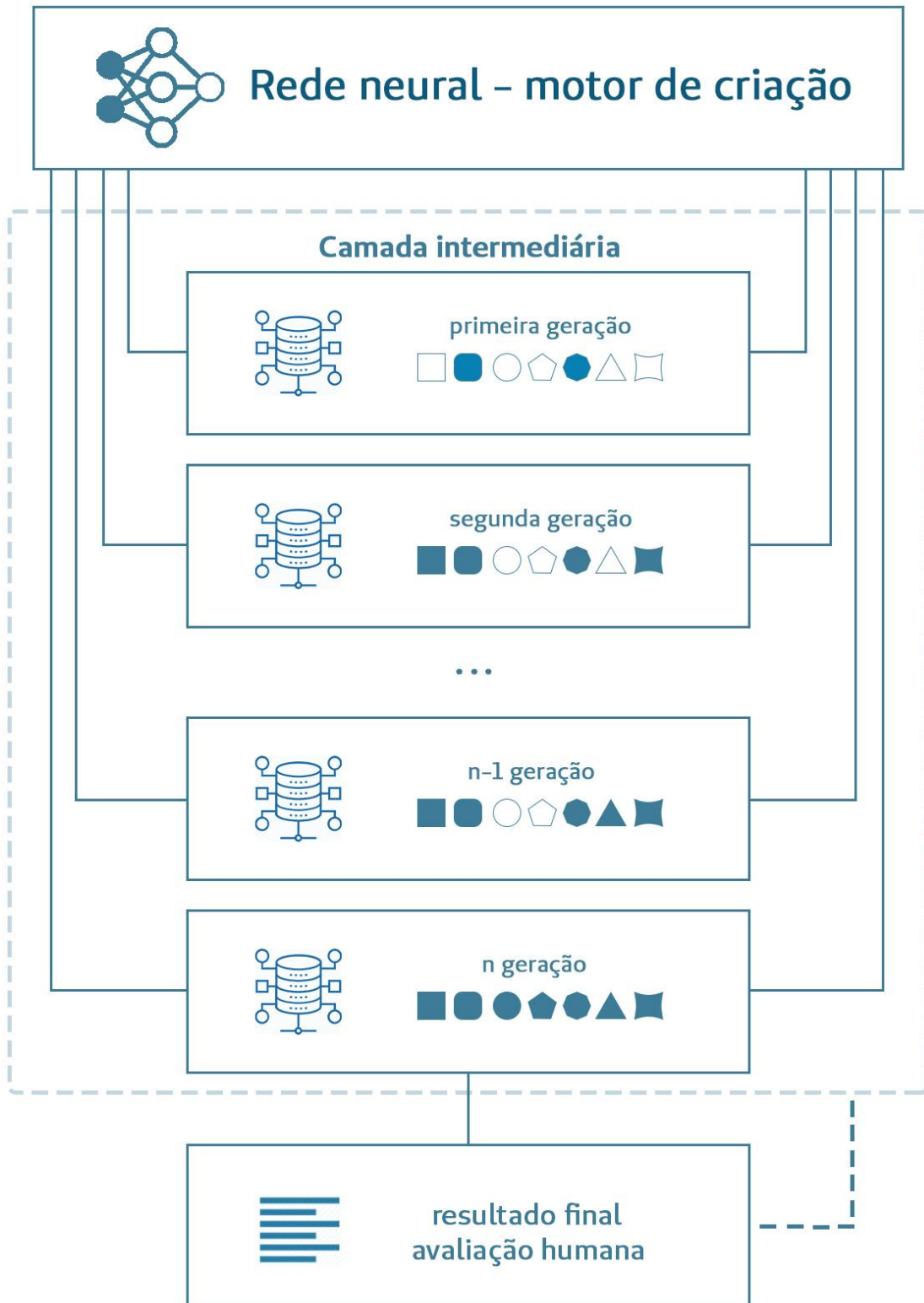


Figura 25: Funcionamento da camada intermediária

As projeções pessoa, espaço e tempo estão na sintaxe do nível discursivo do percurso gerativo do sentido. Estas projeções são fundamentais para o processo de enunciação e podem servir como guia da criação de textos autônomos. A partir da identificação de indícios nos textos iniciais gerados pela rede neural, o sistema pode reconhecer e indicar de volta a construção, agora como um parâmetro, que novos trechos sejam criados a partir de uma projeção em particular. As projeções de pessoa, espaço e tempo também podem servir como elementos balizadores do sistema, com a constante verificação para que não exista a quebra destes pontos. Como exemplo, podemos citar o trecho do primeiro cabeçalho de cena de Orange, em que projeções de espaço diferentes se contrapõem, causando confusão à leitura do texto. Outros pontos de destaque para a programação das projeções estão relacionados a criação e posicionamento dos personagens na narrativa. Com um gigantesco acervo de substantivos próprios dado pelo *corpus*, o WRT constantemente gera um novo personagem. A partir de um completo entendimento de quem são os personagens, e onde eles estão posicionados, o sistema pode parametrizar melhor suas criações.

Os preenchimentos das lacunas sintáticas dependem de diversos fatores a serem analisados nos trechos criados e entre eles está o processo de modalização. Componente da semântica narrativa, a modalização define, antes de tudo, as vontades dos sujeitos. Ela é parte fundamental do sentido do texto e da forma como as transformações serão operadas dentro da narrativa. Assim, o sujeito modalizado será, muitas vezes, o elemento de ligação entre as partes de texto selecionadas, transgredindo o trecho sintático, e mais concreto, para um patamar semântico e abstrato. Esta concretização de um elemento semântico como forma de ligação entre trechos gerados talvez esteja entre as partes mais complexas de se matematizar de todo o sistema proposto. No entanto, é fundamental a sua identificação para a concepção de estruturas de manipulação dos sujeitos. Como em uma fórmula matemática, o índice deve indicar se um sujeito da manipulação está propenso e aceitar ou não um contrato com o sujeito manipulador. Esta forma de exatidão ao determinar uma ação remete a lógica multivalorada (*fuzzy*), de forma que diversos

fatores devem influenciar uma decisão e com diversos pesos. Ainda assim, o cálculo não é exato, já que diversos fatores humanos e criativos devem ser postos.

Outro elemento que deve fazer a passagem do âmbito semântico para atuar diretamente na concretude da narrativa é o de identificação de temas, disposto no nível da semântica discursiva. Segundo DIANA (2005, p. 66) “ tematizar um discurso é formular os valores de modo abstrato e organizá-los em percursos. Em outras palavras, os percursos são constituídos pela recorrência de traços semânticos ou semas, concebidos abstratamente. ” Para a criação de textos com o sistema proposto, o processo de tematização deve ser emulado diretamente na seleção dos elementos. A abstração dos percursos temáticos ganha forma com a figurativização, que recobre com traços sensoriais os percursos. Esta figurativização é expressa em roteiros na maioria das vezes em descrições de ambientes e ações, tornando-se para o sistema uma informação complementar da narrativa criada e esquematizada no nível intermediário. O processo de figurativização durante a etapa de construção e adequação narrativa pela camada intermediária exigem que o sistema possa processar indícios de intenções de temas, para assim, poder recobrir figurativamente estes. Esta etapa exige um constante treinamento do sistema em entender contextos e emoções expressas nos blocos de textos gerados, além de um banco de recursos figurativos que possam ser adequados as temáticas encontradas.

A camada intermediária proposta deve desempenhar uma função mais complexa do que a de simplesmente filtrar os trechos gerados automaticamente de texto. Sua função é de organizar sintaxe e semântica da narrativa seguindo a própria divisão do percurso gerativo do sentido. Sua programação envolve diversas variáveis que dão valores quantitativos para elementos abstratos, arriscando o sucesso semântico do texto. A proposta não se baseia em dar valores matemáticos para elementos da compreensão humana, mas em atribuir uma metodologia de seleção sobre a criação de textos autônomos gerados por uma rede neural com uma camada opaca de entendimento. Os trechos de texto, ao serem gerados, passam por uma camada que tenta avaliar intenções narrativas em frentes sintáticas e semânticas,

atribuindo valores para elementos abstratos que se conectam a elementos concretos. A recursividade da camada permite a chamada do motor de criação para a obtenção de novos trechos parametrizados a partir das necessidades identificadas, permitindo assim, ao sistema, a criação de narrativas de forma não linear.

Além dos elementos propostos neste capítulo, outras alterações na atual configuração de sistemas como o Benjamin e o WRT podem auxiliar na construção de textos com um nível semântico mais aprimorado. O uso de um conjunto menor de textos de exemplo no *corpus* para o treinamento age de forma contrária ao recomendado pelos próprios sistemas de IA, no entanto, este conjunto menor pode possuir mais recorrências e ligações, dando menos possibilidades ao sistema. Dados mais claros e limpos relacionados ao objeto pretendido permitem que o sistema crie de forma limitada. Além disso, a parametrização do sistema também pode auxiliar na forma como os textos serão gerados, indicando palavras chaves para serem abordadas ou trechos inteiros que devem estar contidos.

CONSIDERAÇÕES FINAIS

CONSIDERAÇÕES FINAIS

Sistemas baseados em inteligência artificial possuem uma presença cada vez maior em diversas áreas do cotidiano humano. A mineração de dados em busca novos *insights* representa um mercado bilionário e o modelo de negócio de diversas empresas ligadas a um novo patamar do ecossistema da informação. Não há dúvidas de que avanços em sistemas inteligentes representam as próximas etapas da tecnologia, permitindo grandes conquistas para diversas áreas. Para o pleno avanço da tecnologia, cada vez mais surgem preocupações relacionadas a qualidade dos dados que treinam estes sistemas e a forma como este treinamento acontece.

No audiovisual, sistemas inteligentes inauguram uma nova metodologia de trabalho, podendo auxiliar em diversas tarefas técnicas relacionadas principalmente ao processo de pós-produção. Experimentos como o Benjamin e o WRT, ao emularem os processos criativos de concepção de um roteiro, rompem com a lógica de produção realizando tarefas que antes pertenciam apenas a humanos. Sistemas geradores de textos autônomos ainda estão longe de substituírem a inventividade humana em papéis tão sensíveis, como compor uma narrativa, mas podem ser utilizados como ferramentas de auxílio para diversos profissionais que trabalham diretamente com a criatividade.

O roteiro do curta-metragem Sunspring, escrito por Benjamin, apresenta-se como a primeira narrativa escrita inteiramente por uma inteligência artificial, no entanto, o curta foi organizado por um humano e diversas características ligadas ao sentido, como a projeção de pessoa e espaço, foram adaptadas. Além disso, o sincretismo audiovisual tem grande papel na semântica de Sunspring, projetando seu

sentido através das outras linguagens, como a trilha sonora, atuação e fotografia. Sunspring é a tentativa concreta de dar sentido um texto escrito por uma inteligência artificial.

Benjamin e o WRT apresentam, muitas vezes, criações confusas, que nos fazem questionar o caráter de criativo dos sistemas, no entanto, estes sistemas trabalham suas composições apenas a partir do encontro de recorrências dos conteúdos de seus bancos de dados. Assim, o sistema gera a partir de uma meta semântica de escopo fechada ao processo criação. Esta meta semântica faz sentido a forma quantitativa que o sistema funciona e organiza as palavras em uma composição. Este processo de tomada de decisões em uma rede neural possui um baixo nível de *explicabilidade* e por isso é definido como uma camada opaca.

Nossa proposta de camada intermediadora entre o texto e sua geração autônoma foi a de adicionar elementos parametrizados para a criação, desenvolvendo assim uma nova rede de requisitos, baseados no percurso gerativo de sentido, que devem chamar o motor de busca para seu devido preenchimento. O sistema proposto se apresenta como um desafio, já que este deve qualificar e organizar elementos concretos conectados uns aos outros a partir de elementos semânticos abstratos. Esta nova forma de criação de textos é uma tentativa de emular a seleção humana de trechos gerados pelo computador em uma linearidade e seu sucesso depende de diversos fatores impalpáveis, mas que precisam ser matematizados, como os processos de modalização dos sujeitos. Além disso, o próprio sistema deve estar sob avaliação humana dos resultados, podendo moldar a forma como este seleciona os elementos em busca de um resultado mais aceitável.

A partir das análises do curta-metragem Sunspring e do modo de criação autônoma de Orange, pelo WRT, pode ser identificada como sensível a tarefa de construção e manutenção do sentido nos textos. Como sugestão, esta pesquisa pode expandir-se para estudos mais profundos da semântica, enunciação e das formas de geração de texto. Outro viés que pode ser abordado em pesquisas futuras é o da

construção de níveis intermediários. A partir da implementação de bibliotecas de linguagem natural que possuam um sentido absoluto para as palavras, ou seja, que não funcionem apenas das recorrências do *corpus*, um novo texto pode ser criado a partir de uma parametrização.

Embora a tecnologia utilizada para a geração de texto autônomo esteja em um alto patamar técnico, seus resultados ainda precisam ser aprimorados principalmente no campo semântico. A tecnologia gerativa não é, em seu estágio atual, uma ameaça as profissões criativas. Os diversos modos de se fazer comunicação e arte são pautados na linguagem e no sentido, por isso, o melhor uso destas ferramentas é o de auxílio, assistindo diversos profissionais de criação, agilizando suas formas de escrita e ajudando no desenvolvimento de novas ideias.

REFERÊNCIAS BIBLIOGRÁFICAS

REFERÊNCIAS BIBLIOGRÁFICAS

ALPHAGO. Greg Kohs. Rússia: Moxie Pictures, 2018. Distribuição digital sob demanda.

AMARAL, Fernando. **Introdução à ciência de dados**: mineração de dados e big data. Rio de Janeiro: Alta Books, 2016.

AUGUSTO, Marina, R. A. **Teoria gerativa e aquisição da linguagem**. Sitientibus. Feira de Santana, n.13, p. 115-120, jul/dez, 1995

BARROS, Diana Luz Pessoa de. **Teoria semiótica do texto**. Ed. 4. São Paulo: Atíca, 2005.

BERTRAND, Denis. **Caminhos da semiótica literária**. Tradução do Grupo CASA. Bauru: EDUSC, 2003.

CASTELLS, Manuel. **A galáxia da internet**: reflexões sobre a internet, os negócios e a sociedade. Tradução de Maria Luiza X. de A. Borges. Rio de Janeiro: Jorge Zahar Ed., 2003.

EMC Digital Universe. **The Digital Universe of Opportunities**: Rich data & the increasing value of the internet of things. 2014, disponível em: <<https://uk.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>> acesso em: 21 jan. 2019.

FERNEDA, Edberto. **Redes neurais e sua aplicação em sistemas de recuperação de informação**. Ciência da Informação. Brasília, v. 35, n. 1, 25-30, jan/abr. 2006

FIORIN, José Luiz. **A noção de texto na semiótica**. São Paulo: Contexto, 1999.

_____. **Elementos de análise do discurso**. São Paulo: Contexto, 2002.

FLOCH, Jean-Marie. **Alguns conceitos fundamentais em semiótica geral**. Tradução Analice Dutra Pilar. In: Documentos de estudos do centro de pesquisas sociosemióticas. São Paulo. 2001.

FONSECA FILHO, Clézio. **História da computação: O caminho do pensamento e da tecnologia**. Porto Alegre: EDIPUCRS, 2007.

GOODWIN, Ross; SHARP, Oscar. **Machines Making Movies**. TEDxBoston 2017. Disponível em https://www.youtube.com/watch?v=uPXPQK83Z_Y Acesso em 05 de mar. 2019.

_____. **Machines Making Movies**. GitHub Universe 2016. Disponível em <https://www.youtube.com/watch?v=W0bVxyi38Bc&t=1460s> Acesso em 20 jun. 2019

GREIMAS, Algirdas Julien. **Da imperfeição**. Tradução de Ana Claudia de Oliveira. São Paulo: Hacker Editores, 2002.

_____; COURTÉS, Joseph. **Dicionário de semiótica**. Tradução de Alceu Dias Lima et al. São Paulo: Contexto, 2011.

GUDWIN, Ricardo; QUEIROZ, João. **Semiotics and Intelligent Systems Development**. Hershey: Idea Group Publishing, 2007.

GUARDIAN, The. **Artificial intelligence: 'We're like children playing with a bomb'**. 2016, disponível em <<https://www.theguardian.com/technology/2016/jun/12/nick-bostrom-artificial-intelligence-machine>> acesso em: 20 jan. 2019.

GUNNING, David. **Explainable Artificial Intelligence (XAI)**. Disponível em <https://www.darpa.mil/attachments/XAIProgramUpdate.pdf> Acesso em 20 jun. 2019

GUNKEL, David J. **Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação**. Galaxia (São Paulo, online), n. 34, jan-abr., 2017.

IBM. **Scaling Wimbledon's video production of highlight reels through AI technology**. 2017, disponível em: < <https://www.ibm.com/blogs/research/2017/06/scaling-wimbledons-video-production-highlight-reels-ai-technology/>> acesso em: 20 jan. 2019.

JENKINS, Henry. **Convergence Culture: Where Old and New Media Collide**. Nova York: NYU Press, 2008.

JOHNSON, Steven. **Cultura da interface: como o computador transforma nossa maneira de criar e comunicar**. Tradução de Maria Luiza X. de A. Borges. Rio de Janeiro: Jorge Zahar Ed., 1997.

LIVESCIENCE. **Artificial Intelligence: Friendly or Frightening?** 2014. disponível em: < <https://www.livescience.com/49009-future-of-artificial-intelligence.html>> acesso em: 20 jan. 2019.

MCCARTHY, John. et al. **A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence**. AI Magazine, v. 27, n. 4, jan/mar, 2006.

MCKEOWN, Lewis. **An evaluation of the Impact of Constraints on the Perceived Creativity of Narrative Generating Software**. University of Kent. Reino Unido. 2017.

MÉDOLA, Ana Silvia Lopes Davi. **Televisão digital brasileira e os novos processos de produção de conteúdos: os desafios para o comunicador**. E-Compós, v.12 n.3, set/dez, 2009.

NORVING, Peter; RUSSEL, Stuart. **Inteligência artificial**. Tradução de Regina Célia Simille de Macedo. 3ª ed. Rio de Janeiro: Elsevier, 2013.

OKA, T. Moto. (org) **Fifth generation computer systems**. Amsterdam: Elsevier Science Publishers B. V., 1982.

PAIS, El. **Kate Crawford**: "Estamos injetando nos algoritmos as nossas limitações, a nossa forma de marginalizar". 2018. Disponível em https://brasil.elpais.com/brasil/2018/06/19/actualidad/1529412066_076564.html Acesso em 20 jun. 2019

PRIMO, Alex. **Conhecimento e interação**: fronteiras entre o agir humano e inteligência artificial. In: LEMOS, A.; P. CUNHA (Eds.). Olhares sobre a cibercultura. Porto Alegre: Sulina, 2003. p. 37-56.

_____. **Interação mediada por computador**: comunicação, cibercultura, cognição. Porto Alegre: Sulina, 2007.

SAMEK, Wojciech; WIEGAND, Thomas; MÜLLER, Klaus-Robert. **Explainable artificial intelligence**: understanding, visualizing and interpreting deep learning models. Genebra. ITU Journal: ICT Discoveries, Special Issue No. 1, 13 Oct. 2017

SMITH, John. R. **IBM Research Takes Watson to Hollywood with the First “Cognitive Movie Trailer”**. 2016, disponível em: <
<https://www.ibm.com/blogs/think/2016/08/cognitive-movie-trailer/>> acesso em: 20 jan. 2019.

TURING, Alan Marthison. **Computing machinery and intelligence**. Mind. V. 49, 433-460. 1950.

WAZLAWICK, Raul Sidnei. **História da computação**. 1 ed. Rio de Janeiro: Elsevier, 2016.

APÊNDICES

Roteiro em português do curta-metragem Sunspring.

Nota: O roteiro foi traduzido a partir do original disponibilizado pelos autores e sua diagramação foi alterada com o propósito de melhor leitura

INT. NAVE

Vemos H puxar um livro de uma prateleira, folheá-lo enquanto fala e depois colocá-lo de volta.

H

Num futuro com desemprego em massa, os jovens são forçados a vender sangue. Essa é a primeira coisa que posso fazer

H2

Você deveria ver os meninos e calar a boca. Eu era a pessoa que teria cem anos.

H

Eu o vi novamente. A maneira como você foi enviado para mim ... foi uma grande ideia honesta. Eu não sou uma luz brilhante.

C

Bem, eu tenho que ir ao crânio. Eu não sei.

Ele pega uma tela de luz e luta contra a força de segurança das partículas de uma transmissão em seu rosto.

H

(continuando)

O que você quer dizer?

C

(sorrindo)

Eu não sei nada sobre isso.

H

(para Hawk, tirando os olhos da boca)

Então o que?

H2

Não há resposta.

C

(franzido)

Nós vamos ver o dinheiro.

H

(lendo)

"Tudo bem, você não pode me dizer isso."

Passos para trás. Coffey ainda está passando.

C

Eu estava chegando nessa coisa porque você era tão bonita.

H

Eu não sei. Não sei do que você está falando.

C

Está certo.

H

Então, o que você está fazendo?

H2

Eu não quero ser honesta com você.

Ele olha para ele por um momento e depois sorri para ele.

H

Você não precisa ser médico.

H2

Não tenho certeza. Não sei do que você está falando.

H

Eu quero ver você também.

H2

O que você quer dizer?

H

Sinto muito, mas tenho certeza que você nem me tocou.

H2

Não sei do que você está falando.

H

O princípio é completamente construído para o mesmo tempo.

H2

(sorrindo)

Era tudo sobre você para ser verdade.

H

Você nem viu o filme com o resto da base.

H2

Eu não sei.

H

Eu não ligo

H2

Eu sei que é uma consequência.

O que quer que você queira saber sobre a presença da história, sou um garoto meio no chão.

H

Eu não sei. Eu só tenho que pedir para você me explicar o que você diz.

H2

O que você quer dizer?

H

Porque eu não sei do que você está falando.

H2

Isso foi o tempo todo.

H

Eu sei disso.

H2

Eu não sei.

H

(bravo)

Seria um bom momento. Eu acho que poderia ter sido minha vida.

Ele começa a tremer.

H (continuação)

Pode nunca ser perdoado, mas isso é muito ruim. Eu tenho que sair, mas não estou livre do mundo.

C

Sim. Talvez eu deva tirar daqui. Eu não vou fazer algo.

H

Você não pode levar isso para qualquer lugar. Não é um sonho. Mas tenho um bom momento para ficar lá.

C

Bem, acho que você ainda pode estar de volta à mesa.

H

Mmm. É uma coisa maldita com medo de dizer. Nada vai ser uma coisa, mas fui eu que caí nessa pedra com uma criança e depois deixei as outras duas.

Ele está parado nas estrelas e sentado no chão. Ele se senta no balcão e puxa a câmera para as costas. Ele a encara. Ele está no telefone. Ele corta a espingarda da borda da sala e a coloca na boca.

Ele vê um buraco negro no chão levando ao homem no telhado.

Ele vem atrás dele para protegê-lo. Ele ainda está de pé ao lado dele. Ele olha pela porta e a porta se fecha. Ele olha para a bolsa da mochila e começa a chorar.

T

Bom, existe essa situação entre mim e a luz na nave. O cara estava tentando me parar. Ele é como um bebê e foi embora. Eu estava preocupada com ele. Mas mesmo se ele tivesse feito tudo. Ele não conseguiria voltar. Eu não pretendo ser uma virgem. Quer dizer, ele era fraco. E eu pensei em mudar de ideia. Ele era louco para ir embora. Foi a muito tempo atrás. Ele estava um pouco atrasado. Eu ia ser um momento. Eu

só queria te contar que eu sou muito melhor que ele. Eu tive de pará-lo e não consegui te contar. Eu não queria machuca-lo. Me desculpe. Eu sei que eu não gosto dele. Eu posso ir para casa e ser tão má e eu o amo. Então eu posso pegá-lo todo o caminho até aqui e encontrar o quadrado e ir para o jogo com ele e ela não vai aparecer. Então eu vou dar uma olhada. Mas eu não vou ver ele quando ele chegar para mim. Ele olha para mim e me joga fora dos olhos. Então ele disse que iria para a cama comigo.