



UNIVERSIDADE FEDERAL DE GOIÁS (UFG)  
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE  
COMPUTAÇÃO (EMC)  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA  
E DE COMPUTAÇÃO (PPGEEC)

ALAN HENRIQUE FERREIRA SILVA

**Desenvolvimento de Métrica de Avaliação Bilíngue por  
Estrutura Frasal aplicada à Tradução Automática**

GOIÂNIA

2022



UNIVERSIDADE FEDERAL DE GOIÁS  
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

### E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

#### 1. Identificação do material bibliográfico

Dissertação       Tese

#### 2. Nome completo do autor

Alan Henrique Ferreira Silva

#### 3. Título do trabalho

**“Desenvolvimento de métrica de avaliação bilíngue por estrutura frasal aplicada à tradução automática”**

#### 4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento  SIM       NÃO<sup>1</sup>

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(à) autor(a) e ao(à) orientador(a);
- b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

**Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.**



Documento assinado eletronicamente por **ALAN HENRIQUE FERREIRA SILVA, Discente**, em 17/02/2022, às 16:04, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 18/02/2022, às 06:49, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **2705462** e o código CRC **96AF8261**.

ALAN HENRIQUE FERREIRA SILVA

**Desenvolvimento de Métrica de Avaliação Bilíngue por  
Estrutura Frasal aplicada à Tradução Automática**

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica e de Computação, da Escola de Engenharia Elétrica, Mecânica e de Computação, da Universidade Federal de Goiás (UFG) como requisito para a obtenção do título de Doutor em Engenharia Elétrica e de Computação.

Área de concentração: Engenharia da Computação  
Orientador: Professor Doutor Wesley Pacheco Calixto

GOIÂNIA

2022

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Silva, Alan Henrique Ferreira  
Desenvolvimento de Métrica de Avaliação Bilingue por Estrutura Frasal aplicada à Tradução Automática [manuscrito] / Alan Henrique Ferreira Silva. - 2022.  
96 f.: il.

Orientador: Prof. Dr. Wesley Pacheco Calixto; co-orientador Dr. Gabriel Andres Wainer.  
Tese (Doutorado) - Universidade Federal de Goiás, Escola de Engenharia Elétrica, Mecânica e de Computação (EMC), Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Goiânia, 2022.  
Bibliografia.  
Inclui siglas, símbolos, gráfico, tabelas, algoritmos, lista de figuras, lista de tabelas.

1. ABEF. 2. Tradução Automática. 3. Estrutura Frasal. 4. BLEU. 5. Processamento de Linguagem Natural. I. Calixto, Wesley Pacheco, orient. II. Título.

CDU 621.3



UNIVERSIDADE FEDERAL DE GOIÁS

ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

### ATA DE DEFESA DE TESE

Ata Nº **08** da sessão de Defesa de Tese de **Alan Henrique Ferreira Silva** que confere o título de Doutor em **Engenharia Elétrica e de Computação**, na área de concentração em **Engenharia de Computação**.

Aos **oito dias do mês de outubro de dois mil e vinte e um**, a partir das **14h00min.**, realizou-se a sessão pública de Defesa de Tese intitulada **“Desenvolvimento de métrica de avaliação bilíngue por estrutura frasal aplicada à tradução automática”**. Os trabalhos foram instalados pelo Orientador, Professor Doutor **Wesley Pacheco Calixto (EMC/UFG)** com a participação dos demais membros da Banca Examinadora: Professor PhD. **Gabriel Andres Wainer - Coorientador (VSIM/Carleton University)**, membro titular externo; Professor Doutor **Márcio Rodrigues da Cunha Reis - (ENGPROD/IFG)**, membro titular externo, **cujas participações ocorreu através de videoconferência**; Professor Doutor **Leonardo Garcia Marques - (COMP/IFG)**, membro titular externo; Professor Doutor **Wanderson Rainer Hilário de Araújo - (ELE/PUCGoiás)**, membro titular externo e Professor Doutor **Gelson da Cruz Junior - (EMC/UFG)** membro titular interno. Durante a arguição os membros da banca **não fizeram** sugestão de alteração do título do **trabalho**. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor **Wesley Pacheco Calixto**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos **oito dias do mês de outubro de dois mil e vinte e um**.

#### TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **ALAN HENRIQUE FERREIRA SILVA, Discente**, em 14/10/2021, às 14:35, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Garcia Marques, Usuário Externo**, em 14/10/2021, às 15:04, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **WANDERSON RAINER HILÁRIO DE ARAÚJO, Usuário Externo**, em 14/10/2021, às 16:52, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 15/10/2021, às 09:43, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

Documento assinado eletronicamente por **MÁRCIO RODRIGUES DA CUNHA REIS, Usuário Externo**, em 15/10/2021, às 16:04, conforme horário oficial de Brasília, com fundamento no § 3º do



art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Gabriel Andres Wainer, Usuário Externo**, em 18/10/2021, às 18:47, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Gelson Da Cruz Junior, Professor do Magistério Superior**, em 19/10/2021, às 10:03, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **2397582** e o código CRC **DFE9D4C5**.

---

**Referência:** Processo nº 23070.050908/2021-79

SEI nº 2397582

*“Deus nos concede, a cada dia, uma página de vida nova no livro do tempo. Aquilo que colocarmos nela, corre por nossa conta.”*  
*Chico Xavier*



*Dedico este trabalho à minha mãe Marilene, à minha esposa  
Alana, às minhas filhas Emanuele e Helena e ao meu filho  
Henrique.*



## **AGRADECIMENTOS**

Meus agradecimentos aos que contribuíram direta e indiretamente na construção e finalização deste trabalho. Meus agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro. Meus agradecimentos ao meu orientador Pacheco pela dedicação, orientação e paciência. Meus agradecimentos à minha esposa Alana pela dedicação, ajudas e conselhos. Meus agradecimentos à Dona Maria pelo apoio. Meus agradecimentos aos meus filhos por alegrarem meus dias e me fazerem sempre querer ir em frente. Meus infinitos agradecimentos à minha mãe por ter dedicado a saúde para ensinar os verdadeiros valores da vida, a importância dos estudos e da educação de qualidade.



## RESUMO

O intuito deste trabalho é propor métrica capaz de avaliar traduções automáticas bilíngues pela utilização de estruturas frasais entre os idiomas origem e alvo. Os métodos existentes atualmente proporcionam avaliação dependente de grande volume de dados na forma de textos. A metodologia proposta utiliza como base de dados a associação de estruturas frasais genéricas entre o idioma origem e idioma alvo para avaliar a qualidade do texto resultante da ferramenta de tradução automática. A validação é realizada utilizando estudos comparativos entre as avaliações geradas pela métrica proposta e as métricas BLEU e METEOR. Os resultados obtidos apresentam maior precisão de avaliação da métrica proposta em relação às métricas comparadas, com melhoria de aproximadamente 5%, quando comparados com os resultados de tradutores humanos. Os resultados obtidos nos testes validam a capacidade da métrica como ferramenta auxiliar na avaliação de traduções automáticas.

**Palavras-chave:** ABEF, Tradução Automática, Estrutura Frasal, BLEU, Processamento de Linguagem Natural.



# DEVELOPMENT OF EVALUATION METRIC BILINGUAL BY PHRASAL-STRUCTURE APPLIED TO AUTOMATIC TRANSLATION

## ABSTRACT

This work proposes a metric capable of evaluating automatic bilingual translations, through the use of phrasal structures between the source and target languages. Existing methods currently provide a dependent assessment of large amounts of data as texts examples. The proposed methodology uses as a database the association of generic phrasal structures between the source and target language to assess the text quality, resulting given by an automatic translation machine. The method validation is performed using comparative studies between the evaluations generated by the proposed metric and the BLEU and METEOR metrics. The results obtained shows greater accuracy of evaluation of the proposed metric compared to the other metrics, with an improvement of approximately 5% compared to the results of the evaluation of human translators. The results obtained in the tests validate the capability of the metric as an auxiliary tool in the evaluation of machine translations.

**Keywords:** ABEF, Machine Translation, Phrasal-structure, BLEU, Natural Language Processing



## SUMÁRIO

Pág.

### LISTA DE FIGURAS

### LISTA DE TABELAS

### LISTA DE SÍMBOLOS

### LISTA DE ABREVIATURAS E SIGLAS

<b>CAPÍTULO 1 INTRODUÇÃO</b> . . . . .	<b>25</b>
<b>CAPÍTULO 2 TEORIA DA TRADUÇÃO</b> . . . . .	<b>31</b>
2.1 Tradução interlingual . . . . .	31
2.2 Estratégias de tradução . . . . .	32
2.3 Unidades de tradução . . . . .	33
2.4 Níveis de tradução . . . . .	34
2.5 Estrutura frasal . . . . .	35
2.6 Considerações finais . . . . .	36
<b>CAPÍTULO 3 TRADUÇÃO AUTOMÁTICA</b> . . . . .	<b>37</b>
3.1 História da tradução inter-linguística . . . . .	37
3.2 Definição . . . . .	37
3.3 Sistemas bilíngues e multilíngues . . . . .	38
3.4 Estratégias de tradução automática . . . . .	40
3.4.1 Sistema direto . . . . .	40
3.4.2 Sistema indireto interlíngua . . . . .	40
3.4.3 Sistema indireto por transferência . . . . .	41
3.4.4 Abordagem não-interativa, interativa e uso de dados auxiliares . . . . .	42
3.5 Paradigmas de tradução automática . . . . .	43
3.5.1 Tradução automática baseada em regras . . . . .	43
3.5.2 Tradução automática orientada por dados . . . . .	44
3.5.2.1 Tradução automática baseada em exemplos . . . . .	45
3.5.2.2 Tradução automática estatística . . . . .	45
3.5.2.3 Tradução automática Neural . . . . .	46
3.6 Métricas de avaliação . . . . .	47

3.6.1	<i>Bilingual Evaluation Understudy</i> – BLEU . . . . .	48
3.6.2	<i>National Institute of Standards and Technology</i> – NIST . . . . .	49
3.6.3	<i>Metric for Evaluation of Translation with Explicit Ordering</i> – METEOR . . . . .	50
3.7	Considerações finais . . . . .	51
<b>CAPÍTULO 4 PROCESSO DE OTIMIZAÇÃO . . . . .</b>		<b>53</b>
4.1	Otimização . . . . .	53
4.2	Métodos de Otimização . . . . .	55
4.2.1	Métodos Determinísticos . . . . .	55
4.2.1.1	Método Nelder-Mead Simplex . . . . .	55
4.2.2	Métodos Heurísticos . . . . .	57
4.2.2.1	Algoritmo genético . . . . .	57
4.3	Considerações finais . . . . .	60
<b>CAPÍTULO 5 METODOLOGIA . . . . .</b>		<b>61</b>
5.1	Contextualização . . . . .	61
5.2	Escolha das interfaces de programação de aplicação . . . . .	61
5.3	Métrica para aferir a acurácia . . . . .	62
5.4	Avaliação de tradução pela métrica proposta . . . . .	63
5.5	Tradução utilizando processo de otimização . . . . .	66
5.6	Validação da tradução . . . . .	67
5.7	Considerações finais . . . . .	68
<b>CAPÍTULO 6 RESULTADOS . . . . .</b>		<b>69</b>
6.1	Avaliação da tradução . . . . .	69
6.1.1	Comparação entre as métricas e validação da métrica proposta . . . . .	69
6.1.2	Avaliação de tradução frasal e comparação entre as métricas . . . . .	70
6.1.2.1	Estudo de Caso 1 . . . . .	72
6.1.2.2	Estudo de Caso 2 . . . . .	74
6.1.2.3	Estudo de Caso 3 . . . . .	76
6.1.2.4	Estudo de Caso 4 . . . . .	77
6.2	Tradução automática gerada por processo de otimização . . . . .	80
6.3	Discussão . . . . .	82
<b>CAPÍTULO 7 CONCLUSÃO . . . . .</b>		<b>85</b>
7.1	Contribuições do trabalho . . . . .	86
7.2	Sugestões para trabalhos futuros . . . . .	88
<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>		<b>89</b>

## LISTA DE FIGURAS

	<u>Pág.</u>
2.1	Mapa de disciplinas para a interface com estudos da tradução. . . . . 32
2.2	Tradução utilizando estratégia palavra-a-palavra e tradução livre. . . . . 33
2.3	Exemplo de unidade lexicológica. . . . . 34
2.4	Exemplo de composição de níveis de tradução com base na Figura 2.3. . . 35
2.5	Conjuntos de equivalência de estrutura frasal entre os idiomas Português e Inglês. . . . . 36
2.6	Tradução direta por palavras utilizando estrutura frasal. . . . . 36
3.1	Fluxo básico do sistema de Tradução Automática. . . . . 38
3.2	Ilustração da Pirâmide de Vauquois. . . . . 39
3.3	Esquema de tradução automática de categoria reversível. . . . . 39
3.4	Abordagem de tradução automática por sistema direto. . . . . 40
3.5	Abordagem de tradução automática para o sistema interlíngua de dois idiomas. . . . . 41
3.6	Abordagem de tradução automática para o sistema de transferência para três idiomas. . . . . 42
3.7	Abordagem de tradução automática para o sistema interativo. . . . . 43
3.8	Paradigmas de tradução automática. . . . . 44
3.9	Paradigma de tradução automática baseada em regras. . . . . 44
3.10	Paradigma de tradução automática baseada em exemplo. . . . . 45
3.11	Paradigma de tradução automática estatística. . . . . 46
3.12	Paradigma de tradução automática Neural. . . . . 47
3.13	Divisão de sentença por $n$ -gramas. . . . . 48
4.1	Modelo do processo de otimização. . . . . 54
4.2	Máximos e mínimos locais e globais para função objetivo $f(x_1, x_2)$ . . . . 54
4.3	Ilustração dos coeficientes utilizados no método de Nelder-Mead. . . . . 56
4.4	Fluxograma do algoritmo genético clássico. . . . . 58
4.5	Matriz populacional. . . . . 59
4.6	Operador de cruzamento uniforme. . . . . 60
5.1	Métrica de avaliação bilíngue por estrutura frasal proposta. . . . . 62
5.2	Aquisição da tradução de texto. . . . . 62
5.3	Aquisição da classificação de palavra. . . . . 63
5.4	Avaliação da tradução utilizando métrica BLEU. . . . . 63
5.5	Avaliação da estrutura frasal proposta. . . . . 64

5.6	Bloco de avaliação com base na tradução das palavras. . . . .	64
5.7	Método de Avaliação Bilíngue por Estrutura Frasal (ABEF). . . . .	65
5.8	Correspondência de estrutura frasal. . . . .	66
5.9	Tradução automática utilizando processo de otimização. . . . .	68
6.1	Trecho avaliado pelas métricas ABEF, BLEU e METEOR. . . . .	70
6.2	Dados para avaliação do $T_{O1}$ . . . . .	72
6.3	Análise de variação das distâncias para $T_{O1}$ . . . . .	73
6.4	Dados para avaliação do $T_{O2}$ . . . . .	74
6.5	Análise de variação das distâncias para $T_{O2}$ . . . . .	75
6.6	Dados para avaliação do $T_{O3}$ . . . . .	76
6.7	Análise de variação das distâncias para $T_{O3}$ . . . . .	77
6.8	Trechos extraídos do livro <i>The richest man in Babylon</i> . . . . .	78
6.9	Trechos extraídos do livro <i>Think and grow rich</i> . . . . .	79
6.10	Gráficos de dispersão para os trechos analisados. . . . .	79
6.11	Métricas e avaliações: (a) distâncias entre as avaliações e (b) sobreposição das dispersões. . . . .	80
6.12	Processo de otimização para $T_{O1}$ . . . . .	82
6.13	Processo de otimização para $T_{O2f}$ . . . . .	82
6.14	Processo de otimização para $T_{O3}$ . . . . .	83

## LISTA DE TABELAS

	<u>Pág.</u>
2.1 Equivalência entre tipos gramaticais Português-Inglês. . . . .	35
6.1 Avaliações para o trecho traduzido $T_T$ . . . . .	70
6.2 Textos origem utilizados para avaliação das traduções. . . . .	71
6.3 Avaliações para $T_{O1}$ . . . . .	73
6.4 Análise com relação a avaliação humana para $T_{O1}$ . . . . .	73
6.5 Avaliações para $T_{O2}$ . . . . .	75
6.6 Análise com relação a avaliação humana para $T_{O2}$ . . . . .	75
6.7 Avaliações para $T_{O3}$ . . . . .	77
6.8 Análise com relação a avaliação humana para $T_{O3}$ . . . . .	77



## LISTA DE SÍMBOLOS

$BP$	–	Prena de brevidade
$C_R$	–	Código de referência
$D_{API}$	–	Dicionário em API
$D_{BL}$	–	Dicionário bilíngue
$DIC_{EF}$	–	Dicionário de Estruturas Frasais
$EF_{(I_O, I_A)}$	–	Estrutura Frasal equivalente entre $I_O$ e $I_A$
$EF_{I_A}$	–	Estrutura Frasal de $I_A$
$EF_{I_O}$	–	Estrutura Frasal de $I_O$
$EF_{EN}$	–	Estrutura frasal em inglês
$EF_O$	–	Estrutura frasal do texto origem
$EF_{PT}$	–	Estrutura frasal em português
$EF_{T_T}$	–	Estrutura Frasal do Texto Traduzido
$EF$	–	Estrutura Frasal
$F_{ABEF}$	–	Nota de avaliação do método ABEF
$F_{aval}$	–	Avaliação produzida no otimizador
$F_{EF}$	–	Avaliação da estrutura frasal
$F_n$	–	Avaliação para cada palavra por equivalência validada
$f(x)$	–	Função objetivo
$G$	–	Geração
$I_{O_p}$	–	Idioma origem projetado
$I_1, I_2$ e $I_3$	–	Idiomas correlacionados
$I_A$	–	Idioma alvo
$I_O$	–	Idioma origem
$L_1$	–	Língua origem
$L_2$	–	Língua alvo
$M_O$	–	Modelo no idioma origem
$M_T$	–	Modelo de tradução
$n$ -grama	–	Conjunto de palavras
$P_{BLEU}$	–	Nota de avaliação da métrica BLEU
$P_{METEOR}$	–	Nota de avaliação da métrica METEOR
$P_{MT}$	–	Prena de brevidade da métrica METEOR
$P_{NIST}$	–	Nota de avaliação da métrica NIST
$P(I_O I_A)$	–	Probabilidade de tradução entre $I_O$ e $I_A$
$S_{BL}$	–	Sistema bilíngue
$T_A$	–	Texto alvo
$T_C$	–	Texto candidato
$T_O$	–	Texto origem
$T_T$	–	Texto traduzido
$TA_{API}$	–	API de Tradução Automática
$x_{otim}$	–	Valores otimizados de $x$
$V_h$	–	Avaliação humana



## LISTA DE ABREVIATURAS E SIGLAS

ABEF	–	Avaliação Bilíngue por Estrutura Frasal
AG	–	Algoritmo Genético
API	–	<i>Application Programming Interface</i>
BERT	–	<i>Bidirectional Encoder Representations from Transformers</i>
BLEU	–	<i>Bilingual Evaluation Understudy</i>
METEOR	–	<i>Metric for Evaluation of Translation with Explicit Ordering</i>
MT	–	<i>Machine Translation</i>
NIST	–	<i>National Institute of Standards and Technology</i>
NMT	–	<i>Neural Machine Translation</i>
PF	–	Perspectiva Funcional da Frase
RNA	–	Redes Neurais Artificiais
RNR	–	Redes Neurais Recorrentes
SMT	–	<i>Statistical Machine Translation</i>
TA	–	Tradução Automática
TABR	–	Tradução Automática Baseados em Regras
TAE	–	Tradução Automática Estatística
TAN	–	Tradução Automática Natural
TAO	–	Tradução Automática por Otimização
TAPD	–	Tradução Automática Orientada por Dados



# CAPÍTULO 1

## INTRODUÇÃO

No mundo de pluralidade línguocultural, fortemente conectado física e eletronicamente, surgem constantes desafios para que nações de diferentes idiomas se comuniquem e busquem conteúdo de forma plurilateral. O estabelecimento de comunicação eficaz depende diretamente da capacidade de entendimento e de se fazer entendido em determinado idioma. Aparatos eletrônicos tem auxiliado na compreensão de vários idiomas desde a década de 1970 (NESI, 2009). Pesquisadores vêm criando técnicas e mecanismos para possibilitar a tradução entre um idioma e outro, utilizando tradução automática (*machine translation*) com diversas finalidades como: ofuscação autoral (RAO et al., 2000), identificação autoral (CALISKAN; GREENSTADT, 2012) e, principalmente, compreensão de conteúdos entre idiomas.

Além dos úteis e antigos dicionários bilíngues, nos dias atuais há extensa variedade de ferramentas eletrônicas para auxílio na tradução entre idiomas distintos. Weaver (1955) relata o surgimento da tradução automática (TA) no ano de 1946, dois anos após a Segunda Guerra Mundial, devido ao advento da utilização dos computadores para solução de diversos tipos de problemas, dentre eles a tradução para fim militar.

A técnica de TA ganha relevância devido à capacidade de processamento, velocidade e flexibilidade. Wall (1956) destaca os desafios de realizar tradução através de tradutor humano, comparando custos oriundos deste método com o método computacional de TA. Gil e Pym (2006) destacam os desafios e numerosas vantagens de se explorar o ferramental tecnológico para obter avanços nas técnicas de tradução automática.

Com os avanços tecnológicos e a Internet cada vez mais utilizada e necessária, em meados de 1988 surgem os primeiros serviços de tradução *online*. Contudo, estes recursos são pagos e contam com sérias dificuldades quanto à qualidade dos resultados finais. No ano de 1997 surgem serviços *online* mais elaborados e refinados. As inovações e ferramenta *online* gratuita de tradução são disponibilizados pela Google. A ferramenta disponibilizada, o *Google Translate*, conta com os mais modernos refinamentos desenvolvidos em tradução automática estatística (HUTCHINS, 2009). Algumas tecnologias como a tradução de páginas *web* também são disponibilizadas, mas sempre apresentando dificuldades em expressões coloquiais e termos culturalmente dependentes<sup>1</sup> (HUTCHINS, 1999).

---

<sup>1</sup>Expressões usadas em região geográfica ou comunidades de mesma base cultural.

Métodos como o patenteado por [Koyama et al. \(1996\)](#) criam tecnologias para registro de idioma, evitando a necessidade de armazenamento com redundância e gerando nova opção, ante a tradução direta palavra-a-palavra. Na criação desta solução é apresentada metodologia de divisão de frase, atribuindo símbolo para cada parte, denominado método representativo. Tal técnica apresenta capacidade de categorizar cada termo da frase. O resultado desta metodologia aplicada é possibilitar maior velocidade no registro das sentenças utilizando menor espaço de armazenamento de dados.

[Leffa \(1992\)](#) apresenta pesquisa sobre aplicação de tradução assistida por programa de computador aplicada na tradução de língua inglesa para língua portuguesa. A abordagem desta pesquisa utiliza a comparação sistêmica e quantitativa do tempo utilizado e quantidade de entendimento adquirido, comparando a utilização de glosários eletrônicos em detrimento do modelo físico tradicional. Neste trabalho é apresentado o ganho de velocidade no processo de tradução com aumento da eficiência em 38% e redução do tempo de tradução em 50%, quando utilizada ferramenta computacional. Esta metodologia apresenta o impacto da tecnologia no auxílio das traduções.

[Hutchins \(2009\)](#) implementa e refina metodologia em TA, utilizando a tradução automática estatística (*Statistical Machine Translation*). Neste trabalho [Hutchins \(2009\)](#) desenvolve estudo relacionando tecnologias existentes às ferramentas disponíveis na Internet. O autor apresenta os avanços gerados pela técnica de TA aplicada à ferramentas *online*, acrescentando substancial avanço no campo de comunicação e processamento linguístico.

Várias são as aplicações de TA para tradução ou melhoria de processo de tradução e aprendizado de idiomas. [Dong e Cremers \(2014\)](#) apresentam ferramenta voltada à aplicação de TA no auxílio educacional bilinguístico para aprendizado do idioma Chinês a falantes nativos do idioma Alemão. O processo aplicado neste trabalho consiste no cruzamento entre idiomas utilizando sistema de registro em memória hierárquico. A ferramenta criada apresenta soluções resultantes da aplicação do novo método proposto pelos autores, demonstrando aquisições automatizadas de termos entre as línguas alemã, inglês e chinês. O método se apresenta capaz de resolver problemas específicos da TA tradicional, como técnica facilitadora para acoplagem em sistema de TA de propósito geral.

[Marcu e Wong \(2002\)](#) apresentam modelo de junção probabilística para tradução automática estatística (TAE). No estudo deste modelo é abordada tabela de equi-

valência para treinamento do método, apresentando capacidade de treinamento e aprendizado em nível de software. O sistema proposto modela o problema de tradução e realiza treinamento, utilizando equacionamento, gerando resultado traduzido através de método probabilístico. O método apresenta acurácia na tradução de frases, no entanto, com restrições quanto à precisão da tradução de sentenças da língua inglesa para a francesa.

Cho et al. (2014) estudam como as novas abordagens de TA são acopladas a mecanismos de aprendizagem e formação de base de dados de tradução através da técnica de redes neurais artificiais (RNA). A abordagem deste trabalho utiliza redes neurais recorrentes (RNR), que utiliza estado oculto e saídas opcionais para processar conjunto de soluções candidatas. Os resultados são obtidos através de método codificação/decodificação, baseado em memória e ranqueamento. Métodos como tradução automática neural (TAN) são capazes de alcançar resultados além da TA tradicional, trazendo melhorias em traduções baseadas em frases para tradução inglês/francês.

Lu et al. (2007) apresentam novas tecnologias acopladas à TAE para agregar qualidade em traduções utilizando métodos de otimização para formação de bases de dados *online* ou *offline*, através de seleção de termos baseada na frequência da ocorrência de cada palavra. As informações são processadas, é realizado treinamento *offline* e posteriormente realizada otimização das correlações *online* através da atribuição de pesos para cada conjunto gerado pelo treinamento. Como fruto da metodologia proposta, são obtidas traduções mais precisas se comparadas à modelos TAE tradicionais. Na maioria dos casos com treinamento mais curto e melhor desempenho.

Wang et al. (2016) propõem método utilizando três idiomas para treinamento de RNA com intuito de aumentar a taxa de acerto nas traduções utilizando método de comparação paralela entre as combinações inglês/japonês/chinês. Na construção do método proposto é utilizada TAN baseada em RNR, traduzindo conteúdo original para o **idioma alvo** desejado, paralelamente com outra tradução de mesmo sentido origem alvo de forma síncrona através de vetores origem-1 e origem-2 previamente relacionados. Após a aplicação do método, as traduções obtidas são avaliadas como tendo precisão maior que o ponto de partida dos testes usando unicamente TAN. Tal inovação apresenta capacidade de equiparar ou produzir resultados de maior precisão, comparadas ao método base utilizado preliminarmente.

Yang et al. (2020) apresentam estudo da aplicação de *Deep Learning* no processamento de linguagem natural através de processo de tradução automática neural

(*Neural Machine Translation - NMT*). A metodologia apresenta a avaliação de diferentes variações de análise, orientação de tradução por categoria. Após o processamento da sentença original, o texto é codificado e decodificado, tendo os termos das sentenças extraídos e categorizados. O processo de desestruturação dos textos utiliza diferentes camadas de pesquisa, fornecendo base de dados para treinamento das redes neurais voltadas à tradução automática de textos. Este método ainda encontra limitações impostas pela abrangência de vocabulário, contudo trazendo significativo avanço na cobertura de termos.

Apesar dos avanços na área de TA, com as inovações e melhorias criadas pela TAN, ainda existem problemas a serem resolvidos no campo da tradução via software. Existem problemas como os relatados por [Luong et al. \(2014\)](#), onde determinadas frases ainda apresentam falta de ordenação de termos nas sentenças resultantes dos processos de tradução. [Luong et al. \(2014\)](#) apresentam proposta para preencher espaços deixados pela tradução automática neural tradicional, através de método para tratar problemas de palavras fora de vocabulário (*Out-of-vocabulary*). O método consiste em identificar e emitir o posicionamento de cada uma das palavras fora de vocabulário. Posteriormente estas palavras são substituídas com auxílio de dicionário. Após processamento das frases e avaliação da precisão, o resultado apresenta maior acerto na correspondência e fidelidade de tradução.

Há métricas como a *Bilingual Evaluation Understudy* (BLEU) ([PAPINENI et al., 2002](#)), criada para dar celeridade ao processo de tradução via software, contrapondo alto custo e longo tempo para tradução realizada por humanos. [Papineni et al. \(2002\)](#) propõem métrica para avaliação de TA que seja rápida, de baixo custo, independente de idioma e que possua alto nível de correlação com a ponderação humana. A métrica consiste em comparar a sentença candidata a solução com sentença reduzida esperada, realizando contagem de termos correspondentes entre candidata e referência correta. São avaliados resultados de tradução entre o idioma inglês e os idiomas árabe, chinês, francês e espanhol. O trabalho apresenta técnica capaz de gerar valoração numérica, robusta e simples para classificar traduções obtidas por TA.

[Callison-Burch et al. \(2006\)](#) descrevem sobre limitações da métrica BLEU quanto à produção do mesmo acerto para milhões de possibilidades de tradução de determinada sentença. Na pesquisa apresentada por [Callison-Burch et al. \(2006\)](#) são realizados testes de permuta de frases e modificação de termos através de conjuntos de **palavras candidadas**. Nos testes realizados o método aplicado apresenta-se, em

casos diversos, incapaz de garantir fielmente a qualidade do julgamento humano sobre as traduções, mas ainda superior a outros sistemas de avaliação como o método *National Institute of Standards and Technology* (NIST) e o método *Metric for Evaluation of Translation with Explicit Ordering* (METEOR) (DORR et al., 2018). Existem qualidades e limitações da métrica BLEU, com destaque às falhas de incapacidade de senso de desambiguação<sup>2</sup> de palavras e de contornar variações léxicas<sup>3</sup>.

Vários são os trabalhos realizados na área de TA, no entanto ainda existem lacunas para produção de tecnologias aplicadas no auxílio da tradução. Ferramentas eletrônicas são aparatos que favorecem o avanço em TA e o processo de otimização computacional oferece novo arcabouço matemático na solução de problemas de tradução. Um dos pontos-chave das TA são as métricas de avaliação dos resultados de tradução. Métricas de pontuação apresentam semelhança ao método BLEU, no entanto, estão com credibilidade inferior (CALLISON-BURCH et al., 2006; LEE; PRZYBOCKI, 2005; KOEHN, 2009).

Mathur et al. (2020) apresentam comparativo entre a métrica BLEU e outras técnicas derivadas. Pesquisas recentes demonstram limitações de codificadores multilíngue quando apreciadas/estimadas por metodologias de avaliação de tradução automática livres de referência (ZHAO et al., 2020). As limitações destas métricas deixam lacunas para novas pesquisas/desenvolvimentos, possibilitando a exploração de novas abordagens para melhoria da qualidade de avaliação em TA, justificando assim este trabalho.

Levando em consideração o fértil campo da pesquisa a ser explorado e as lacunas ainda existentes, pode-se levantar a hipótese: **se** é possível realizar a tradução de texto da língua origem para outra língua através de softwares de tradução automática e se é possível realizar o processo inverso, voltando o texto no novo alvo para o idioma de origem, **então** é possível utilizar as estruturas frasais entre os idiomas para avaliar a precisão da tradução atribuindo avaliação. Portanto, de posse de software de tradução automática (TA), pode-se realizar traduções que podem ser avaliadas por funções matemáticas e utilizando bibliotecas relacionais de tradução bilíngue, avaliando as palavras, frases e outras expressões para auxiliar na qualidade da tradução através da nota atribuída.

O objetivo geral deste trabalho é criar metodologia para mensurar a qualidade da

---

<sup>2</sup>Efeito de desambiguar, retirando ambiguidade onde pode haver mais de um sentido em palavra ou expressão.

<sup>3</sup>Diferentes palavras para denominar mesmo objeto, ação ou sujeito.

tradução de textos de determinada língua para outra, além da aplicação da métrica como função de avaliação em mecanismo de tradução automática por processo de otimização. Ainda como objetivos específicos têm-se: i) desenvolvimento de metodologia de aquisição e tratamento dos textos originais, ii) desenvolvimento de função de avaliação para mensurar a qualidade final da tradução, e iii) comparação das avaliações geradas pela métrica proposta e métricas predecessoras às avaliações de tradutores humanos.

Este trabalho está estruturado de forma a abordar teoria envolvida no desenvolvimento da métrica proposta. No Capítulo 2 é apresentada a teoria da tradução. No Capítulo 3 são apresentados métodos de tradução automática, as principais técnicas de tradução e métricas de avaliação de tradução. No Capítulo 4 é descrita a teoria de processos de otimização computacional. No Capítulo 5 é apresentada metodologia aplicada no desenvolvimento da métrica. O Capítulo 6 dispõe de estudos de caso e resultados obtidos, enquanto o Capítulo 7, dispõe das conclusões, produções e sugestões para trabalhos futuros.

## CAPÍTULO 2

### TEORIA DA TRADUÇÃO

Neste capítulo será apresentada a teoria da tradução utilizada como base fundamental na formatação e desenvolvimento de métodos de tradução automática. São descritos o histórico sobre a tradução, linguística e conceitos sobre classificação gramatical de palavras.

#### 2.1 Tradução interlingual

A pluralidade de etnias, geolocalizações e diferentes origens culturais fez do planeta Terra o local que abriga incontáveis tipos e formas de comunicação. Para que possa ocorrer diálogo entendível entre pessoas que se expressam em diferentes idiomas, é necessário que ocorra a tradução. Segundo o dicionário *The Concise Oxford English Dictionary* a tradução é o ato ou a instância de transformar uma expressão escrita ou falada da significação de uma palavra, fala ou livro em outro idioma (SOANES; STEVENSON, 2004). Hatim e Munday (2004) definem tradução em duas perspectivas. A primeira foca nas regras de tradução entre o texto origem  $T_O$  e o texto alvo  $T_A$ . A segunda perspectiva é centrada no produto concreto gerado pelo tradutor.

Há, no entanto, confusão gerada entre a tradução literal e o significado da mensagem a ser transmitida. A potencial confusão provém da interpretação entre o significado direto e a significação pretendida. Mesmo excluída a interpretação, os campos cobertos pela tradução são vastos e complexos quando levados em consideração todas as nuances de significação de expressões utilizadas em distintos idiomas (SHUTTLEWORTH, 2014).

Segundo Hatim e Munday (2004) há três tipos de tradução: i) tradução intralingual, ii) tradução interlingual e iii) tradução intersemiótica. A tradução intralingual ocorre quando há a tradução dentro do mesmo idioma, envolvendo reordenação ou parafraseamento dos conteúdos. Na tradução interlingual ocorre a tradução entre dois idiomas distintos. Em tradução intersemiótica envolve a transmissão de mensagem traduzida através de sinais não-verbais, tais como música ou imagem. Sendo considerada pela maioria dos autores como tradução somente o modelo interlingual (SOANES; STEVENSON, 2004).

O estudo da tradução compreende várias disciplinas interconectadas, como ilustra a Figura 2.1, adaptada de Hatim e Munday (2004). Como composição para cada campo disciplinar têm-se: i) **linguística**: semântica, pragmática, sociolinguística,

corpo linguístico, linguística cognitiva, análise textual e discursiva; ii) **filosofia**: hermenêutica, desconstrução, pós-estruturalismo; iii) **engenharia de idioma**: tradução automática, corpora, terminologia, lexicologia, multimídia; iv) **estudos culturais**: filmes, idioma e poder, ideologia, estudo de gênero, história, pós-colonialismo e v) **estudos literários**: poesia, retórica, crítica literária, narratologia, análise de discurso crítico, literatura comparativa.

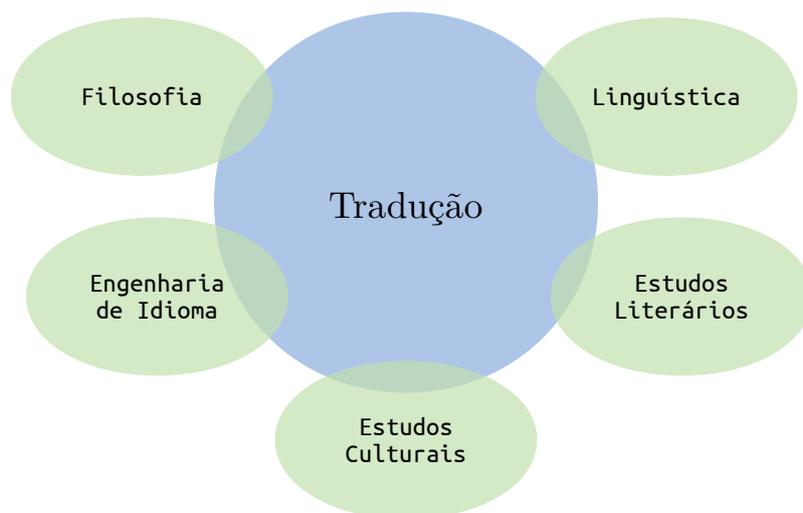


Figura 2.1 - Mapa de disciplinas para a interface com estudos da tradução.

Além da composição multidisciplinar, existe a delimitação de campos técnicos de pesquisa como análise textual de tradução, assistência em qualidade de tradução, tradução de literatura e outros gêneros, tradução audio-visual, tradução e tecnologia, tradução histórica, tradução ética, terminologia e glossários, processo de tradução, treinamento de tradução e as características do ofício de tradução. Trata-se de campo vasto e pode estar conectado à áreas de conhecimento das mais variadas origens e natureza.

## 2.2 Estratégias de tradução

A tradução trata-se de processo de renderização do significado, com aderência à forma e ênfase na acurácia geral. Há expressões que não podem ser literalmente traduzidas do idioma origem ao idioma alvo ou até mesmo que não possuem tradução direta por questões culturais. Há neste caso, a dicotomia na tradução entre sentido/conteúdo e do outro lado a forma/estilo. O sentido deve ser traduzido, enquanto que a forma geralmente não pode ser. Exemplos deste tipo de situação são

as traduções de poesias, músicas, propaganda e outros, na qual a rima ou o duplo significado, na maioria dos casos, não podem ser recriados no idioma alvo.

A divisão entre forma e conteúdo está ligada em várias maneiras e resulta em dois tipos básicos de tradução: i) literal e ii) livre. A tradução literal, realizada palavra-a-palavra, pode ocasionar a perda do significado da mensagem original. Por outro lado, a tradução livre leva em consideração aspectos como significado dentro do contexto, valores culturais e o contexto que origina a tradução. A Figura 2.2 ilustra a expressão em português traduzida para o inglês. Nela é utilizada estratégia literal palavra-a-palavra e em seguida a estratégia de tradução livre. No caso da Figura 2.2, a tradução palavra-a-palavra, que não considera contexto, distorce completamente o significado resultante na tradução do português para o inglês.

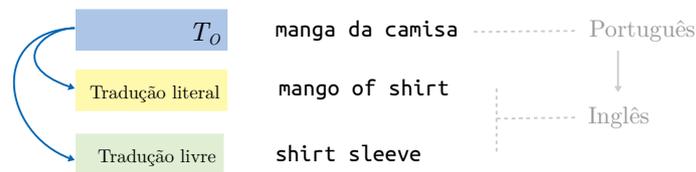


Figura 2.2 - Tradução utilizando estratégia palavra-a-palavra e tradução livre.

### 2.3 Unidades de tradução

As unidades de tradução são definidas como o nível linguístico em que o texto origem  $T_O$  é recodificado para o idioma alvo  $I_A$ . É o elemento considerado pelo tradutor quando está conduzindo o processo de tradução. Diferentes tipos de blocos podem ser levados em consideração quando da tradução, como: i) palavras, ii) grupo de palavras, iii) cláusula, iv) sentença e v) todo o texto. A divisão entre texto origem e texto traduzido ( $T_O$  e  $T_T$ ) para as unidades de tradução são de extrema importância na estruturação do processo de tradução (SHUTTLEWORTH; COWIE, 1997).

Há necessidade de definir o que é **significador** e **significado**. O **significador** representa o conteúdo ao assunto ou objeto da tradução, seja som, imagem ou palavra. O **significado** trata diretamente do conceito aplicado. A estes dois termos é dado o nome de assinatura linguística (VINAY; DARBELNET, 1995). Assim, o **significador árvore** refere-se ao significado real da palavra sendo planta com caule fixo no solo, com raízes, com base despida e com folhagem na parte superior. A unidade lexicológica é o grupo de elementos léxicos reunidos para formar único elemento de pensamento.

A Figura 2.3, adaptada de Hatim e Munday (2004), ilustra o exemplo de texto origem  $T_O$  em espanhol traduzido para inglês e português. A expressão  $T_O$  representa única unidade de pensamento e que pode ter diferentes significados quando traduzidas, na qual as mudanças podem ser verificadas quando traduzidas para o inglês e para o português. Entre os três idiomas há equivalência entre o termo **surto**, no entanto, é verificável a divergência quanto ao significador da doença da **febre aftosa**, que possui equivalência de tradução direta somente entre espanhol-português.

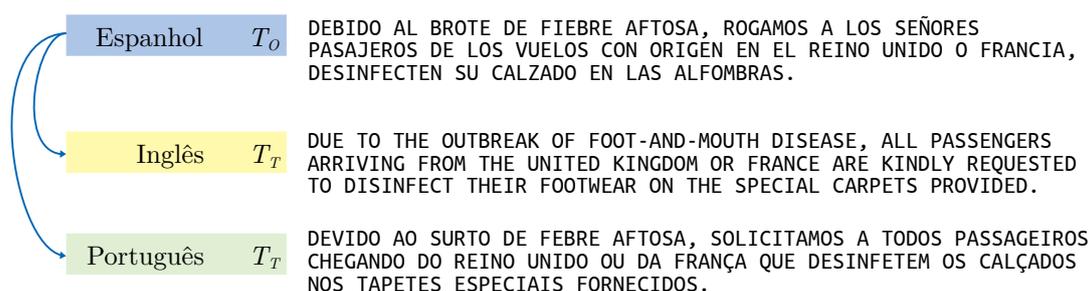


Figura 2.3 - Exemplo de unidade lexicológica.

## 2.4 Níveis de tradução

Dentro do processo de tradução existem diferentes níveis de tradução. Podem ser considerados: i) morfema, ii) palavra, iii) frase, iv) oração e/ou v) período. No decorrer do procedimento de tradução são examinadas equivalências em nível de palavra, colocação, gramática, temática e também na estrutura da informação, coesão e pragmática. A análise isolada de palavras gera traduções incoerentes, como ilustrado na Figura 2.2. O segundo nível de análise aborda partes relacionadas entre os grupos de palavras e período. Neste ponto, há avaliação do significado da expressão antes de ser diretamente traduzida para o idioma alvo  $I_A$  (NEWMARK, 1988).

Considerando a estrutura dos textos ilustrados na Figura 2.3, é extraído conjunto de termos que definem as unidades de nível de tradução, como ilustrado na Figura 2.4. Palavras e grupos de palavras são os níveis de tradução iniciais, no entanto, nem sempre fornecem tradução fidedigna. A análise sistêmica dos demais níveis oferecem maior significância no contexto comunicativo, tendo assim seu papel ainda mais importante que os dois níveis anteriores (HALLIDAY; MATTHIESSEN, 2013).

No processo de tradução devem ocorrer transposições e rearranjos frequentemente. Estas modificações ou mesmo divisões devem respeitar a perspectiva funcional da

morfema – *chega-ndo*  
 palavra – *chegando*  
 grupo – *febre aftosa*  
 oração – *devido ao surto de febre aftosa*  
 período – *Devido ao surto de febre aftosa, solicitamos a todos passageiros  
 chegando do Reino Unido ou da França que desinfetem os  
 calçados nos tapetes especiais fornecidos.*

Figura 2.4 - Exemplo de composição de níveis de tradução com base na Figura 2.3.

frase (PFF). Esta perspectiva define que a estrutura sintática, ou modificação linear, é importante dispositivo de estruturação. Entretanto, a comunicação é dirigida primariamente pelo dinamismo-comunicativo, ou seja, por elementos que são independentes de conceito e contribuem majoritariamente para novas informações (NEWMARK, 1988).

## 2.5 Estrutura frasal

As frases e períodos são a representação central do sentido da mensagem a ser transmitida. Levando em consideração o nível da frase, a avaliação e transferência da mensagem entre idioma origem ( $I_O$ ) e idioma alvo ( $I_A$ ), pode-se realizar a tradução diretamente por estruturas frasais compatíveis. Por exemplo, considerando os idiomas português e inglês, na Tabela 2.1 são dispostas equivalências entre classes gramaticais destes dois idiomas, na qual  $C_R$  é o código de referência.

Tabela 2.1 - Equivalência entre tipos gramaticais Português-Inglês.

#	Português	Inglês	$C_R$
1	Adjetivo	<i>Adjective</i>	A
2	Adverbio	<i>Adverb</i>	B
3	Artigo	<i>Determiner</i>	C
4	Conjunção	<i>Conjunction</i>	D
5	Interjeição	<i>Interjection</i>	E
6	Numeral	<i>Determiner</i>	F
7	Preposição	<i>Preposition</i>	G
8	Pronome	<i>Pronoun</i>	H
9	Substantivo	<i>Noun</i>	I
10	Verbo	<i>Verb</i>	J

Utilizando estruturas gramaticais e com base em conhecimento pré-definidos, por exemplo, consolidados por tradutores humanos, é possível definir modelos de equivalência entre  $I_O$  e  $I_A$  com base em conjuntos frasais equivalentes, como ilustra a Figura 2.5, na qual  $EF_{PT}$  é o conjunto de estruturas frasais em português e  $EF_{EN}$  é o conjunto de estruturas frasais em inglês. Podem existir mais de uma estrutura

frasal ( $EF$ ) equivalente entre as  $EF$  de cada idioma, possibilitando distintas configurações, porém com mesmo sentido e mensagem final.

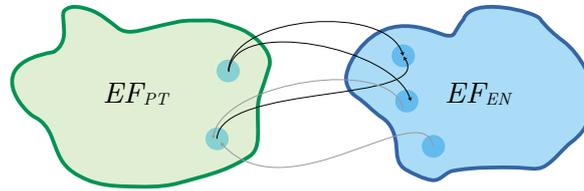


Figura 2.5 - Conjuntos de equivalência de estrutura frasal entre os idiomas Português e Inglês.

Como técnica de tradução automática (TA) a tradução por estrutura frasal oferece ferramenta parametrizada através de banco de dados de equivalência de estruturas frasais entre os idiomas, possibilitando a sistematização da tradução através de sistema computacional. A Figura 2.6 ilustra a tradução direta por palavras, utilizando como âncora a correspondência da estrutura frasal de  $T_O$  dada por  $EF_O = C + I + J + A$  e  $EF_A = C + I + J + A$ . Neste exemplo  $EF_O \equiv EF_A$  é a mesma estrutura frasal  $C + I + J + A$ .

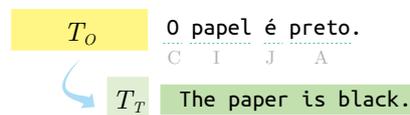


Figura 2.6 - Tradução direta por palavras utilizando estrutura frasal.

No exemplo ilustrado na Figura 2.6 foi utilizada como base de comparação e busca a estrutura frasal equivalente nos dois idiomas. Após a localização, é procurada a tradução direta para cada um dos elementos da frase e em seguida gerada a tradução final para o  $I_A$ , inglês.

## 2.6 Considerações finais

Neste capítulo foram apresentadas as fundamentações teóricas e práticas utilizadas no processo de tradução. Foram apresentados os níveis fundamentais e exemplos de resultados de tradução. As definições apresentadas neste capítulo formam a base para a teoria e técnicas de tradução automática apresentada no próximo capítulo.

## CAPÍTULO 3

### TRADUÇÃO AUTOMÁTICA

Neste capítulo serão apresentadas as técnicas de tradução automática. É realizada contextualização histórica e definidos os sistemas bilinguais e multilinguais, as estratégias de tradução automática de sistema direto, os paradigmas de tradução automática e as métricas de avaliação, por conseguinte, sua estruturação e desdobramentos.

#### 3.1 História da tradução inter-linguística

É fato que a multiplicidade de idiomas por vezes impede o intercâmbio intelectual e cultural entre pessoas do planeta (SPINKS; WELLS, 1997). A tradução realizada por humanos, dotados do conhecimento de idiomas, entre os interessados em manter conversação ou intercâmbio de informações é a principal maneira de realizar tal comunicação (SLOCUM, 1985). Com o passar dos anos são criados dicionários bilinguais impressos para auxiliar na tarefa das traduções, mas somente como ferramenta auxiliar nesta demanda.

Com o advento tecnológico dos computadores e a impulsão dada pelas soluções computacionais na elucidação de enigmas criptografados, surge em 1947 o que seria o início da tradução automática (TA) (WEAVER, 1955). A criação de método para tradução automatizada por computadores é alavancada no pós Segunda Guerra Mundial, devido à preocupação da *United Nations Educational, Scientific and Cultural Organization* (UNESCO) quanto à comunicação entre povos dos diversos países do planeta envolvidos na guerra (AC03561797, 1957).

#### 3.2 Definição

A tradução automática é a técnica computacional utilizada para transformar texto digital, escrito em determinado **idioma origem**, em texto resultado no **idioma alvo**, diferente do idioma do texto origem. A concepção original de TA aborda modelo bilingual de tradução (HUTCHINS; SOMERS, 1992). O processo compreendido por sistemas de TA pode utilizar métodos totalmente automáticos e modelos que utilizam intervenção humana.

Para todo sistema de TA o principal intuito é alcançar a tradução de qualidade e fidelidade ao significado original. Assim como em traduções realizadas por humanos, as traduções realizadas por computadores são revisadas por outros tradutores antes da divulgação ou utilização em larga escala. Por vezes o resultado da tradução

computadorizada pode ser utilizada como pré-tradução ou como rascunho para que o restante do processo seja continuado por pessoas, conduzindo na sequência a geração do texto final (HUTCHINS, 1986). O fluxo básico de TA é ilustrado na Figura 3.1, onde  $I_O$  é texto digital no **idioma origem**,  $I_A$  é texto digital no **idioma alvo** e  $TA$  o sistema computacional de tradução automática.

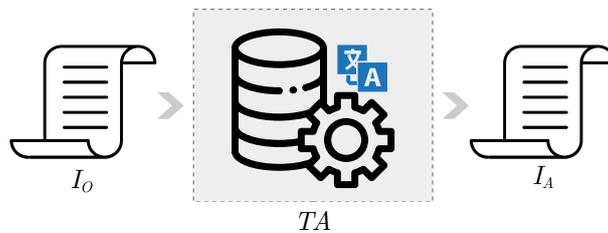


Figura 3.1 - Fluxo básico do sistema de Tradução Automática.

Sistemas de TA tem em sua composição interna pelo menos dois componentes para efetuar o processo de tradução: análise e geração. Estes componentes são partes separadas do sistema e tem funcionamento independente. A **análise** consiste na etapa de avaliação da estrutura frasal, identificando os elementos gramaticais do texto a ser traduzido. A **geração** trata da criação dos termos correspondentes, criando como resultado os elementos gramaticais traduzidos no idioma alvo (NIRENBURG et al., 1993).

Os sistemas primários de TA podem ser representados através da pirâmide de Bernard Vauquois. A pirâmide representa o comparativo funcional entre sistema direto e os sistemas indiretos, interlíngua e baseado em transferência. A Figura 3.2, adaptada de (PRASAD; MUTHUKUMARAN, 2013), ilustra a estratégia direta na base da pirâmide, passando pelo método por transferência no meio da pirâmide e no topo, abordagem interlíngua. Todos estes métodos possuem ao menos os passos de análise e geração para traduzir  $I_O \rightarrow I_A$  (OKPOR, 2014).

### 3.3 Sistemas bilíngües e multilíngües

Na TA clássica é abordada tradução unidirecional  $I_O \rightarrow I_A$ . Em derivações da TA clássica, são utilizados modelos bidirecionais  $I_O \leftrightarrow I_A$ , podendo ser divididos em sistemas reversíveis ou não-reversíveis (HUTCHINS, 1986). A Figura 3.3, ilustra os dois sentidos possíveis da realização da tradução. Para este caso, os passos do processo de geração de linguagem são diretamente inversos aos da análise de linguagem. Os

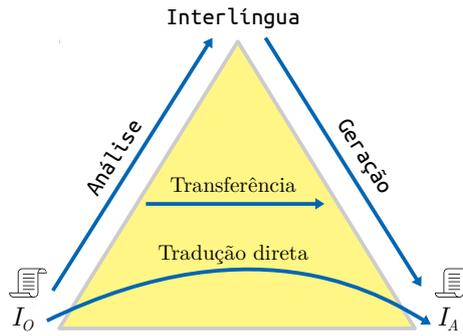


Figura 3.2 - Ilustração da Pirâmide de Vauquois.

processos de geração e de análise são realizados separadamente.

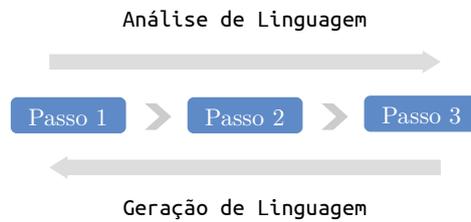


Figura 3.3 - Esquema de tradução automática de categoria reversível.

Os sistemas reversíveis não garantem eficácia e confiabilidade necessária na maioria dos casos. Em geral, sistemas bilíngues são compostos por dois sistemas unidirecionais processados pelo mesmo computador. Um sistema bilingue  $S_{BL}$  é representado por (3.1), onde  $L_1$  é a língua origem e  $L_2$  a língua alvo. Sistemas de tradução bilíngue são tipicamente projetados para traduzir do idioma origem ao idioma alvo em apenas uma direção (HUTCHINS; SOMERS, 1992).

$$S_{BL} : L_1 \rightarrow L_2 \quad (3.1)$$

Os sistemas envolvendo mais de duas línguas são denominados multilíngues. Por vezes sistemas multilíngues não cobrem todas as possibilidades de tradução, tendo em vista que envolvem múltiplos idiomas origem e apenas um idioma alvo (HUTCHINS; SOMERS, 1992). Não há decisão absoluta sobre qual dos sistemas é o melhor, sendo necessário avaliar a situação específica para determinar a solução mais apropriada.

### 3.4 Estratégias de tradução automática

Existem três estratégias básicas de TA: i) sistema direto de tradução, ii) sistema de tradução por transferência e iii) sistema interlíngua. Estratégias de tradução direta são considerados como sistemas de primeira geração, enquanto os demais são considerados mecanismos de abordagem indireta.

#### 3.4.1 Sistema direto

O sistema **direto** dispensa qualquer estágio intermediário no processo de tradução. Este modelo de tradução é o primeiro idealizado, tendo em vista o poder computacional disponível à época de sua criação, no qual a tecnologia existente é limitante (BHATTACHARYYA, 2015). O sistema direto processa o texto no idioma original  $I_O$  através da análise morfológica, seguida por busca em dicionário bilíngue, reordenação dos elementos da sentença e por fim, remontagem da sentença no idioma alvo  $I_A$ , como ilustra a Figura 3.4.

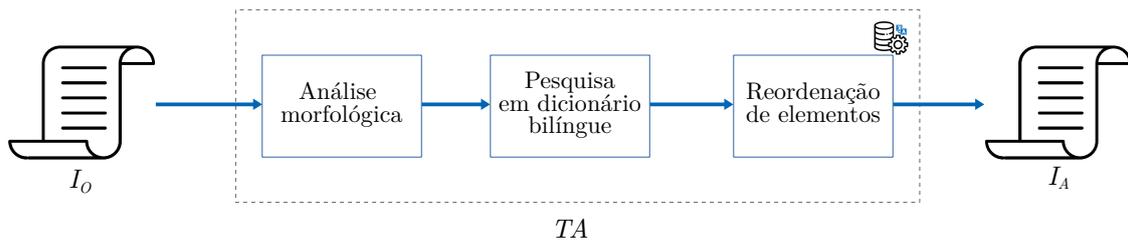


Figura 3.4 - Abordagem de tradução automática por sistema direto.

Sistemas de TA diretos possuem limitações como a tradução palavra-por-palavra, onde é realizada pesquisa em dicionário bilíngue e posteriormente realocação do posicionamento dos termos na frase em  $I_A$ . As limitações se dão pelo simples fato da estratégia direta não fazer uso da análise gramatical, desconsiderando a significação da sentença  $I_O$  (HAUENSCHILD; HEIZMANN, 2011).

#### 3.4.2 Sistema indireto interlíngua

No sistema indireto **interlíngua** é realizada análise unidirecional do texto original  $I_O$ , gerando representações que são utilizadas quando da geração do texto alvo  $I_A$ . Por ser de caráter unidirecional não-reversível, a análise não consulta novamente  $I_O$ . É criada projeção de  $I_O$ , formando novo texto fonte projetado  $I_{Op}$ , que não é tomado diretamente como fonte para geração de  $I_A$ , porém será utilizado como base

(DAVE et al., 2001). A estratégia interlíngua é utilizada várias vezes em sistema de TA multilínguas. Métodos interlíngua geram mapas em  $I_{Op}$ , correlacionando os idiomas  $I_1$  e  $I_2$ , que posteriormente são utilizados na tradução entre os idiomas analisados e gerados, como ilustra a Figura 3.5.

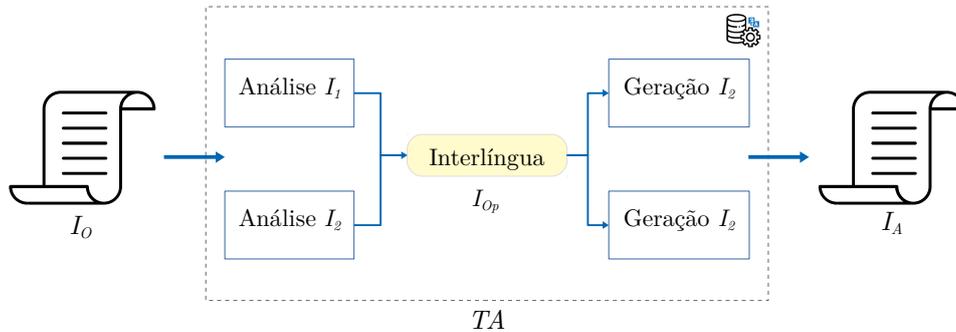


Figura 3.5 - Abordagem de tradução automática para o sistema interlíngua de dois idiomas.

Os blocos de análise e geração funcionam isoladamente para cada linguagem. Uma das vantagens desta abordagem é o fato de que a inclusão de novo idioma no sistema gera acréscimo de pares de possibilidades de tradução, sendo que este número de pares cresce em escala exponencial a cada novo idioma inserido (LEAVITT et al., 1994). Em sistema bilíngue originalmente com os idiomas  $S_{BL} = \{I_1 \leftrightarrow I_2\}$ , ao acrescentar terceiro idioma  $I_3$ , aumenta sua capacidade para sistema multilíngue  $S_{ML} = \{I_1 \leftrightarrow I_2, I_1 \leftrightarrow I_3, I_2 \leftrightarrow I_3\}$  e assim sucessivamente.

### 3.4.3 Sistema indireto por transferência

No modelo de TA indireto por **transferência**, a tradução é realizada através da criação de blocos intermediários de representação entre dois idiomas. Este método é dotado de três módulos separados: análise, transferência e geração. Sistemas por transferência geram representações dependentes entre cada par de idioma, onde são geradas duas correlações de transferência para cada par de idioma (WILKS, 2008).

Na Figura 3.6 são utilizados idiomas  $I_1$ ,  $I_2$  e  $I_3$ . Para cada idioma é criada representação do idioma  $I_1$  com os outros dois idiomas, gerando por combinação, seis módulos de transferência de  $I_1 \rightarrow I_2$ ,  $I_1 \rightarrow I_3$ ,  $I_2 \rightarrow I_1$  e assim sucessivamente. O acréscimo de novos idiomas elevaria a quantidade de módulos de transferência, tendo seu quantitativo total representado por  $n \cdot (n - 1)$  (WILKS, 2008). Apesar do considerável aumento de módulos, o sistema ainda apresenta-se mais simples que o modelo

interlúngua (Figura 3.5) por não necessitar criar o complexo mapa  $I_{Op}$ .

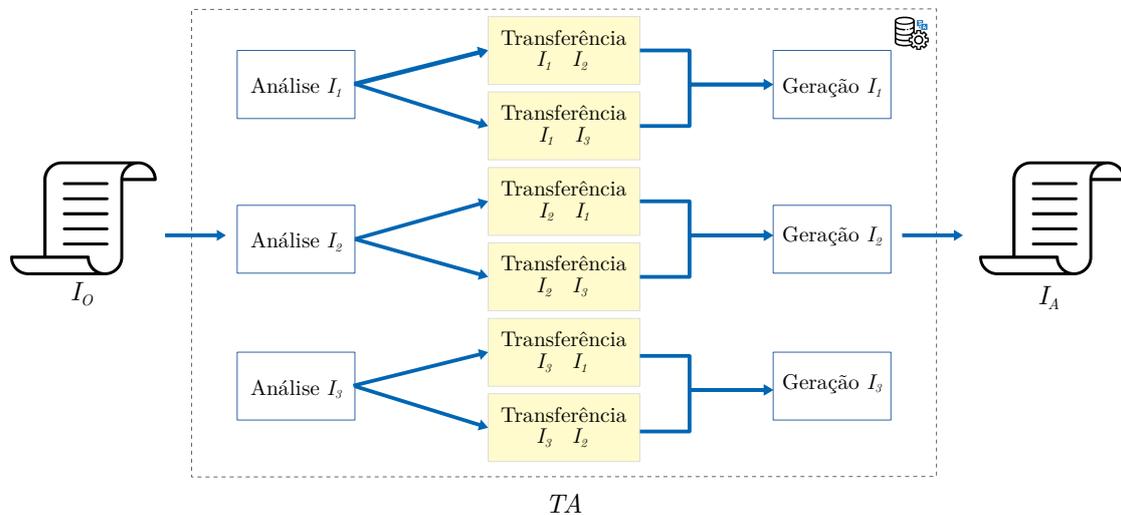


Figura 3.6 - Abordagem de tradução automática para o sistema de transferência para três idiomas.

### 3.4.4 Abordagem não-interativa, interativa e uso de dados auxiliares

Além dos modelos direto e indireto de TA, podem ser acoplados outros mecanismos que interferem diretamente na qualidade final das traduções. Nos primeiros sistemas de TA criados utiliza-se somente abordagem não-interativa, devido à limitação de interface computador/usuário. Em abordagens não-interativas o processamento de tradução é realizada em lotes. Não há nenhum tipo de intervenção do operador até que o processo de tradução produza o resultado final em  $I_A$  (HUTCHINS, 1986). O funcionamento do sistema não-interativo é dado pelo modelo ilustrado na Figura 3.1.

Com o avanço tecnológico e melhoria das interfaces computador/usuário são criados sistemas onde o usuário pode interagir diretamente com o sistema de TA, a fim de conduzi-lo no processo da tradução. Nesta abordagem há interação direta, em tempo real, com o operador do sistema (HOANG et al., 2018). O computador realiza questionamentos ao usuário, tais como: seleção de alternativas, complementação linguística, desambiguação de termos, identificação da classificação do tópico, dentre outros.

Após a série de questionamentos e processamento interno, o sistema dá como saída o texto final  $I_A$ , como ilustra a Figura 3.7. Apesar da existência da abordagem interativa, os sistemas de TA da atualidade são considerados sistemas não-interativos,

pois internamente tratam a tradução de blocos textuais ou de sentenças inteiras sem questionar o operador.

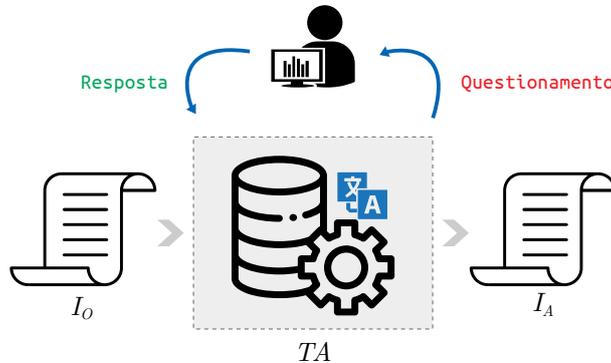


Figura 3.7 - Abordagem de tradução automática para o sistema interativo.

Além consultar dicionário digital, os sistemas de TA contam com a possibilidade de consultar outros tipos de dados auxiliares. O uso da gramática se dá na consulta das regras de construção-frasal, classificação de palavras, combinações e sequência aceitável de palavras nos módulos de análise e de geração dos idiomas  $I_O$  e  $I_A$ . Como apoio adicional podem ser consultados dados léxicos. Os dados léxicos podem conter informações adicionais específicas sobre palavras tanto no idioma  $I_O$  como em  $I_A$ . Em todos os sistemas, os dados léxicos contam com determinados termos no idioma  $I_O$  e seus equivalentes no idioma alvo  $I_A$ . Os dados léxicos são utilizados no módulo de geração (GUZMÁN, 2008).

### 3.5 Paradigmas de tradução automática

Existem paradigmas diferentes quanto à construção de sistemas de tradução automática. Os paradigmas dividem-se em: TA baseada em conhecimento e TA orientada por dados. Os sistemas baseados em conhecimento são regras pré-estabelecidas, enquanto os sistemas orientados por dados são conduzidos por informações previamente compiladas e categorizadas (BHATTACHARYYA, 2015). A Figura 3.8 ilustra a subdivisão interna de cada tipo de paradigma de TA.

#### 3.5.1 Tradução automática baseada em regras

Sistemas de Tradução Automática Baseados em Regras (TABR) dependem diretamente de dados linguísticos como: dicionários morfológicos, dicionários bilíngues, gramáticas e regras estruturais de transferência (FORCADA et al., 2011). Sistemas ba-

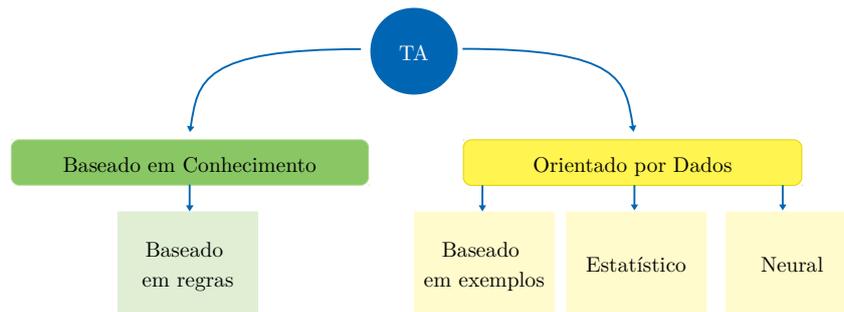


Figura 3.8 - Paradigmas de tradução automática.

seados em regras tendem a produzir traduções mecanizadas e altamente repetitivas (GUZMÁN, 2008).

Os sistemas TABR criam textos sem fluidez, gerando retrabalho de conferência e suavização por parte de revisores e editores. A Figura 3.9 ilustra o sistema com diferentes subdivisões e funcionalidades dentro dos passos de análise, transferência e sintetização. As variações ocorrem em cada sistema específico e não obedecem categorias gerais (SIMARD et al., 2007).

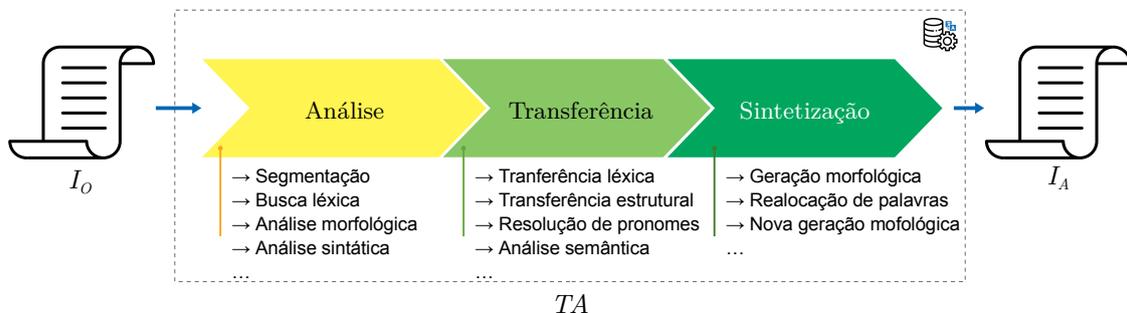


Figura 3.9 - Paradigma de tradução automática baseada em regras.

### 3.5.2 Tradução automática orientada por dados

Como ilustrado na Figura 3.8, sistemas de Tradução Automática Orientados por Dados (TAOD) podem ser classificados pela natureza dos dados utilizados, como: i) baseados em exemplos, ii) dados baseados em estatísticas e iii) dados baseados Rede Neural Artificial (RNA). Para os três tipos de TAOD é necessário a coleta prévia, categorização e preparação dos dados a serem utilizados durante o processo de tradução.

### 3.5.2.1 Tradução automática baseada em exemplos

Algoritmos baseados em exemplos tem funcionamento simples. O processo de tradução é todo baseado em estruturas frasais previamente conhecidas e armazenadas entre os idiomas  $I_O$  e  $I_A$ . Este tipo de TA é construída sobre a ideia de que a tradução pode ser realizada por analogia (FOURNIER, 2008). O método de tradução inicia seu processo dividindo o texto origem  $I_O$  em frases, em seguida traduzindo estas frases de forma análoga às frases similares à que é previamente treinado (NAGAO, 1984).

Levando em consideração o trecho de texto origem  $I_O$  e base de dados bilíngue  $D_{BL}$ , são gerados e armazenados textos  $I'_O$ . Após a geração do textos  $I'_O$ , são realizadas comparações deste texto com texto no idioma alvo parcial  $I'_A$ . Após a geração de opções de  $I'_A$ , o texto original  $I_O$  é comparado com as opções  $I'_O$  alinhadas a  $I'_A$  e candidato a tradução que atestar melhor correspondência, determinado por métrica, é armazenado no idioma alvo  $I_A$ , como ilustra a Figura 3.10 (NIRENBURG et al., 1993).

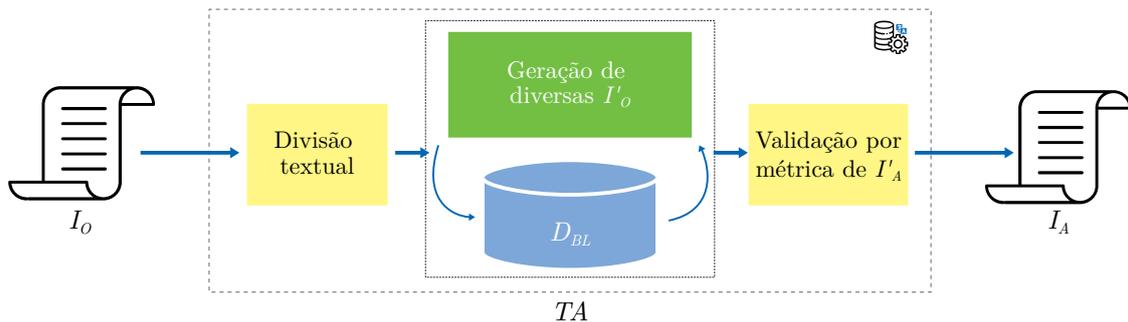


Figura 3.10 - Paradigma de tradução automática baseada em exemplo.

### 3.5.2.2 Tradução automática estatística

A abordagem estatística para tradução automática consiste em considerar possíveis opções de tradução de determinada sentença. Para cada par de texto origem/alvo ( $I_O, I_A$ ) é associada probabilidade  $P(I_O|I_A)$ . Esta probabilidade é interpretada como a possibilidade do sistema de TA produzir tradução alvo  $I_A$ , tendo como fonte a origem  $I_O$ . Utilizando o teorema de Bayes, o método procura estabelecer como tradução de  $I_A$  em  $I_O$ , aquela que apresentar maior probabilidade (BROWN et al., 1990). Em outras palavras, o intuito é maximizar o valor de  $P(I_O|I_A)$  a fim de encontrar tradução mais adequada. O mecanismo de verificação da probabilidade é dado por:

$$P(I_O|I_A) = \frac{P(I_O) \cdot P(I_A|I_O)}{P(I_A|I_A)} \quad (3.2)$$

Em (3.2), o denominador é independente de  $I_O$ . Possibilitando que  $I_A$  maximize o produto de  $P(I_O) \cdot P(I_A|I_O)$ . O primeiro fator deste produto é a probabilidade de  $I_O$ , enquanto o segundo fator é a probabilidade de  $I_A$  dado  $I_O$  (BROWN et al., 1990). O sistema de TA estatística necessita de método para computar a probabilidade do modelo de linguagem, o método para cálculo das probabilidades de tradução e por fim, método para vasculhar as possíveis sentenças de  $I_O$  que produzam maior valor para  $P(I_O) \cdot P(I_A|I_O)$ , como ilustra a Figura 3.11.

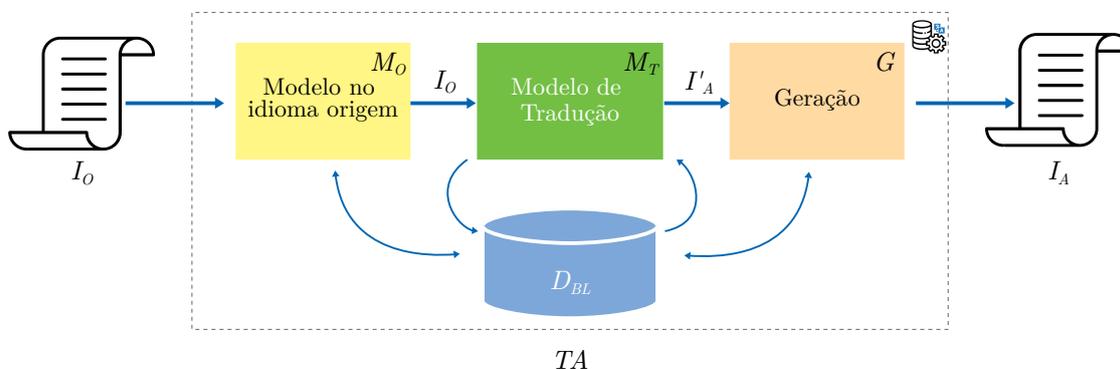


Figura 3.11 - Paradigma de tradução automática estatística.

No sistema de TA estatística ilustrado na Figura 3.11, os blocos  $M_O$  e  $M_T$  fornecem a distribuição de probabilidade sobre o par de sentença  $(I_O, I_A)$ , enquanto que o bloco da geração  $G$  efetua a tradução, dando como resultado o texto no idioma alvo  $I_A$ . O gerador  $G$  seleciona a sentença correspondente ao máximo valor da probabilidade  $P(I_O|I_A)$ . Todo o processo faz comunicação com base de dados bilíngue  $D_{BL}$  de conhecimento prévio.

### 3.5.2.3 Tradução automática Neural

Sistemas de Tradução Automática Neural (TAN) são RNA que modelam diretamente a probabilidade condicional  $P(I_A|I_O)$  de tradução de texto origem,  $I_{O1}, \dots, I_{On}$ , ao texto alvo  $I_{A1}, \dots, I_{An}$  (LUONG et al., 2015). O sistema TAN é composto por dois componentes: i) codificador e ii) decodificador. O codificador calcula a representação  $s$  para cada sentença origem, enquanto o decodificador gera uma palavra chave por vez para cada  $s$  do texto origem. A decomposição dos termos

da sentença, através do decodificador gera desdobramento da probabilidade, dada por:

$$\log P(I_A|I_O) = \sum_{j=1}^m \log P(I_{A_j}|I_A < j, s) \quad (3.3)$$

A decomposição do sistema possibilita a utilização da arquitetura de Redes Neurais Recorrentes (RNR) (KALCHBRENNER; BLUNSOM, 2013). A Figura 3.12, adaptada de Luong et al. (2015), ilustra a organização conceitual da TAN que apresenta a estrutura básica da RNA, dividida entre codificador e decodificador. A região de abrangência do codificador trata do treinamento da rede sobre os termos  $o_1, \dots, o_n$  da sentença origem, sendo a região pós-conectada do decodificador, responsável pela geração decodificada da sentença no idioma alvo, representado pela sequência  $a_1, \dots, a_n$  (LUONG et al., 2015).

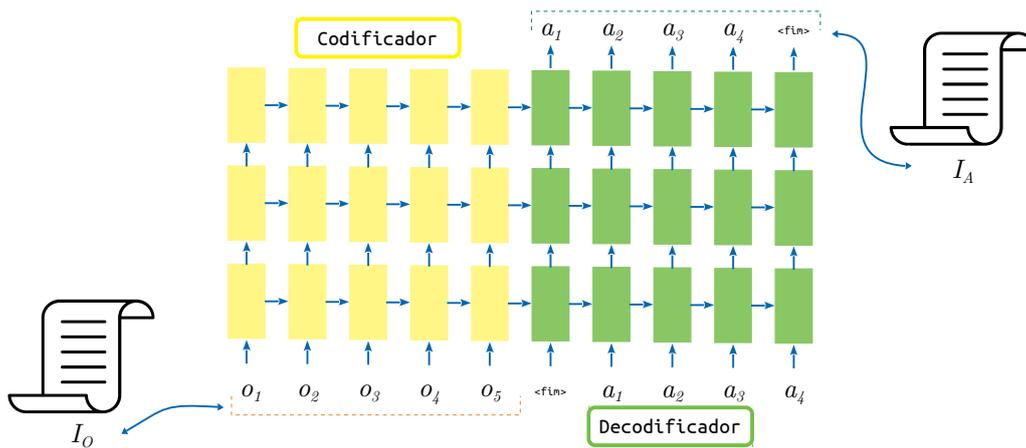


Figura 3.12 - Paradigma de tradução automática Neural.

### 3.6 Métricas de avaliação

Para verificação da qualidade e fidelidade das traduções realizadas, os textos resultantes do processo de tradução devem ser avaliados quanto à correspondência entre os idiomas origem  $I_O$  e idioma alvo  $I_A$ . O método mais eficaz, porém mais oneroso, é a utilização de tradutor humano para avaliar o resultado da tradução. Existem várias técnicas de avaliação para traduções digitais, sendo as principais e mais utilizadas: *Bilingual Evaluation Understudy* (BLEU), *National Institute of Standards and Technology* (NIST) e *Metric for Evaluation of Translation with Explicit Orde-*

*ring* (METEOR) (DORR et al., 2018). Estas métricas são do tipo automatizadas e não contam com interação humana no processo de avaliação.

### 3.6.1 Bilingual Evaluation Understudy – BLEU

O sistema de avaliação BLEU é considerado como a métrica padrão para avaliação da tradução de texto digital. O método consiste na utilização de exploração direta de múltiplas referências. São necessários dois itens para que este sistema seja aplicado: i) capacidade de estimar numericamente a similaridade da tradução e ii) base de dados de traduções de referência (BABYCH, 2014). A pontuação atribuída por este algoritmo é calculada pela contagem do número de letras ou sequência de palavras  $n$ -grama. Cada  $n$ -grama representa a menor parte analisada no decorrer do processo de atribuição de pontuação (PAPINENI et al., 2002). A Figura 3.13 ilustra a divisão da sentença de referência  $S_r$ , através da contagem de cada palavra da frase base de comparação.

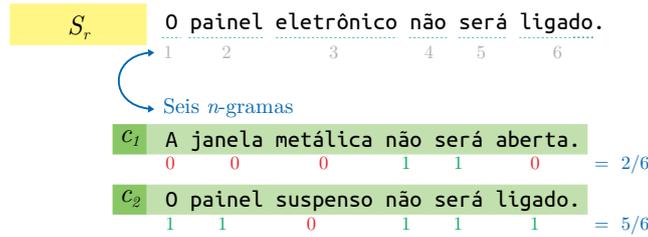


Figura 3.13 - Divisão de sentença por  $n$ -gramas.

No exemplo da Figura 3.13 existem duas sentenças candidatas e uma tradução de referência,  $c_1$  e  $c_2$  com precisões  $p_n = 2/6$  e  $p_n = 5/6$ , respectivamente, na qual a precisão  $p_n$  é dada por:

$$p_n = \frac{\sum_{c \in \{Can\}} \sum_{n\text{-grama} \in c} Cnt_{clip}(n\text{-grama})}{\sum_{c' \in \{Can\}} \sum_{n\text{-grama}' \in c'} Cnt(n\text{-grama}')} \quad (3.4)$$

onde  $p_n$  é o valor da precisão para as partes  $n$ -gramas de tamanho  $n$ ,  $c$  é o candidato avaliado,  $Can$  é o conjunto de sentenças de teste,  $Cnt(n\text{-grama})$  é o número de vezes que o  $n$ -grama ocorre no candidato,  $c'$  a sentença referência comparada,  $n\text{-grama}'$  é parte analisada na sentença de referência e  $Cnt_{clip}(n\text{-grama})$  é o mínimo

de ocorrências não cortadas e o máximo<sup>1</sup> de vezes que estas ocorrências aparecem na tradução de referência (DORR et al., 2018).

No auxílio do refinamento da produção do valor BLEU é utilizado uma **pena de brevidade** relacionando o tamanho entre candidados  $c$  e referências  $r$  de tradução. A pena de brevidade penaliza as sentenças que apresentam tamanho inferior às sentenças de referência e seu cálculo é dado por:

$$BP = \begin{cases} 1 & \text{se } c > r \\ e^{(1-\frac{r}{c})} & \text{se } c \leq r \end{cases} \quad (3.5)$$

O valor da pontuação de precisão BLEU é calculado por (3.6), no qual  $N$  é o número máximo de sequência de  $n$ -grama considerada, por definição como sendo  $1 \leq N \leq 4$  e  $w_n$  os pesos uniformes para cada  $n$ -grama considerada para dado  $N$ , podendo ocorrer variação  $0 \leq w_n \leq 1/N$  (KOEHN; MONZ, 2006). O valor da precisão BLEU pode variar entre 0 e 1, onde o valor 1 é considerado a máxima precisão.

$$P_{BLEU} = BP \cdot \exp \sum_{n=1}^N w_n \log p_n \quad (3.6)$$

### 3.6.2 *National Institute of Standards and Technology* – NIST

A métrica de qualificação de traduções NIST é desdobramento do método BLEU (FINCH et al., 2004). Este método utiliza abordagem teórica de informações baseada nos  $n$ -gramas, mas neste tipo de técnica específico, os  $n$ -gramas considerados variam de 1 até 5, diante do máximo valor  $N = 4$  para o método BLEU (DODDINGTON, 2002). Diferentemente do método BLEU, onde é calculada média geométrica, nesta métrica é calculada média aritmética dos  $n$ -gramas com peso atribuído por contribuição específica de cada  $n$ -grama.

A pontuação gerada nesta métrica representa a informação média por cada palavra, dado por cada  $n$ -grama da tradução que corresponde à referência no conjunto da base de dados de referências. O algoritmo NIST penaliza as traduções mais curtas, enquanto que sentenças mais próximas às das sentenças na fonte de referência são menos penalizadas se comparadas à penalidade de brevidade  $BP$  usada no método BLEU.

---

<sup>1</sup>Ocorre o truncamento da contagem de palavras, se necessário, para não exceder a maior contagem encontrada para única referência para aquela palavra.

A pontuação de cada candidato à tradução é dada por:

$$P_{NIST} = BP \cdot \sum_{n=1}^N \sum_{w_1, \dots, w_n \in C} \frac{Info(w_1, \dots, w_n)}{Cnt(w_1, \dots, w_n)} \quad (3.7)$$

$$Info(w_1, \dots, w_n) = \log_2 \left( \frac{\text{número de ocorrências de } w_1, \dots, w_{n-1}}{\text{número de ocorrências de } w_1, \dots, w_n} \right) \quad (3.8)$$

$$BP = \begin{cases} 1 & \text{se } c > r \\ \exp \left( \beta \ln^2 \left( \frac{c}{r} \right) \right) & \text{se } c \leq r \end{cases} \quad (3.9)$$

onde  $\beta$  é escolhido para forçar o valor de  $BP = 0,5$  quando o número de palavras no sistema de saída for até  $2/3$  da média de palavras na tradução de referência (FINCH et al., 2004). Devido à diferença no tratamento da penalidade de brevidade, onde é considerado detalhes diferentes do que é realizado no método BLEU, o método NIST apresenta pontuação de qualidade de tradução mais precisa como julgamento de fluência (DODDINGTON, 2002).

### 3.6.3 *Metric for Evaluation of Translation with Explicit Ordering – METEOR*

A métrica METEOR é outra variante do algoritmo BLEU. Este modelo de pontuação de tradução fundamenta-se no ajuste e correção de vários pontos falhos identificados no modelo BLEU (BANERJEE; LAVIE, 2005). Enquanto a métrica BLEU é orientada por precisão, o método METEOR é orientado por chamada. Para atribuição de pontuação é calculada precisão  $P$  e a chamada  $R$ , ambos valores são combinados em média harmônica, dado por:

$$F_m = \frac{P \cdot R}{\alpha P + (1 - \alpha)R} \quad (3.10)$$

Com propósito de aproximar o julgamento da métrica ao julgamento humano, o valor padrão de  $\alpha$ , que por definição é  $\alpha = 0,9$ , pode ser otimizado. Como complemento dos termos componentes do cálculo da avaliação a penalidade  $P_{MT}$  é dada por:

$$P_{MT} = \gamma \cdot \left( \frac{n_b}{n_{n\text{-gramas}}} \right)^\phi \quad (3.11)$$

onde  $n_b$  é o número de sequências encontradas de dois ou mais  $n$ -gramas juntapostos,  $n_{n\text{-gramas}}$  a quantidade de  $n$ -gramas relacionados isoladamente entre a tradução candidata e a referência e por definição,  $\gamma = 0,5$  e  $\phi = 3,0$ . A pontuação final do algoritmo METEOR é dada por:

$$P_{METEOR} = F_m \cdot (1 - P_{MT}) \quad (3.12)$$

Como resultado final de sua classificação, o método METEOR gera valor de pontuação para cada candidata a tradução separadamente. Em seguida, escolhe aquela candidata a tradução que tiver obtido maior pontuação. Comparado ao sistema BLEU, o sistema METEOR não faz diretamente a verificação de múltiplas fontes de referência de uma só vez. Assim como a métrica NIST, esta métrica apresenta maior similaridade ao julgamento humano quanto à avaliação de traduções (LAVIE; DENKOWSKI, 2009).

### 3.7 Considerações finais

Neste capítulo foram apresentados conceitos de tradução automática, explanando as principais características, métodos de tradução direta, indireta e por transferência. Foram abordados os métodos não-interativos, interativos e de utilização de dados auxiliares, apresentando vantagens e desvantagens de cada método. Ao final, foi abordada a categoria de métricas de qualidade para os resultados gerados por tradução automática. As metodologias de tradução e as métricas de avaliação de tradução são utilizadas como base para a junção das técnicas de TA nos métodos de otimização apresentados no próximo capítulo.



## CAPÍTULO 4

### PROCESSO DE OTIMIZAÇÃO

Este capítulo apresenta a contextualização teórica e a aplicação de processos de otimização. São apresentados métodos de otimização heurístico e determinístico, apresentando as principais características, vantagens e desvantagens de aplicação de cada método de otimização.

#### 4.1 Otimização

Otimização pode ser formulada de formas diversas e pode atender à resolução de diversos tipos de questionamentos ao buscar soluções para problemas reais através de cálculos computacionais. Partes dos problemas reais envolvem utilização de algum tipo de otimização (KOZIEL; YANG, 2011). O processo de otimização é a metodologia utilizada para encontrar melhor solução para problemas como: redução de custos, maximização de produção, redução de gasto energético, aumento de lucro, desempenho e vários outros (ALTIPARMAK et al., 2006).

A otimização pode ser definida como grupo de procedimentos executados na busca do melhor desempenho de dado processo, obtendo assim a melhor função objetivo. A função objetivo, também chamada função de avaliação, pode ser única ou composta por restrições específicas de igualdade, desigualdade e restrições laterais, definidas por:

$$\begin{aligned} \max f(x) \quad \text{ou} \quad \min f(x) \\ \text{Sujeito à :} \\ g(x) \leq 0 \\ h(x) = 0 \end{aligned} \tag{4.1}$$

no qual  $x$  é o vetor de possíveis soluções,  $f(x)$  é a função objetivo,  $g(x)$  a função restrição de desigualdade,  $h(x)$  a função restrição de igualdade.

Na forma básica o processo de otimização possui os componentes: otimizador e simulador, como ilustrado na Figura 4.1. A representação matemática é a modelagem numérica e equacional do problema físico real e esta representação pode oferecer numericamente a solução ao problema abordado. O processo de modelagem matemática do problema é a parte crucial em problemas de otimização (MAGALHAES et

al., 2020).

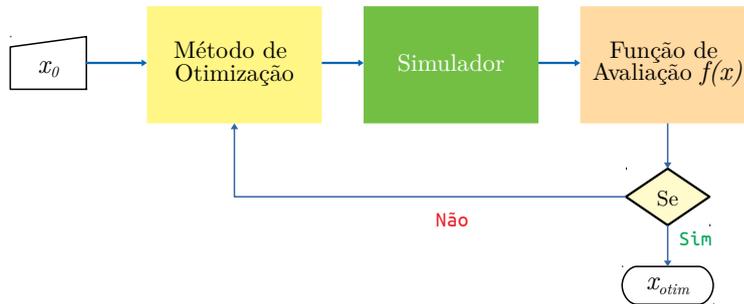


Figura 4.1 - Modelo do processo de otimização.

Na Figura 4.1  $x_0$  são os valores iniciais de  $x$  e  $x_{otim}$  são os valores ótimos ou otimizados de  $x$ , que levam ao menor valor da função de avaliação (no caso da minimização). Funções objetivas são, na maioria dos problemas, multivaloradas, apresentando diferentes valores ótimos ou otimizados como solução de determinado problema. Os valores de  $x_{otim}$  encontrados podem caracterizar-se como valores ótimos locais ou ótimos globais, como ilustra a Figura 4.2. Problemas que apresentam configuração similar ao exemplo da Figura 4.2 são de difícil solução, pois ao buscarem solução, podem acabar presos dentro de ótimos locais, sem condições de alcançar os ótimos globais (CALIXTO et al., 2015).

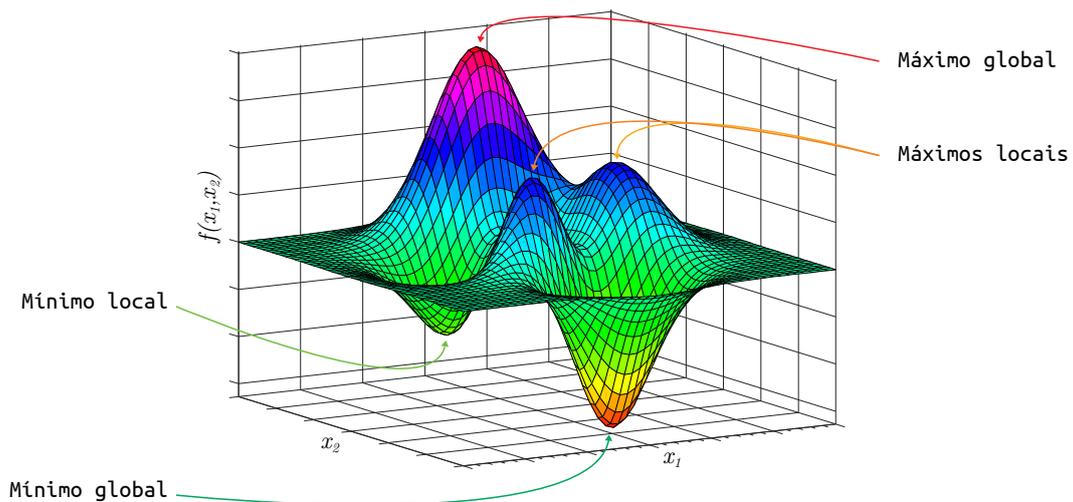


Figura 4.2 - Máximos e mínimos locais e globais para função objetivo  $f(x_1, x_2)$ .

## 4.2 Métodos de Otimização

Entre as classes definidas para os métodos de otimização existem duas mais utilizadas: determinística e heurística. Cada classe adota abordagem distinta e com variações de modelagem e considerações, como presença de semente inicial de candidatos a solução, quantidade de iterações, dentre outros aspectos.

### 4.2.1 Métodos Determinísticos

Os métodos determinísticos de otimização tem seu funcionamento baseado na utilização de técnicas analíticas aplicadas ao modelo matemático, a fim de minimizar o valor da função objetivo. É método que dada semente inicial de possível solução, considerando que seus passos são previamente conhecidos, o resultado final apresentado é sempre o mesmo (FLOUDAS, 1997).

Como aspecto de desvantagem, este método, na maioria das vezes, requer conhecimento da primeira derivada da função de avaliação em relação às variáveis do problema. Métodos determinísticos garantem sempre a convergência para o valor de solução ótimo, mas que não é necessariamente o valor ótimo global para solução do problema. Esta classe de abordagem utiliza o gradiente da função objetivo, matriz Jacobiana ou matriz Hessiana. Partindo deste pressuposto, o método parte em busca do valor de máximo ou de mínimo da função objetivo (KOZIEL; YANG, 2011).

O método determinístico pode ser classificado em problemas de programação linear e problemas de programação não-linear, podendo haver variações destas duas classes através de métodos mistos. Problemas de programação linear buscam encontrar solução através da prerrogativa de que a função objetivo e todas as restrições são dadas como funções lineares baseadas nas variáveis do problema (OLIVERI et al., 2011), enquanto problemas de programação não-linear trata da solução de problemas onde a função de avaliação ou pelo menos uma das funções de restrição seja não-linear (HAFTKA; WALSH, 1992).

#### 4.2.1.1 Método Nelder-Mead Simplex

O método de otimização de Nelder-Mead utiliza técnicas analíticas na obtenção de valores mínimos para a função objetivo. Para obtenção deste mínimo, o método minimiza funções não-lineares sem basear sua busca nas informações das derivadas e sem restrições, sendo classificado como método de busca irrestrita (LAGARIAS et al.,

1998). O método é estruturado pela utilização do simplex<sup>1</sup>. Este tipo de abordagem busca a solução ótima para dado problema com  $k$  variáveis com a função a ser minimizada variando suavemente (DANTZIG, 1987). O conceito trata de polítopo<sup>2</sup> contido no espaço  $\mathbb{R}^\delta$ , com  $\delta + 1$  vértices em  $\delta$  dimensões.

No decorrer da busca pelo valor mínimo, o pior vértice do poliedro de  $\delta + 1$  vértices é substituído por novo vértice colinear ao antigo vértice (WRIGHT et al., 2010). A cada iteração os  $\delta + 1$  vértices do simplex  $x_1, x_2, \dots, x_{\delta+1} \in \mathbb{R}^\delta$  são ordenados com o crescimento dos valores de  $f(x_i)$ , dado por:

$$f(x_1) \leq f(x_2) \leq \dots \leq f(x_{\delta+1}) \quad (4.2)$$

onde  $x_1$  e  $x_{\delta+1}$  representam o melhor e o pior vértice, respectivamente.

Quatro coeficientes são considerados quando do reposicionamento dos vértices: o coeficiente de reflexão  $\rho$ , o coeficiente de expansão  $\chi$ , o coeficiente de contração  $\gamma$  e o coeficiente de redução  $\sigma$ , os quais devem obedecer as seguintes condições:  $\rho > 0$ ,  $\chi > 1$ ,  $0 < \gamma < 1$  e  $0 < \sigma < 1$  (NELDER; MEAD, 1965). A Figura 4.3 ilustra as projeções utilizando cada um dos coeficientes de reposicionamento dos vértices para as iterações, onde  $x_1, x_2$  e  $x_3$  são os vértices iniciais,  $x_c$  o vértice da contração,  $x_r$  o vértice da reflexão e  $x_e$  o vértice de expansão. O novo vértice é obtido por expansão, reflexão ou contração do pior vértice ao longo da reta que passa pelo novo vértice.

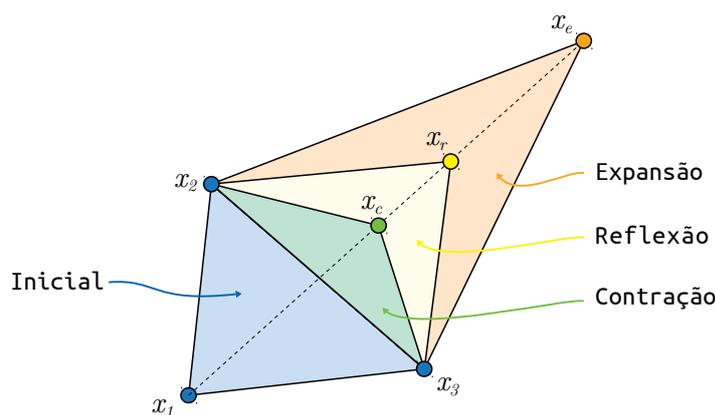


Figura 4.3 - Ilustração dos coeficientes utilizados no método de Nelder-Mead.

<sup>1</sup>Geometria gerada na construção do polítopo com número de dimensões igual à quantidade de variáveis do problema. A solução ótima sempre será o conjunto de coordenadas de um dos vértices deste polítopo.

<sup>2</sup>Região contida em  $\mathbb{R}^\delta$  que é resultante da intersecção do conjunto de semi-espacos.

Cada nova iteração executa a substituição do pior vértice ou reduz em torno do melhor vértice. O centróide dos  $\delta$  vértices é dado por:

$$\bar{x} = \sum_{i=1}^{\delta} \frac{x_i}{\delta} \quad (4.3)$$

na qual  $x_i$  é o vértice. O critério de parada do método é dado por:

$$\sqrt{\sum_{i=1}^{\delta+1} \frac{[f(x_i) - f(\bar{x})]^2}{\delta}} \leq \eta \quad (4.4)$$

onde  $\eta$  representa a tolerância de erro.

## 4.2.2 Métodos Heurísticos

Métodos de otimização heurísticos são algoritmos exploratórios que buscam por soluções independentemente do conhecimento específico relacionado ao problema abordado, podendo ser classificados como busca cega (CALIXTO et al., 2011). O valor ótimo do problema por vezes não é o alvo principal dos métodos heurísticos, uma vez que, tendo como ponto de partida uma possível solução, são realizadas sucessivas aproximações, direcionado ao ponto ótimo. Desta forma, estes métodos podem encontrar a solução ótima ou otimizada. A subjetividade deste modelo é peculiaridade análoga à inteligência humana. Por se aproximar da solução ótima, este método é indicado em alguns casos onde outras abordagens determinísticas falham por se prenderem no ótimo local ou mesmo por não encontrar solução satisfatória (CALIXTO et al., 2010b).

### 4.2.2.1 Algoritmo genético

Baseado no mecanismo de seleção natural, este método utiliza conceitos da evolução e seleção natural das espécies, atrelado à geração randômica de valores, para encontrar a solução otimizada do problema. A principal característica de determinado indivíduo observada por Charles Darwin é sobrevivência do mais apto. Baseados nos conceitos evolutivos de Darwin, algoritmos genéticos combinam conceitos como seleção, troca de material genético e sobrevivência dos mais capazes (HOLLAND et al., 1992).

Comparados a outros métodos de busca, pode-se indicar as características do algoritmo genético: i) baseado em conjunto de possíveis soluções, ii) baseia-se na infor-

mação da função objetivo e não da derivada, iii) apresenta conjunto de candidatos a solução ao invés de candidato único e iv) usa regras de transição probabilística na busca do espaço de soluções em detrimento de regras determinísticas (KOZA, 1997).

O algoritmo genético clássico conta com a geração aleatória da população inicial, simulação, avaliação, manipuladores genéticos e operadores genéticos, como ilustra a Figura 4.4. A população inicial, gerada aleatoriamente, é o primeiro conjunto de candidatos à solução do problema, formada pela matriz  $m \times n$ , na qual  $m$  é o número de indivíduos na população e  $n$  é o número de genes  $G_j$  em cada indivíduo  $I_i$  (número de parâmetros a serem otimizados) (CALIXTO et al., 2010). O cromossomo é o conjunto de genes que forma um indivíduo. O gene é a menor unidade de informação do cromossomo.

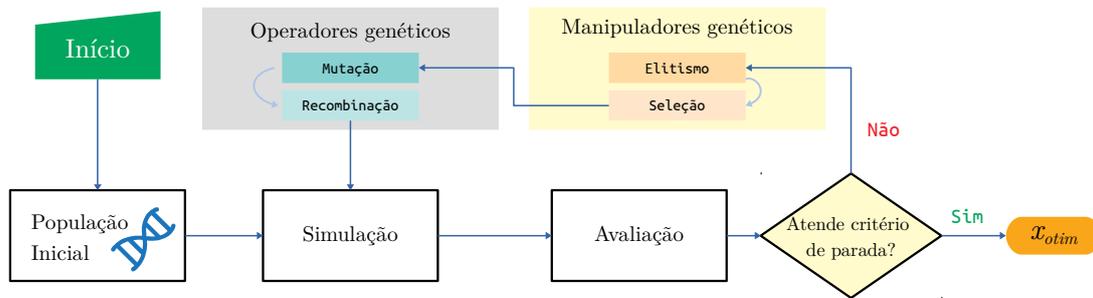


Figura 4.4 - Fluxograma do algoritmo genético clássico.

Dado o problema a ser otimizado, cada indivíduo  $I_i$  é um cromossomo. A população representada pela matriz  $m \times n$ , de  $m$  indivíduos e  $n$  genes, é ilustrada na Figura 4.5. O tamanho  $m$  da população é definido com certo crit rio para que n o ocasione lentid o no processo de otimizaç o quando numeroso, ou para que n o produza populaç o com baixo  ndice de qualidade por possuir diversidade reduzida (ALTIPARMAK et al., 2006).

A funç o de avaliaç o  $f(x)$  atribui valor de aptid o para cada indivíduo, avaliando cada um quanto   sua capacidade de solucionar o problema. Levando em consideraç o o car ter universal dos algoritmos genéticos, a funç o de avaliaç o   a ligaç o mais pr xima com a natureza geral do problema abordado. Quanto mais precisa na caracterizaç o do problema, melhor   sua capacidade de gerar notas condizentes para cada indivíduo avaliado (CALIXTO et al., 2010a). No algoritmo genético, cada iteraç o   classificada como uma geraç o  $g$  que abarca o universo de indiv dos

	G <sub>1</sub>	G <sub>2</sub>	...	G <sub>n</sub>
I <sub>1</sub>	c <sub>11</sub>	c <sub>12</sub>	...	c <sub>1n</sub>
I <sub>2</sub>	c <sub>21</sub>	c <sub>22</sub>	...	c <sub>2n</sub>
⋮	⋮	⋮	⋮	⋮
I <sub>m</sub>	c <sub>m1</sub>	c <sub>m2</sub>	...	c <sub>mn</sub>

Figura 4.5 - Matriz populacional.

avaliados e manipulados.

No bloco de manipuladores genéticos ocorre a escolha/seleção dos progenitores para cada nova geração de população. No torneio, são escolhidos os indivíduos que apresentarem melhor nota de desempenho, passando assim a incorporarem a lista de progenitores para gerar novos indivíduos. Definindo o tamanho do torneio  $\tau$ , considerando  $2 \leq \tau \leq m$ , é utilizada variável aleatória  $\vartheta$ , onde  $1 \leq \vartheta \leq m$ . Desta maneira,  $\tau$  é a definição de quantos indivíduos são selecionados para o torneio e  $\vartheta$  define aleatoriamente quais indivíduos  $\tau$  serão selecionados na matriz populacional  $m \times n$  (KOZA, 1997).

A competição pelo direito de compor a nova geração populacional leva em consideração a nota de cada  $\tau$  indivíduo selecionado, sendo escolhidos aqueles com maior nota. Outro manipulador genético é o elitismo, que se encarrega de selecionar o melhor indivíduo da população e garante a participação deste na próxima geração. Isto possibilita que, no cenário menos favorável, o melhor indivíduo da geração anterior será o melhor indivíduo da próxima geração, no pior caso (CALIXTO et al., 2015).

A fase da aplicação dos operadores genéticos pode utilizar operador de cruzamento e operador de mutação. Dada a população em determinada iteração, é necessário a definição da taxa de cruzamento  $P_c$  e da taxa de mutação  $P_m$ . Tais taxas podem sofrer ajustes no decorrer de cada iteração ou serem determinadas aleatoriamente. O operador de cruzamento pode ser definido como mecanismo de troca de carga genética de exploração global (*exploration*), como ilustra a Figura 4.6, exemplo do operador de cruzamento uniforme operando com os indivíduos  $p_1$  e  $m_1$  selecionados como progenitores. O critério para a combinação dos genes dos progenitores se dá pelo vetor de variáveis de decisão binário probabilístico  $\phi$ , que neste caso utiliza o gene de  $p_1$  caso  $\phi = 0$  e o gene de  $m_1$ , caso  $\phi = 1$  (ARAUJO, 2020).

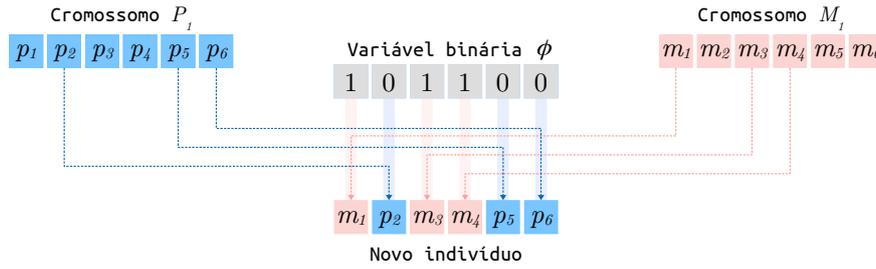


Figura 4.6 - Operador de cruzamento uniforme.

O operador de mutação realiza modificação na carga genética de cada gene  $c_{(i,j)}$ , realizando a exploração local (*exploitation*), visitando a vizinhança dos novos locais indicados pelo operador de cruzamento. O operador de mutação cria a diversificação genética na população selecionando e alterando aleatoriamente informações dos genes, possibilitando a busca de novos valores nas redondezas de cada gene, impedindo a estagnação do algoritmo em ótimos locais. Como operador de mutação, existe o operador de mutação adaptativo (CALIXTO et al., 2010b). Este operador atua aumentando ou diminuindo os valores das variações em  $c_{(i,j)}$  e da  $P_m$  de mutação com base nos valores de  $f(x)$  obtidos no decorrer das iterações, dado por:

$$P_m = v \cdot \frac{f(x_{max}) - f(x_{best})}{f(x_{max}) - f(x)} \quad \text{dado que } 0 < v \leq 1 \quad (4.5)$$

onde  $v$  é a constante relacionada à diversidade populacional,  $f(x_{max})$  o maior valor obtido para a função objetivo,  $f(x_{best})$  é o valor obtido pela função de avaliação para o melhor indivíduo da população atual e  $\overline{f(x)}$  a média dos valores  $f(x)$  para cada indivíduo da população da iteração atual.

### 4.3 Considerações finais

Neste capítulo foram apresentados conceitos sobre otimização, além da abordagem de ferramentas de otimização como algoritmo genético e o método de Nelder-Mead. Foram expostas as principais vantagens e desvantagens dos métodos e observada a natureza de funcionamento de cada um. Estes conceitos, são utilizados na validação da metodologia proposta, apresentada no próximo capítulo.

## CAPÍTULO 5

### METODOLOGIA

Neste capítulo será apresentada a metodologia utilizada para o desenvolvimento e a formatação do método proposto. Serão descritos os objetivos da pesquisa, processo de avaliação da tradução automática, criação da métrica para avaliação da tradução automática e a comparação da tradução final obtida por método de tradução automática tradicional.

#### 5.1 Contextualização

Com a globalização e a necessidades de comunicação entre povos de diferentes idiomas, surgiram ferramantes computacionais e dispositivos eletrônicos para facilitar a comunicação entre as pessoas. Porém, existem ainda limitações em ferramentas de tradução automática (TA). Na metodologia proposta estão estabelecidas novas técnicas para avaliação de TA existentes e a geração de textos traduzidos com qualidade. A metodologia proposta é composta por três blocos: i) tradução por processo de otimização, ii) avaliação da qualidade da tradução por função de avaliação da estrutura frasal e iii) ajuste de tradução utilizando processo de otimização.

A sequência de validação da metodologia proposta compreende os passos: i) definição da TA online que ofereça interface de programação de aplicativos (*Application Programming Interface* – API) de conexão ao *software* desenvolvido, ii) escolha do dicionário online com API, iii) escolha da métrica-base para avaliação das traduções candidatas, iv) definição da nova métrica proposta, v) avaliação dos textos traduzidos utilizando métrica-base e nova métrica e vi) avaliação dos resultados por tradutor humano. A Figura 5.1 ilustra o fluxo utilizado para desenvolvimento da metodologia proposta, na qual  $D_{API}$  é a API do dicionário,  $DIC_{EF}$  é o dicionário de estrutura frasal,  $EF$  é a estrutura frasal e o  $TA_{API}$  é a API de tradução automática.

#### 5.2 Escolha das interfaces de programação de aplicação

São necessárias as definições de duas API distintas, sendo a primeira API de TA que será utilizada como *software* de tradução. A segunda API constitui de dicionário de consulta para coleta de sinônimos e definições de palavras para alimentação da métrica proposta. A API de TA deve oferecer requisição de tradução através de outros *softwares*, de preferência, via requisição remota. O critério de escolha para esta API é a gratuidade de utilização das traduções. A Figura 5.2 ilustra o fluxo para aquisição de texto traduzido. O processo consiste no envio de chave de acesso à

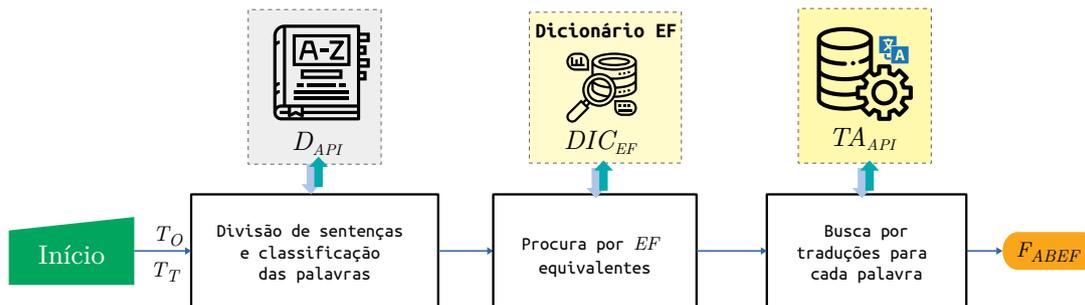


Figura 5.1 - Métrica de avaliação bilíngue por estrutura frasal proposta.

API, em seguida é encaminhado o texto original  $T_O$ , o idioma origem  $I_O$  e o idioma alvo  $I_A$ , enviados via parâmetros ao *software*. Na última etapa do processo o texto traduzido  $T_T$  para o idioma alvo é retornado pela TA.

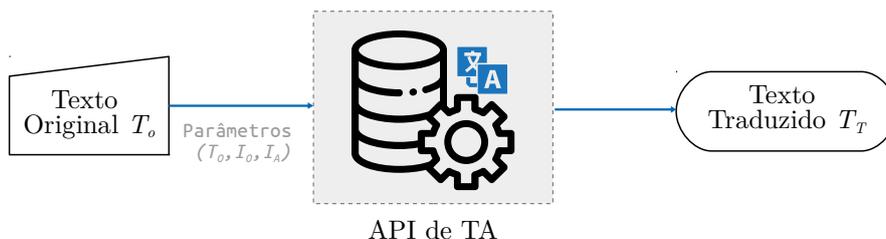


Figura 5.2 - Aquisição da tradução de texto.

A API de dicionário deve oferecer consulta nas palavras do idioma inglês com vasta base de dados. O critério de escolha para esta API também é a gratuidade e possibilidade de integração com outras API. A Figura 5.3 ilustra o processo de aquisição dos dados da palavra hipotética  $P$  consultada na base da API de dicionário. O parâmetro de entrada é a palavra  $P$  e o resultado obtido é o conjunto de informações sobre a palavra  $P$ , como: classe gramatical, variação e sinônimos. Este dicionário é definido por  $D_{API}$ .

### 5.3 Métrica para aferir a acurácia

Como métrica de aferição da acurácia das traduções obtidas é utilizada a métrica *Bilingual Evaluation Understudy* (BLEU). Por ser a métrica mais difundida na comunidade científica, a métrica BLEU é utilizada como base de comparação para os resultados das traduções obtidas. A Figura 5.4 ilustra a sequência de passos utilizados no *software* para obtenção da tradução, no qual o bloco tradução representa o

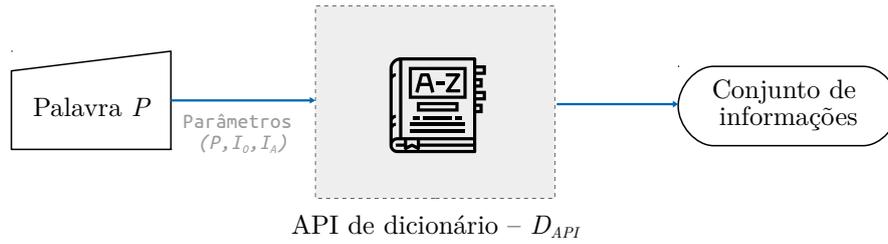


Figura 5.3 - Aquisição da classificação de palavra.

mecanismo de tradução  $TA_{API}$  e o bloco de avaliação representa a métrica BLEU.

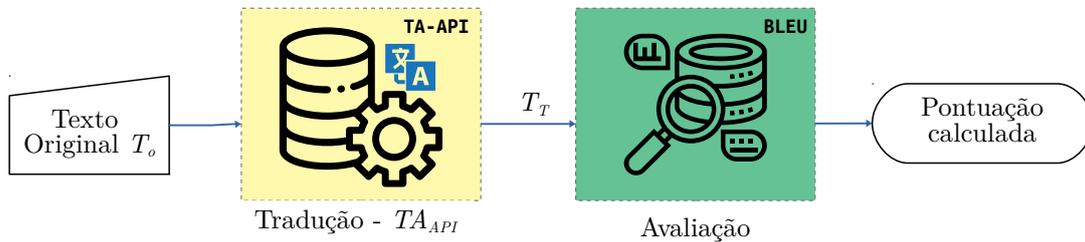


Figura 5.4 - Avaliação da tradução utilizando métrica BLEU.

#### 5.4 Avaliação de tradução pela métrica proposta

A métrica de avaliação da tradução textual proposta é denominada **Avaliação Bilingue por Estrutura Frasal** – (ABEF). O método é dividido em dois blocos, no qual o primeiro utiliza como parâmetros iniciais a frase no idioma origem  $T_O$  e a frase traduzida  $T_T$ . De posse destes parâmetros, é realizada consulta da estrutura frasal para cada frase juntamente com as estruturas frasais equivalentes  $EF_{(I_O, I_A)}$  dados idiomas  $I_O$  e  $I_A$ , como ilustra a Figura 5.5.

É realizada a classificação de cada segmento das frases recebidas através de consulta a dicionário externo. Em seguida, é realizada busca de  $EF_{I_O, I_A}$  e como última etapa deste bloco, é realizada a comparação entre as estruturas frasais de  $T_T$  e a  $EF$  esperada, dado  $T_O$ . Por fim, é realizado método de comparação de palavras (WAGNER; FISCHER, 1974), resultando em valor  $F_{EF}$  no intervalo  $[0,1]$ .

No segundo bloco do método ABEF, é dada continuidade no processo de avaliação da tradução candidata a solução. Este bloco recebe  $T_O$ ,  $T_T$ ,  $EF_{T_T}$ ,  $EF_{I_O, I_A}$  e  $F_{EF}$  como parâmetros de entrada. De posse destes parâmetros, é realizada a busca da  $EF_{I_A}$  mais próxima de  $EF_{T_T}$ , utilizando o método de comparação de palavras. Escolhida

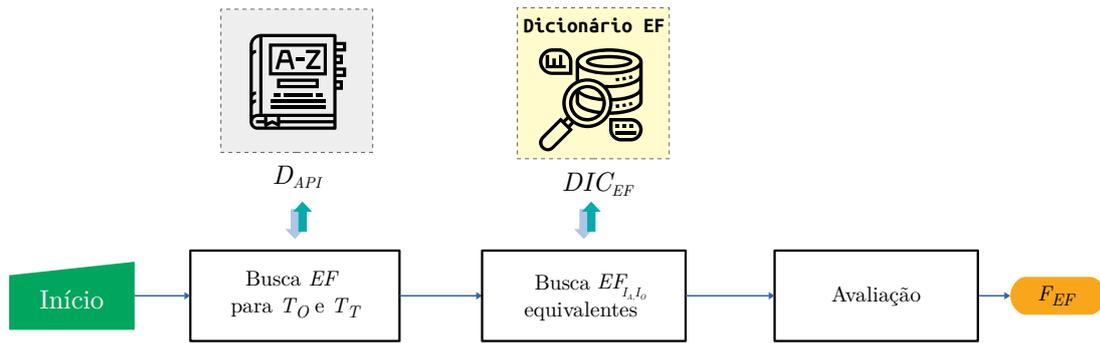


Figura 5.5 - Avaliação da estrutura frasal proposta.

a  $EF$  mais próxima a  $EF_{T_T}$ , cada termo da frase  $T_T$  é comparado com a tradução equivalente à  $T_O$ . Para cada tradução equivalente validada,  $F_n$  é incrementado em uma unidade. Ao final do processamento de todos os elementos de  $T_T$ , o valor final de  $F_{ABEF}$  é dado por:

$$F_{ABEF} = \frac{F_n}{cont_p} \cdot F_{EF} \quad (5.1)$$

na qual  $F_{EF}$  é o valor da avaliação considerando somente a estrutura frasal,  $cont_p$  é a quantidade de palavras correspondentes encontradas no dicionário e dada a estrutura frasal de referência  $EF_{T_T}$ ,  $F_n$  é o somatório das traduções corretas encontradas em  $T_T$  e  $F_{ABEF}$  é o valor final da avaliação e está no intervalo  $[0, 1]$ . A Figura 5.6 ilustra os passos até a obtenção do valor final de  $F_{ABEF}$ .

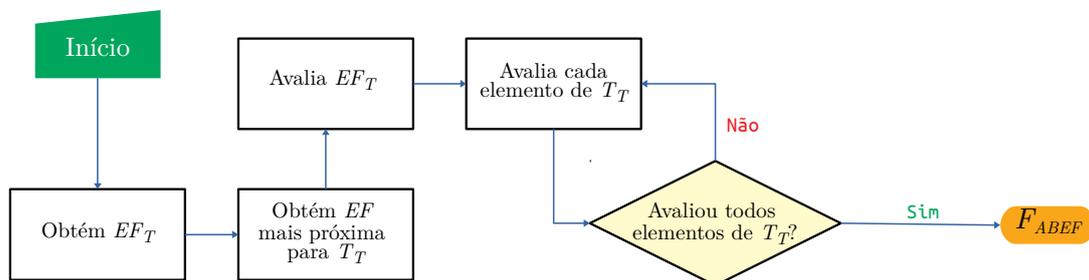


Figura 5.6 - Bloco de avaliação com base na tradução das palavras.

Para frases compostas com mais de uma sentença, é realizada avaliação individual

de cada sentença utilizando a expressão (5.1) e o resultado final da avaliação total é calculado através da média ponderada entre as avaliações resultantes e a quantidade de elementos de cada sentença, dado por:

$$F_{ABEF_T} = \frac{\sum_{i=1}^n F_{ABEF_i} \cdot EL_i}{\sum_{i=1}^n EL_i} \quad (5.2)$$

na qual  $n$  é a quantidade de sentenças avaliadas,  $EL_i$  é a quantidade de elementos/palavras da sentença avaliada,  $F_{ABEF_i}$  é a avaliação para a sentença  $i$  com  $i = 1, 2, \dots, n$  e  $F_{ABEF_T}$  é a média ponderada que representa a nota total para o texto avaliado.

A métrica utiliza o conceito de tradução bidirecional  $I_O \leftrightarrow I_A$  na obtenção de texto traduzido  $T_T$ . A avaliação é dividida em duas etapas de tradução e uma etapa de avaliação. A primeira etapa traduz  $T_O$  utilizando o processo  $I_O \rightarrow I_A$  através da  $TA_{API}$ . Esta tradução consiste na tradução de ida. Como resultado da primeira etapa é obtido  $T_T$  no idioma alvo  $I_A$ . A segunda etapa faz a tradução de volta no processo  $I_A \rightarrow I_O$  de  $T_T$ , resultando em novo texto traduzido  $T'_O$ . A terceira e última etapa utiliza banco de dados de regras das estruturas frasais  $EF_{(I_O, I_A)}$  e aplica a função de avaliação em  $T_T$  e  $T'_O$ , como ilustrado na Figura 5.7.

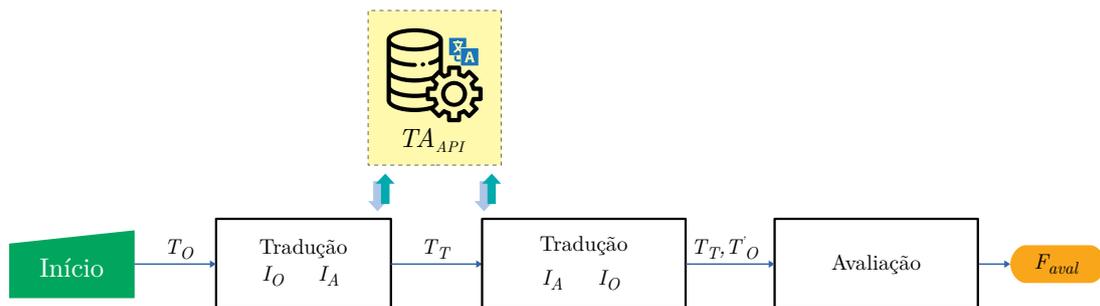


Figura 5.7 - Método de Avaliação Bilingue por Estrutura Frasal (ABEF).

O Algoritmo 1 dispõe o detalhamento dos passos necessários para obtenção da avaliação do texto traduzido  $T_T$ . A avaliação utiliza duas funções auxiliares, na qual a primeira função monta as estruturas frasais  $EF_O$  e  $EF_A$ . A segunda função calcula

o valor da avaliação da tradução de ida  $T_T$  e da tradução de volta  $T'_O$ .

---

**Algoritmo 1:** Modelo de Avaliação Bilingue por Estrutura Frasal.

---

**Dados:**  $T_O, T_T, T'_O, I_O, I_A$

**Resultado:**  $F_{aval}$

```

1 início
2    $EF_O \leftarrow \text{BuscarEstruturaFrasal}(T_O, I_O);$ 
3    $EF_A \leftarrow \text{BuscarEstruturaFrasal}(T_T, I_A);$ 
4    $F_{T_T} \leftarrow \text{CalcularFaval}(T_T, I_A, EF_A);$ 
5    $F_{T'_O} \leftarrow \text{CalcularFaval}(T'_O, I_A, EF_A);$ 
6    $F_{aval} \leftarrow \frac{F_{T_T} + F_{T'_O}}{2};$ 
7 fim

```

---

### 5.5 Tradução utilizando processo de otimização

No processo de tradução automática por otimização (TAO) é utilizada base de dados de estrutura frasal correspondente entre os idiomas  $I_O$  e  $I_A$ . O sistema utilizado é o de tradução bilingue direta **orientado por dados**. O processo é composto pela divisão textual, busca por modelo frasal em base de dados, busca no dicionário  $D_{API}$  por tradução de cada segmento da sentença  $T_O$  e pela geração da sentença traduzida  $T_T$ .

Na divisão textual é realizada segmentação da sentença  $T_O$ , categorizando quanto a classe gramatical de cada palavra. A categorização das palavras como adjetivo, advérbio, artigo, conjunção, interjeição, numeral, preposição, pronome, substantivo e verbo são necessárias para realizar a correlação entre a estrutura frasal do idioma origem  $EFI_O$  e idioma alvo  $EFI_A$ . A Figura 5.8 ilustra o exemplo de correspondência de estrutura frasal entre a sentença em português  $S_{pt}$  e a sentença em inglês  $S_{en}$ .



Figura 5.8 - Correspondência de estrutura frasal.

No exemplo da Figura 5.8, o código de referência  $C_R$  foi extraído da Tabela 2.1, na qual é apresentado que para dada sentença em português  $S_{pt}$ , há estrutura frasal definida  $EF_{pt} = I + A$  e sua equivalente em inglês é  $EF_{en} = A + I$ . Por existirem estruturas frasais distintas existem bancos de dados de estruturas frasais em português e suas correspondentes no idioma inglês. A mesma lógica é utilizada para outros pares de idiomas origem-alvo.

Com o banco de dados de estruturas frasais correspondentes definido, o passo seguinte no processo da tradução é a conexão do *software* com dicionário online para coletar os candidatos a tradução. É enviado ao dicionário  $D_{API}$  individualmente cada palavra da sentença e o idioma origem. O dicionário retorna o conjunto de informações sobre cada palavra enviada. Com as informações de classificação de cada palavra, é realizada nova chamada ao dicionário enviando a palavra no idioma origem juntamente com os parâmetros do idioma origem-alvo.

De posse dos dados, o dicionário devolve a tradução de cada palavra, na qual é realizada a primeira montagem de sentença aleatoriamente. Para definição de sequência, é utilizada função de avaliação  $F_{aval}$  modelada utilizando  $EFI_A$ . O bloco otimizador utiliza como critério de parada a função objetivo  $F_{aval}$  quando esta atinge o valor  $F_{aval} = 1$  ou quando o otimizador atinge quantidade predefinida de iterações. A Figura 5.9 ilustra o processo de tradução desde a entrada do texto origem  $T_O$ , passando pela construção do texto candidato a tradução  $T_c$  até o processamento pelo otimizador, gerando como saída final o texto traduzido  $T_T$ .

No intuito de verificar a aplicação da métrica ABEF como ferramenta auxiliar para mecanismo de tradução automática, foram utilizados o método determinístico de Nelder-Mead Simplex e o método heurístico algoritmo genético. Para ambos os métodos de otimização, será utilizada a métrica ABEF como função objetivo, analisando as frases geradas por cada método e devolvendo o valor da avaliação a cada iteração.

## 5.6 Validação da tradução

Como etapa final da metodologia é aplicada avaliação dos textos por tradutor humano. Para tanto, são submetidos à avaliação pelas métricas e por tradutores humanos e comparadas as notas para verificar qual das métricas se aproxima mais do julgamento humano. O tradutor humano faz a avaliação de cada tradução e atribui nota entre 0 e 1 para cada candidato a tradução  $T_n$ . Sendo que a nota final para cada sentença é a média aritmética das notas atribuídas pelos tradutores. É calculada a

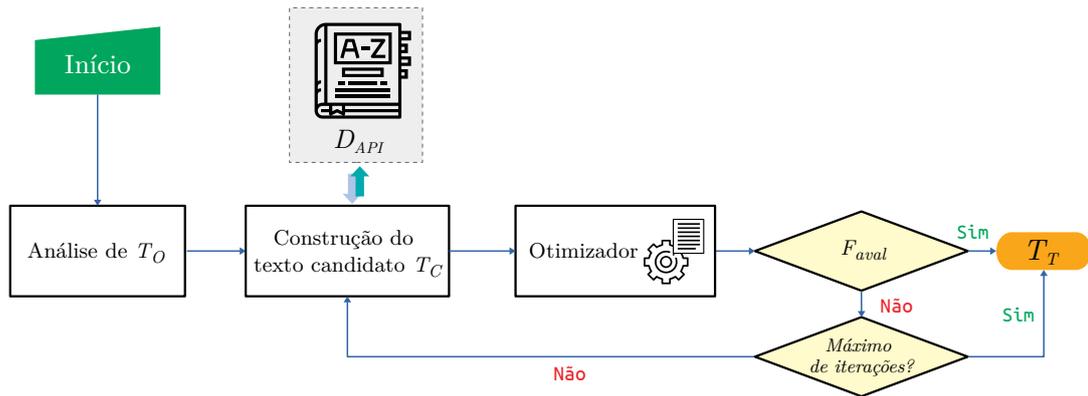


Figura 5.9 - Tradução automática utilizando processo de otimização.

distância absoluta entre os valores, distância média entre os valores e distância percentual média para cada uma das métricas, comparadas com as notas dos tradutores humanos.

## 5.7 Considerações finais

Neste capítulo foram apresentados os passos compreendidos pela metodologia proposta, além da conexão entre os diferentes API e mecanismos de busca e organização dos dados. Foram expostas as características desejadas de cada componente e o formato em que deve ser realizada a acoplagem de cada ferramenta na obtenção da metodologia proposta. Estes conceitos, são aplicados para obtenção dos resultados, apresentados no próximo capítulo.

## CAPÍTULO 6

### RESULTADOS

Neste capítulo serão apresentados os resultados obtidos utilizando a metodologia proposta. São apresentados estudos de caso com a aplicação das métricas: i) Avaliação Bilíngue por Estrutura Frasal, ii) *Bilingual Evaluation Understudy* e iii) *Metric for Evaluation of Translation with Explicit Ordering*. É realizada a aplicação da métrica proposta como função de avaliação em tradução automática gerada por processo de otimização e os resultados dos processos de tradução automática são avaliados por tradutores humanos.

#### 6.1 Avaliação da tradução

A primeira ação para avaliar a tradução automática (TA) é escolher as interface de programação de aplicativos (*Application Programming Interface* – API) de dicionário e de tradução. Como API de dicionário foi utilizado o *Oxford Dictionary API*, que é o conjunto de ferramentas construído para desenvolvedores. Este conjunto de ferramentas fornece dados monolíngues e bilíngues atualizados e com ambiente acessível e escalonável, podendo ser utilizado em áreas distintas. Como software auxiliar na TA, foi utilizado o *Translate API*. Foi desenvolvido pela *Yandex Technologies* e permite que desenvolvedores integrem a tradução automática em aplicativos. O *Translate API* suporta mais de 90 idiomas e pode traduzir palavras ou textos completos, tornando possível incorporá-lo nos mais diversos tipos de aplicativos.

##### 6.1.1 Comparação entre as métricas e validação da métrica proposta

No primeiro estudo de caso, avalia-se a tradução utilizando a métrica proposta ABEF, a métrica BLEU e a métrica METEOR. Neste estudo de caso é analisado trecho do livro *Trading in the zone: master the market with confidence, discipline and a winning attitude*, do autor Douglas (2000), publicado em inglês (texto original  $T_O$ ) e com edição impressa. A versão oficial em português (texto traduzido  $T_T$ ) foi traduzida por tradutores humanos e assim, trecho do mesmo livro em ambas as línguas foi utilizado para avaliar as métricas. O trecho selecionado para ser analisado consta na página 58 da versão em inglês e na página 81 da versão em português, na qual ambos trechos são apresentados na Figura 6.1.

Cada métrica possui característica específica e depende de conjunto de dados inicial para que possa executar a avaliação. Neste sentido, o trecho  $T_O$  foi traduzido para o português  $T_T$  utilizando os tradutores online do Google e Bing. As estruturas frasais

$T_O$	<p>Your state of mind is a by-product of your beliefs and attitudes. You can try to create consistency without having the appropriate beliefs and attitudes, but your results won't be any different than if you try to be happy when you're not having fun. When you're not having fun, it can be very difficult to change your perspective to one where you, all of a sudden, start enjoying yourself.</p>
$T_T$	<p>O estado mental é um subproduto das convicções e das atitudes. Podemos tentar ser consistentes sem ter as convicções e as atitudes adequadas, mas os resultados não serão diferentes do que tentar divertir-se quando não está feliz. Quando não está feliz, é muito difícil mudar subitamente a disposição para se divertir.</p>

Figura 6.1 - Trecho avaliado pelas métricas ABEF, BLEU e METEOR.

foram extraídas das traduções geradas pelas TA online e dos textos em inglês e português do livro. Com o intuito de garantir a igualdade da base de dados para as avaliações, as métricas receberam as versões traduzidas pelas TA e o texto traduzido da versão do livro em português para comparação e subsequente avaliação.

Considerando que a tradução do trecho do livro é da versão oficial que passou por vários avaliadores humanos e considerando que a tradução é a melhor possível, a métrica que apresentar o valor mais próximo da nota máxima pode ser considerada a métrica com maior eficácia. Na Tabela 6.1 são dispostas as avaliações obtidas para cada métrica, com base em conjunto de dados  $T_O$  e  $T_T$ . As métricas apresentam valores no intervalo  $[0, 1]$  na qual 1 é indicativo de tradução precisa/total acurácia. Os resultados das avaliações apresentam diferença de  $\approx 42\%$  entre as métricas, na qual a métrica METEOR apresentou a pior avaliação, seguida pela métrica BLEU e a métrica ABEF foi a que apresentou nota mais alta, indicando que o  $T_T$  é passível de ser a tradução do  $T_O$ .

Tabela 6.1 - Avaliações para o trecho traduzido  $T_T$ .

	ABEF	BLEU	METEOR
Avaliação	0,46785	0,37827	0,270600

### 6.1.2 Avaliação de tradução frasal e comparação entre as métricas

Para os demais estudos de casos foram utilizados três textos originais ( $T_O$ ), sendo dois em português e um em inglês, como disposto na Tabela 6.2. Utilizando o código

de referência  $C_R$  disposto na Tabela 2.1, as estruturas frasais correspondente aos  $T_{O1}$ ,  $T_{O2}$  e  $T_{O3}$  são dadas respectivamente por (6.1), (6.2) e (6.3). Nestes casos (Tabela 6.2), as estruturas frasais entre idioma origem  $I_O$  e idioma alvo  $I_A$  possuem estruturas frasais  $EF_{I_O I_A}$  correspondentes em base de dados predefinida.

Tabela 6.2 - Textos origem utilizados para avaliação das traduções.

#	Texto Origem
$T_{O1}$	O cavalo é branco.
$T_{O2}$	<i>If you don't find a way to make money while you sleep, you will work until you die.</i>
$T_{O3}$	A manga da camisa rasgou.

$$EF_{T_{O1}} = C + I + J + A \quad (6.1)$$

$$EF_{T_{O2}} = [D + H + J + J + F + I + G + I + D + H + I] \quad (6.2)$$

$$[H + J + J + G + H + J]$$

$$EF_{T_{O3}} = C + I + G + I + J \quad (6.3)$$

Os textos  $EF_{T_{O1}}$  e  $EF_{T_{O3}}$  possuem estrutura frasal com única sentença, enquanto o texto  $EF_{T_{O2}}$  possui duas sentenças. O procedimento de avaliação é realizado em cada estrutura frasal separadamente quando utilizada a métrica ABEF, sendo a avaliação final dada por (5.2).

Para a aplicação da métrica BLEU e METEOR, são necessárias frases de referência para possibilitar o cálculo da avaliação. Nos textos  $T_{O1}$  e  $T_{O3}$  foram utilizadas frases relacionadas com o mesmo tema do texto origem, trocando verbos e substantivos e variando o posicionamento dos elementos para cada referência. No texto  $T_{O2}$  foram utilizadas frases traduzidas pelos tradutores automáticos do Google, Bing e Yandex. Na aplicação da métrica ABEF, foram utilizadas estruturas frasais previamente definidas levando em consideração possíveis traduções com estruturas frasais similares às dos textos traduzidos.

Aos tradutores humanos foram fornecidos os  $T_O$  e as traduções esperadas para cada

par  $(T_O, T_T)$ . Foram consultados dez tradutores humanos e a nota final humana para cada candidato a tradução foi calculada através da média aritmética das notas dos tradutores. Todas as notas foram atribuídas no intervalo entre  $[0, 1]$ , na qual 0 representa a pior avaliação e 1 representa a avaliação máxima da sentença candidata.

### 6.1.2.1 Estudo de Caso 1

Para a avaliação do  $T_{O1}$  são utilizadas as frases apresentadas na Figura 6.2, na qual constam o texto original  $T_{O1}$ , o texto traduzido desejado  $T_{T1}$ , as sentenças de referência para as métricas BLEU e METEOR e seis candidatos a tradução  $TC_{1i}$ . Por se tratarem de frases pequenas, foram consultados tradutores humanos sobre as possibilidades de variações necessárias para testar a capacidade de avaliação das métricas. Foram então sugeridas: i) repetição de palavras, ii) troca de posição dos elementos das frases, iii) remoção de palavras e iv) mudanças nas posições dos substantivo e verbo.

$T_{O1}$  – “o cavalo é branco”  
 $T_{T1}$  – “the horse is white”  
Referências – “a white cow” “the horse is galloping” “the horse is alive” “a cow is falling”  
Candidatos –  $TC_{1a}$  → “the white horse”  
 $TC_{1b}$  → “the chicken is white”  
 $TC_{1c}$  → “the the the the”  
 $TC_{1d}$  → “this horse is black”  
 $TC_{1e}$  → “white horse is the”  
 $TC_{1f}$  → “the horse is white”

Figura 6.2 - Dados para avaliação do  $T_{O1}$ .

Na Tabela 6.3 são dispostos os valores obtidos da avaliação *Aval* e da distância absoluta *Dist* entre as avaliações obtidas pelas métricas e a avaliação realizada por humano para cada candidato  $TC_{1i}$  da tradução do  $T_{O1}$ . Como disposto na Tabela 6.4, a métrica ABEF apresentou menor distância absoluta quando comparada com a avaliação humana e na contagem das vezes que as métricas apresentaram a menor *Dist*, a métrica ABEF obteve o menor valor para cinco dos seis candidatos.

Na Tabela 6.4 a distância total  $Dist_T$  é a somatória das *Dist* de todas as frases candidatas (neste caso  $n = 6$ ) para cada métrica, com relação a avaliação dos tradutores humanos. A distância média  $Dist_M$  é a  $Dist_T$  dividida pela quantidade de frases candidatas avaliadas. A distância percentual média  $Dist_P$  é dada por:

Tabela 6.3 - Avaliações para  $T_{O1}$ .

Candidatos	ABEF		BLEU		METEOR		Humana $V_h$
	Aval	Dist	Aval	Dist	Aval	Dist	
$TC_{1a}$	0,333	0,222	1,000	0,445	0,384	0,171	0,555
$TC_{1b}$	0,675	<b>0,000</b>	0,750	0,075	0,638	0,037	0,675
$TC_{1c}$	0,062	<b>0,013</b>	0,750	0,075	0,125	0,050	0,075
$TC_{1d}$	0,500	<b>0,125</b>	0,500	<b>0,125</b>	0,468	<b>0,157</b>	0,625
$TC_{1e}$	0,500	<b>0,225</b>	1,000	<b>0,725</b>	0,500	<b>0,225</b>	0,275
$TC_{1f}$	1,000	<b>0,000</b>	1,000	<b>0,000</b>	0,992	<b>0,008</b>	1,000

Tabela 6.4 - Análise com relação a avaliação humana para  $T_{O1}$ .

	ABEF	BLEU	METEOR
$Dist_T$	0,58417	1,54500	0,64556
$Dist_M$	0,09736	0,25750	0,10759
$Dist_P$	<b>20,27%</b>	<b>36,17%</b>	24,47%

$$Dist_P = \frac{1}{n} \cdot \sum_{i=1}^n \frac{|Aval_i - V_{h_i}|}{Aval_i} \quad (6.4)$$

na qual  $V_h$  é o valor obtido pela avaliação humana. A distribuição de valores das distâncias são apresentadas na Figura 6.3.

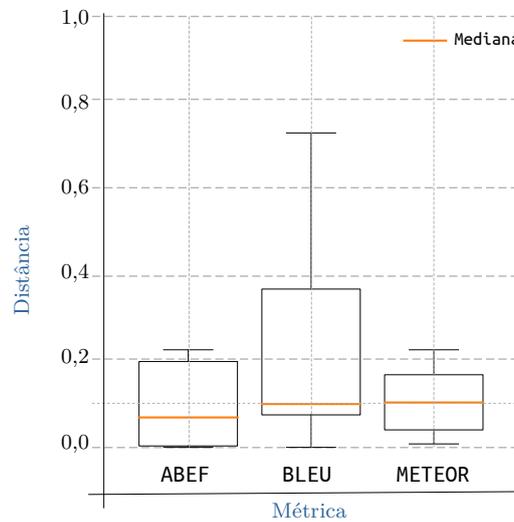


Figura 6.3 - Análise de variação das distâncias para  $T_{O1}$ .

Comparando as métricas com relação a avaliação humana, a métrica ABEF não apresentou  $Dist$  maior para nenhum dos candidatos, a métrica METEOR apresentou a maior  $Dist$  para dois candidatos e a métrica BLEU apresentou maior  $Dist$  para quatro dos seis candidatos. A métrica ABEF se apresentou mais eficaz que

as métricas BLEU e METEOR em todas as análises realizadas em  $T_{O1}$ , ficando mais próxima da tradução humana, com  $Dist_P \approx 4\%$  menor que a segunda melhor métrica.

### 6.1.2.2 Estudo de Caso 2

Para a avaliação do  $T_{O2}$  são utilizados os dados apresentados na Figura 6.4. Neste estudo específico, foi necessário utilizar o  $T_{O2}$  com múltiplas sentenças separadas por vírgula. Por ser sentença maior, foram fornecidas como base de referência para o método BLEU e método METEOR a sentença traduzida utilizando as TA: i) *Google Translator*, ii) *Bing Translator* e iii) *Yandex Translator*.

$T_{O2}$	–	<i>“if you don't find a way to make money while you sleep, you will work until you die”</i>
$T_{T2}$	–	<i>“se você não encontrar um jeito de ganhar dinheiro enquanto dorme, você vai trabalhar até morrer”</i>
Referências	–	<i>“se você não encontra uma maneira de ganhar dinheiro enquanto dorme, você vai trabalhar até morrer”</i> – Google Translator  <i>“se você não encontrar uma maneira de ganhar dinheiro enquanto você dorme, você vai trabalhar até que você morrer”</i> – Bing Translator  <i>“se não encontrares uma maneira de ganhar dinheiro enquanto dormes, trabalharás até morreres”</i> – Yandex Translator
Candidatos	–	$TC_{2a}$ → <i>“se não achar um jeito de fazer dinheiro enquanto dorme, trabalhará até morrer”</i> $TC_{2b}$ → <i>“você irá trabalhar até morrer, se não achar um jeito de ganhar dinheiro enquanto dorme”</i> $TC_{2c}$ → <i>“se tu não encontrares uma forma de ganhar dinheiro enquanto dormes, trabalharás até morrer”</i> $TC_{2d}$ → <i>“se você não encontrar um caminho para fazer dinheiro enquanto você dorme, você trabalhará até você morrer”</i> $TC_{2e}$ → <i>“se não encontrar você meio para criar dinheiro quando dorme você, trabalhar até morrer você irá”</i> $TC_{2f}$ → <i>“se você não encontrar um jeito de ganhar dinheiro enquanto dorme, você vai trabalhar até morrer”</i>

Figura 6.4 - Dados para avaliação do  $T_{O2}$ .

Na Tabela 6.5 são dispostos os valores obtidos das avaliações *Aval* e das distâncias absolutas  $Dist$  para todos os candidatos  $TC_{2i}$  avaliados no  $T_{O2}$  e a distribuição das distâncias são apresentadas na Figura 6.5. Na análise da Tabela 6.5, a métrica ABEF apresentou menor  $Dist$  em quatro dos seis candidatos para tradução, a métrica BLEU apresentou melhor  $Dist$  para dois dos seis candidatos. A métrica ABEF apresentou  $Dist$  maior para o candidato  $TC_{2b}$  e a métrica METEOR apresentou  $Dist$  maior para todos os demais candidatos da tradução. Na Tabela 6.6 são dispostas as

distância total  $Dist_T$ , distância média  $Dist_M$  e  $Dist_P$ , dada por (6.4).

Tabela 6.5 - Avaliações para  $T_{O_2}$ .

Candidatos	ABEF		BLEU		METEOR		Humana $V_h$
	Aval	Dist	Aval	Dist	Aval	Dist	
$TC_{2a}$	0,933	<b>0,002</b>	0,615	0,320	0,179	<b>0,756</b>	0,935
$TC_{2b}$	0,000	<b>0,645</b>	0,561	<b>0,084</b>	0,178	0,467	0,645
$TC_{2c}$	0,939	0,134	0,857	<b>0,052</b>	0,278	<b>0,527</b>	0,805
$TC_{2d}$	0,944	<b>0,084</b>	0,705	0,125	0,251	<b>0,609</b>	0,860
$TC_{2e}$	0,640	<b>0,205</b>	0,562	0,725	0,135	<b>0,710</b>	0,845
$TC_{2f}$	0,939	<b>0,061</b>	0,875	0,000	0,282	<b>0,718</b>	1,000

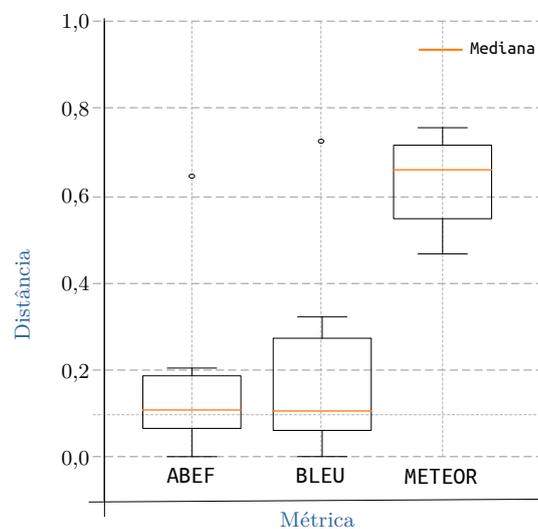


Figura 6.5 - Análise de variação das distâncias para  $T_{O_2}$ .

Tabela 6.6 - Análise com relação a avaliação humana para  $T_{O_2}$ .

	ABEF	BLEU	METEOR
$Dist_T$	1,03030	0,91708	3,68385
$Dist_M$	0,17172	0,15285	0,61398
$Dist_P$	23,91%	18,03%	73,80%

Avaliando as métricas e levando em consideração a avaliação humana das traduções, a métrica METEOR apresentou maior  $Dist$  para cinco dos seis candidatos. A métrica ABEF apresentou maior  $Dist$  para um dos seis candidatos enquanto a métrica BLEU não apresentou maior  $Dist$  para nenhum dos candidatos. A métrica ABEF apresenta-se mais precisa que as métrica BLEU e métrica METEOR, quando considerada a quantidade de vezes que apresentou a menor  $Dist$ . No entanto, apresentou

$Dist_P > 5\%$  da métrica BLEU que obteve somente dois resultados menores de  $Dist$ . Esta diferença se deve ao fato de que as referências fornecidas aos métodos BLEU e METEOR possuem quantidade alta de  $n$ -gramas (palavras) coincidentes com a tradução esperada.

### 6.1.2.3 Estudo de Caso 3

Para a avaliação de  $T_{O3}$  foram utilizados os dados apresentados na Figura 6.6. Neste estudo de caso foram escolhidas palavras de múltiplos significados, que dependendo da aplicação pode alterar totalmente o sentido final da frase. Foram consultados tradutores humanos para que pudessem ser utilizadas referências coerentes aos métodos BLEU e METEOR, assim como candidatos a tradução que pudessem testar a acurácia das métricas aplicadas.

Na Tabela 6.7 são dispostas as avaliações obtidas pelas três métricas e pelo tradutor humano. Para o  $T_{O3}$ , a métrica ABEF obteve menor  $Dist$  seguida pela métrica METEOR e posteriormente pela métrica BLEU. Na contagem das vezes que as métricas apresentaram menor  $Dist$  considerando a avaliação humana, a métrica ABEF obteve a menor  $Dist$  para quatro dos seis candidatos, a métrica BLEU para dois candidatos, enquanto a métrica METEOR apresentou menor distância para um candidato, como apresentado na Figura 6.7.

$T_{O3}$  – “a manga da camisa rasgou”  
 $T_{O3}$  – “the shirt sleeve tore”  
Referências – “the shirt” “the shirt is tore” “the shirt is black” “a shirt can tear” “shirt sleeve”  
Candidatos –  $TC_{3a}$  → “the tore sleeve”  
 $TC_{3b}$  → “the collar has tear”  
 $TC_{3c}$  → “the short sleeve”  
 $TC_{3d}$  → “the shirt is torn”  
 $TC_{3e}$  → “shirt sleeve the rent”  
 $TC_{3f}$  → “the shirt sleeve tore”

Figura 6.6 - Dados para avaliação do  $T_{O3}$ .

Com o intuito de verificar a diferença das notas obtidas por cada métrica comparada à avaliação dos tradutores humanos, foram calculados e dispostos na Tabela 6.8 a  $Dist_T$ ,  $Dist_M$  e  $Dist_P$ .

Avaliando as três métricas e levando em consideração a avaliação da tradução hu-

Tabela 6.7 - Avaliações para  $T_{O3}$ .

Candidatos	ABEF		BLEU		METEOR		Humana $V_h$
	Aval	Dist	Aval	Dist	Aval	Dist	
$TC_{3a}$	0,333	<b>0,017</b>	1,000	<b>0,650</b>	0,256	0,094	0,350
$TC_{3b}$	0,136	<b>0,319</b>	0,500	<b>0,045</b>	0,227	0,228	0,455
$TC_{3c}$	0,675	<b>0,585</b>	0,500	0,410	0,227	<b>0,137</b>	0,090
$TC_{3d}$	0,150	<b>0,055</b>	0,500	<b>0,295</b>	0,468	0,263	0,205
$TC_{3e}$	0,500	<b>0,410</b>	0,750	0,660	0,852	<b>0,762</b>	0,090
$TC_{3f}$	1,000	<b>0,000</b>	1,000	<b>0,000</b>	0,852	<b>0,148</b>	1,000

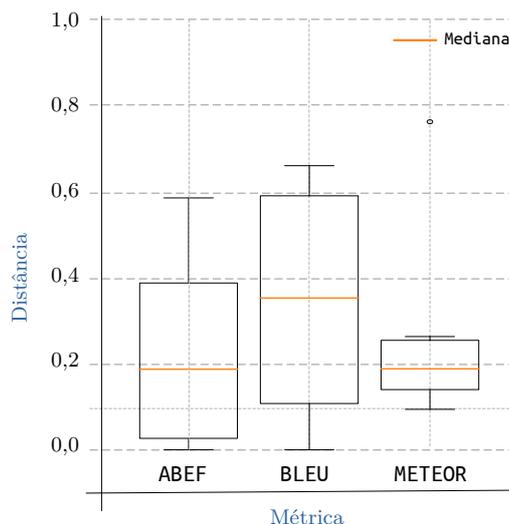


Figura 6.7 - Análise de variação das distâncias para  $T_{O3}$ .

Tabela 6.8 - Análise com relação a avaliação humana para  $T_{O3}$ .

	ABEF	BLEU	METEOR
$Dist_T$	1,38498	2,06000	1,063234
$Dist_M$	0,23083	0,34333	0,27206
$Dist_P$	<b>45,04%</b>	<b>50,50%</b>	49,61%

mana, as métricas apresentaram  $Dist_P$  similares. Cada métrica apresentou maior  $Dist$  para dois dos seis candidatos. A métrica ABEF se apresentou mais precisa do que a métrica BLEU e a métrica METEOR, quando levada em consideração a quantidade de vezes que apresentou a menor distância, superando a métrica BLEU em 5,46%.

#### 6.1.2.4 Estudo de Caso 4

No quarto estudo de caso foram submetidos frases e parágrafos dos livros *Think and grow rich* do autor Hill (2011) e *The richest man in Babylon* do autor Classon (2020). Com o objetivo de validar a capacidade das métricas, foram utilizados como

textos origem tanto a versão em inglês quanto a versão em português, dado que os livros em ambos idiomas são considerados como tradução oficial. Os textos avaliados são apresentados na Figura 6.8 e Figura 6.9.

Inglês	Português
<i>Ahead of you stretches your future like a road leading into the distance.</i>	<i>À nossa frente se estende o futuro como uma estrada que se perde na distância.</i>
<i>Along that road are ambitions you wish to accomplish. Desires you wish to gratify.</i>	<i>Ao longo dela há ambições que queremos realizar. Desejos que queremos satisfazer.</i>
<i>Like the law of gravity, they are universal and unchanging May they prove for you, as they have proven to so many others, a sure key to a fat purse, larger bank balances and gratifying financial progress.</i>	<i>Como a lei da gravidade esses princípios são universais e imutáveis. Espero que lhe propiciem como o fizeram para muitos outros os meios eficazes para uma carteira cheia uma bela conta bancária e um satisfatório progresso financeiro.</i>
<i>Our prosperity as a nation depends upon the personal financial prosperity of each of us as individuals.</i>	<i>Nossa prosperidade como nação depende de nossa prosperidade financeira como indivíduos.</i>
<i>To the business executives who have distributed these tales in such generous quantities to friends, relatives, employees and associates, the author takes this opportunity to express his gratitude. No endorsement could be higher than that of practical men who appreciate its teachings because they, themselves, have worked up to important successes by applying the very principles it advocates.</i>	<i>A todos os homens de negócios que distribuíram estas narrativas em tão generosas quantidades a amigos parentes empregados e associados o autor aproveita a oportunidade para expressar sua gratidão. nenhum respaldo pode ser mais valioso do que o desses homens práticos que apreciaram seus ensinamentos, porque eles, por si mesmos, conseguiram o que o presente livro defende.</i>
<i>Money is governed today by the same laws which controlled it when prosperous men thronged the streets of Babylon, six thousand years ago.</i>	<i>O dinheiro é hoje governado pelas mesmas leis que o controlavam quando, há seis mil anos, homes prósperos enchiam as ruas da Babilônia.</i>
<i>Consult with wise men. Secure the advice of those experienced in the profitable handling of gold. Let their wisdom protect thy treasure from unsafe investments.</i>	<i>Sigam a opinião daqueles que lidam habitualmente com dinheiro. Deixem que o tirocínio deles proteja seus tesouros contra os investimentos de alto risco.</i>
<i>If a man be lucky, there is no foretelling the possible extent of his good fortune. Pitch him into the Euphrates and like as not he will swim out with a pearl in his hand.</i>	<i>Se um homem tem sorte, não há como prever extensão possível de sua boa fortuna. Joguem-no no Eufrates, e ele se salvar a nado com uma pérola na mão.</i>

Figura 6.8 - Trechos extraídos do livro *The richest man in Babylon*.

Para a aplicação das avaliações foram geradas as sentenças correspondentes para a métrica ABEF e foram gerados, através do *Google Translator* e *Bing Translator*, pares de traduções Inglês-Português e Português-Inglês para cada sentença avaliada. Os textos foram submetidos às métricas ABEF, BLEU e METEOR. Os conjuntos de dados de avaliação obtidos foram analisados quanto à distância da avaliação oficial, considerada com valor 1 (um) e levando em consideração a quantidade de palavras em cada trecho avaliado, como apresentado na Figura 6.10.

Cada ponto de dispersão na Figura 6.10 corresponde à projeção da relação entre a quantidade de palavras e a avaliação atribuída por cada método. A reta e a região ao redor desta representam o comportamento de cada métrica em atribuir maior ou

Inglês	Português
<p><i>To the business executives who have distributed these tales in such generous quantities to friends, relatives, employees and associates, the author takes this opportunity to express his gratitude. No endorsement could be higher than that of practical men who appreciate its teachings because they, themselves, have worked up to important successes by applying the very principles it advocates.</i></p> <p><i>Psychologists have correctly said that when one is truly ready for a thing, it puts in its appearance.</i></p> <p><i>If you are ready to put it to use, you will recognize this secret at least once in every chapter. I wish I might feel privileged to tell you how you will know if you are ready, but that would deprive you of much of the benefit you will receive when you make the discovery in your own way.</i></p> <p><i>What is education, anyway? This has been answered in full detail.</i></p> <p><i>Previously you may have had a logical excuse for not having forced life to come through with whatever you asked, but that alibi is now obsolete, because you are in possession of the master key that unlocks the door to life's bountiful riches.</i></p> <p><i>It is the satisfaction that comes to all who conquer self and force life to pay whatever is asked</i></p>	<p><i>É absolutamente verdade que pensamentos são coisas, aliás, coisas muito poderosas, quando conjugadas com um objetivo definido, perseverança e um desejo ardente de que se transformem em riqueza ou outros bens materiais.</i></p> <p><i>Psicólogos acertaram ao dizer que quando alguém está realmente pronto para alguma coisa, essa coisa não deixa de aparecer.</i></p> <p><i>Se você estiver pronto para pô-lo em uso, reconhecerá o segredo pelo menos uma vez em cada capítulo. Gostaria de ter o privilégio de dizer-lhe como o saberá quando estiver pronto a recebê-lo, mas isso o privaria de muitos dos benefícios que receberá ao fazer a descoberta a seu modo.</i></p> <p><i>O que é educação afinal? Isso foi respondido pormenorizadamente.</i></p> <p><i>Antigamente você poderia ter uma desculpa lógica por não ter forçado a vida a lhe dar o que queria, mas esse hábito agora obsoleto, porque você está de posse da chave-mestra que abre a porta da vida a riquezas fartas.</i></p> <p><i>É a satisfação o que atinge aos que conquistam a si mesmos e forçam a vida a pagar a que lhe pedem.</i></p>

Figura 6.9 - Trechos extraídos do livro *Think and grow rich*.

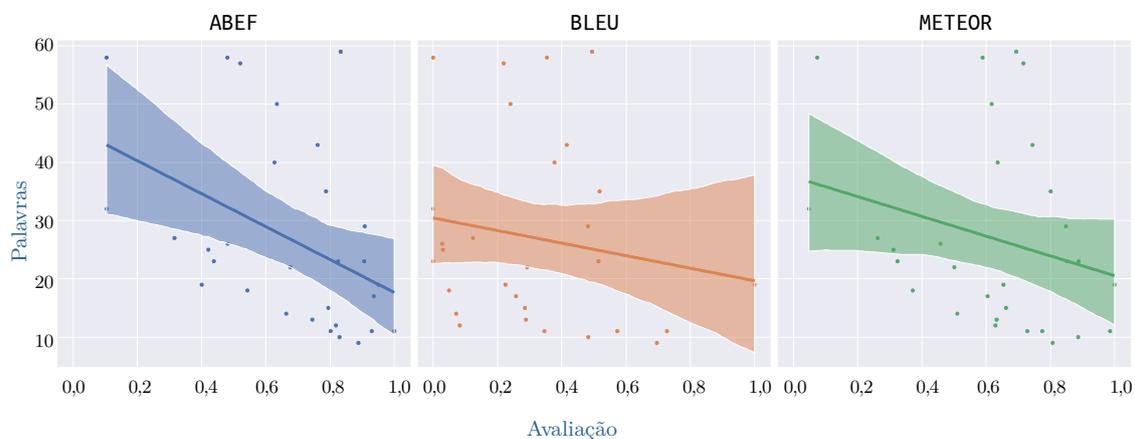


Figura 6.10 - Gráficos de dispersão para os trechos analisados.

menor avaliação de acordo com a quantidade de palavras. Apesar da inclinação da reta indicar tendência das métricas a atribuírem maior nota para textos menores, não é observada a permanência das dispersões diretamente sobre a linha.

Na Figura 6.11 são sobrepostos os gráficos de dispersão dos valores de avaliação de

cada uma das métricas. É possível verificar que as métricas apresentam comportamento similar quanto à quantidade de palavras nos trechos avaliados e o comportamento não possui relação linear direta com a quantidade de palavras.

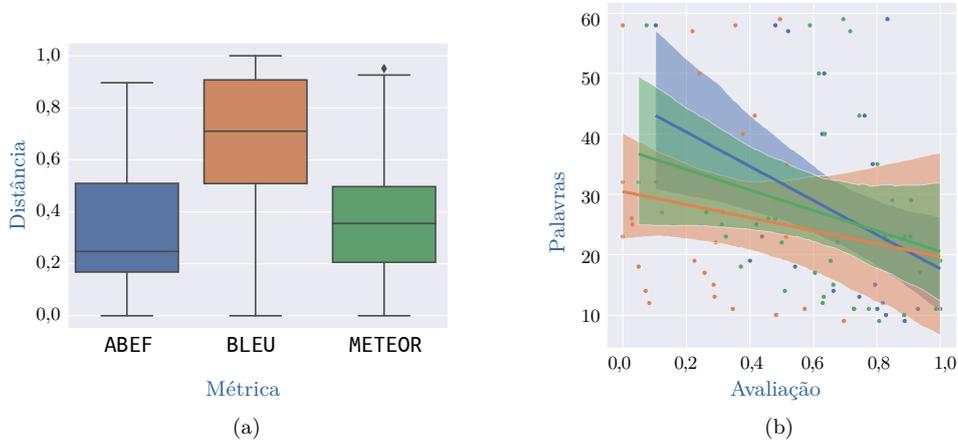


Figura 6.11 - Métricas e avaliações: (a) distâncias entre as avaliações e (b) sobreposição das dispersões.

Observa-se na Figura 6.11 que a concentração das distâncias das avaliações da métrica BLEU estão na região acima de 50% na qual a distância da avaliação ideal é 1 (um). As métricas ABEF e METEOR possuem valores de distâncias parecidas, no entanto a métrica ABEF apresentou mediana mais próxima à menor distância da avaliação ideal.

## 6.2 Tradução automática gerada por processo de otimização

Com o objetivo de abordar novo método de TA, foi proposto e analisado o processo de otimização aplicado na tradução de texto. Para este estudo de caso foram utilizados método determinístico de Nelder-Mead e método heurístico algoritmo genético. A ideia é, o processo de otimização recebe o  $T_O$  e realiza busca no dicionário para tradução palavra por palavra, construindo sentenças aleatoriamente. A sentença construída é avaliada pela métrica ABEF, e se o candidato (a sentença) obter  $Aval = 1$ , o otimizador cessa seu processo e a sentença é considerada traduzida. Caso o candidato obtenha  $0 \leq Aval < 1$ , o método de otimização gera novas sentenças candidatas e realiza novamente a avaliação utilizando a métrica ABEF.

Para realizar o processo de otimização, foi necessário construir todos os algoritmos, pois, as estruturas das ferramentas de otimização encontradas disponíveis não eram

compatíveis com os formatos de dados de TA. O primeiro algoritmo de otimização a ser testado foi o Nelder-Mead, no entanto, este método de otimização não conseguiu otimizar/traduzir os  $T_O$ . O algoritmo ficou estagnado em ótimos locais e não tinha energia para prosseguir, produzindo sentenças sem significados, com termos totalmente embaralhados e sem sentido, cessando o processo de otimização.

O algoritmo genético (AG), foi implementado com as características: i) operador de cruzamento uniforme, ii) operador de mutação adaptativo, iii) manipulador de seleção por torneio, iv) número de indivíduos na população inicial de dez, v) número máximo de gerações  $G_{max} = 500$ , vi) taxa de cruzamento  $20\% \leq P_c \leq 90\%$ , vii) taxa de mutação  $5\% \leq P_m \leq 40\%$ , viii) tamanho do torneio  $\tau = 4$  e ix) critério de parada por  $G_{max}$  ou pela função de avaliação  $F(x) = 0$ , dada por:

$$F(x) = 1 - F_{ABEF} \quad (6.5)$$

na qual  $F_{ABEF}$  é dada por (5.1), sendo esta a avaliação atribuída para cada candidato avaliado e  $F(x)$  é o valor da função de avaliação do processo de otimização, no qual  $x$  é a matriz  $m \times n$  com  $m = 10$  indivíduo. Pela própria natureza da métrica ABEF, os resultados de  $F(x)$  ficam no intervalo  $[0 \ 1]$ . Quando  $F(x) \rightarrow 0$ , melhor é considerado o resultado. Neste estudo de caso foram traduzidas pela TA proposta os  $T_{O1}$ ,  $T_{O2f}$  e  $T_{O3}$ . Como a proposta é de TA, o método de otimização deve receber além dos parâmetros tradicionais do AG, os dicionários de correspondência de estrutura frasal nos idiomas português e inglês (para este caso), o dicionário com as respectivas classificações gramaticais com sinônimos e o  $T_O$ .

As Figura 6.12, Figura 6.13 e Figura 6.14 apresentam o mapeamento dos valores de  $F(x)$  pelas gerações para os  $T_{O1}$ ,  $T_{O2f}$  e  $T_{O3}$ , respectivamente. Para a tradução de  $T_{O1}$  a otimização utilizou 34 gerações, para  $T_{O2f}$  96 gerações e para  $T_{O3}$  39 gerações. O  $T_{O2f}$  possui estrutura frasal mais elaborada que os  $T_{O1}$  e  $T_{O3}$ , necessitando assim de mais iterações para minimizar  $F(x)$  até a tradução desejada.

Com os resultados obtido nos três processos de tradução otimizados, foi possível observar a capacidade de tradução da TA por otimização (TAO) proposta. A métrica ABEF assim como as demais métricas apresentadas podem servir de função de avaliação para TAO e a obtenção das frases em igual teor e forma como as classificadas previamente por tradutores humanos, validam a capacidade do tradutor automático gerar traduções válidas baseado em dicionário de dados.

$T_{O1}$  o cavalo é branco

$c_1$  the the horse horse  $F(x) = 0,8$

⋮

$c_{34}$  the horse is white  $F(x) = 0,0$

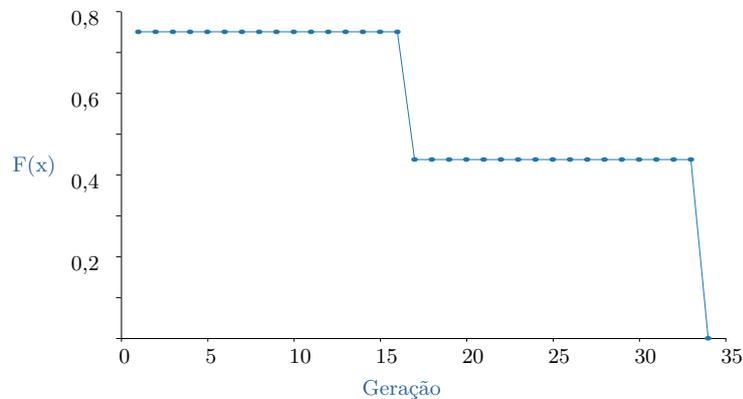


Figura 6.12 - Processo de otimização para  $T_{O1}$ .

$T_{O2f}$  você vai trabalhar até morrer

$c_1$  will you will will you will  $F(x) = 0,77778$

⋮

$c_{96}$  you will work until you die  $F(x) = 0,0$

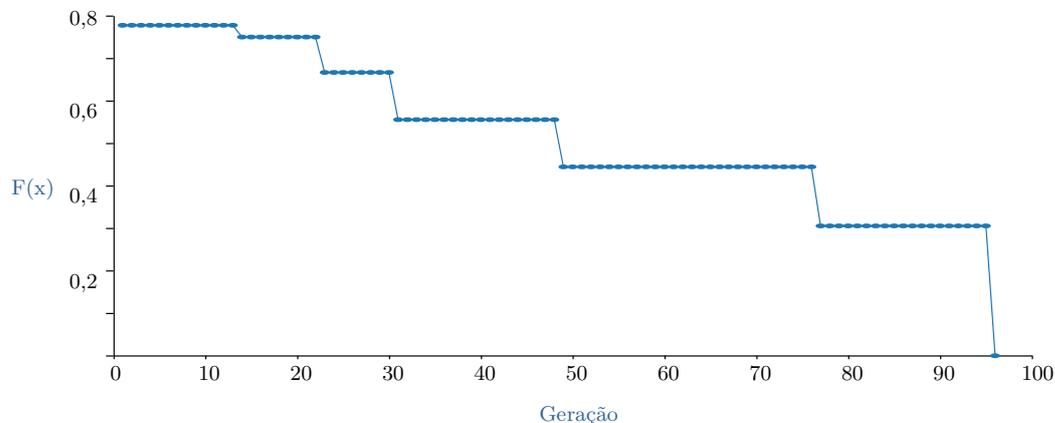


Figura 6.13 - Processo de otimização para  $T_{O2f}$ .

### 6.3 Discussão

Os resultados obtidos nos estudos de caso apresentaram a capacidade da métrica ABEF como mecanismo de avaliação de TA. As notas obtidas pela métrica proposta superou, na grande maioria, a acurácia das outras métricas empregadas e consolidadas na literatura. A métrica ABEF pode ser utilizada como mecanismo auxiliar

$T_{O3}$	a manga da camisa rasgou	
$C_1$	ripped the shirt shirt	$F(x) = 0,625$
	:	
$C_{39}$	the shirt sleeve tore	$F(x) = 0,0$

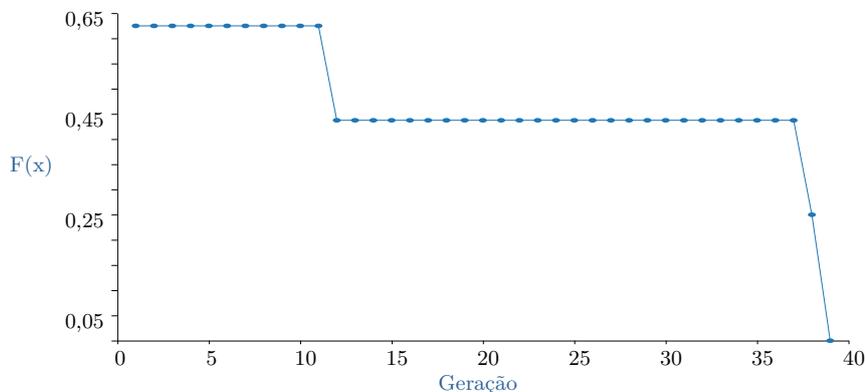


Figura 6.14 - Processo de otimização para  $T_{O3}$ .

de TA, pois, apresentou precisão de  $\approx 5\%$  superior as demais métricas.

Para a aplicação da métrica proposta foram encontradas algumas dificuldades quanto à criação das estruturas frasais equivalentes. No entanto, uma vez definidas estas estruturas, basta armazená-las no conjunto simples de dados para que possa ser utilizada em novas avaliações de tradução. A utilização somente de frases simples e com subsequentes estruturas frasais correspondentemente simples, atrapalhou no processo de definição da métrica. A utilização de estruturas frasais mais elaboradas possibilitou a depuração e aperfeiçoamento da métrica para aplicação em textos de extensão variada, haja vista que qualquer texto pode ser dividido em partes.

Foram realizados três diferentes testes para validação da métrica. Em cada frase de cada teste, foi abordado cenário diferente, sendo: i) no primeiro foi utilizado tradução oficial de livro como fonte de comparação absoluta da tradução, ii) no segundo ( $T_{O1}$ ,  $T_{O2}$  e  $T_{O3}$ ) a fonte de comparação absoluta foram os textos traduzidos e validados por dez tradutores humanos distintos e iii) no terceiro teste, na TAO, foi utilizado como fonte de comparação absoluta as frases validadas dos Estudo de Caso 1, Estudo de Caso 2 e Estudo de Caso 3. Analisando o comportamento da métrica, observa-se que ela não é precisa como a avaliação do tradutor humano, assim como as demais métricas. No entanto, apresenta capacidade de mensurar a TA. A métrica proposta apresenta limitações quando os textos são extensos, necessitando dividi-los em partes menores, limitando nas frases o número de verbos.

Levando em consideração a métrica BLEU que foi introduzida no ano de 2002 e a métrica METEOR apresentada no ano de 2005, há a necessidade do estudo e criação de novas métrica levando em consideração aspectos não considerados pelas métricas predecessoras. Para obter melhores traduções para os mais diversos fins, é importante construir/desenvolver métricas mais próximas à avaliação de tradutores humanos, sendo que quanto mais eficientes, mais rápido o processo de tradução por otimização irá convergir para o resultado desejado.

Os avanços nos estudos de tradução automática alcançados nos últimos anos possibilitaram a criação de ferramentas complexas e alto custo atrelado, como o *Google Translator* que utiliza a rede *Bidirectional Encoder Representations from Transformers* (BERT). O algoritmo BERT foi inicialmente idealizado para a utilização na avaliação de textos nos mecanismos de busca. A aplicação da rede BERT é utilizada no processo de treinamento das redes neurais do *Google Translator*. O *Google Translator* foi utilizado como ferramenta de tradução das frases também traduzidas pelo TAO. No entanto, o TAO não conseguiu obter resultados de traduções consistentes para textos mais elaborados. Neste sentido, abre-se a oportunidade de estudos para possível aplicação da rede BERT como forma de hibridizar a métrica ABEF e consequentemente o TAO para obtenção de avaliações e traduções mais precisas.

No início do trabalho ocorreram erros/equívocos quanto à escolha das tecnologias/ferramentas a serem utilizadas. Por exemplo, como o trabalho tem o intuito de somente apresentar a métrica e a TA proposta, tudo seria desenvolvido em Matlab. No entanto, esta ferramenta não fornece os recursos necessários. Em seguida, após iniciar os estudos, foi adotada a linguagem Python por contar com biblioteca com as métricas BLEU e METEOR, além de oferecer aparato completo para integração do sistema desenvolvido com as API externas.

No desenvolvimento do trabalho, foram identificadas possibilidades que ainda podem ser exploradas, como por exemplo, a utilização de inteligência artificial (IA) aplicada em conjunto com o método de otimização nas TAO e a criação de banco de dados para frases de referências, entre várias outras. Certamente a incorporação de IA melhorará a acurácia do método de TAO proposto. Por existirem estudos avançados utilizando redes neurais e IA, não seria plausível o investimento de estudos de aplicação da métrica ABEF na avaliação de textos com expressões idiomáticas, como expressões regionais e gírias.

## CAPÍTULO 7

### CONCLUSÃO

Este trabalho apresentou a criação de métrica para avaliação de tradução de texto e a aplicação desta métrica na construção de tradutor automático baseado em processo de otimização. O método proposto de avaliação por estrutura frasal possibilitou a criação de nova alternativa para avaliação de tradução automática. A hipótese levantada inicialmente foi comprovada através dos resultados obtidos, apresentando a capacidade de utilizar a estrutura frasal de textos como base para ferramenta de avaliação de TA, contribuindo na criação de nova metodologia.

A utilização de API possibilitou a extração dos dados de cada palavra. A extração dos dados gramaticais possibilitou a formação do dicionário de estruturas frasais e o tratamento das informações dos textos originais e candidatos a tradução. A geração das informações processadas e do dicionário de estruturas frasais possibilitou a criação da métrica baseada na correspondência frasal entre textos de idiomas distintos, fornecendo ferramenta de avaliação de candidatos a tradução. As avaliações obtidas pela métrica proposta puderam ser comparadas às métricas existentes e à avaliação de tradutores humanos.

O método proposto foi utilizado na avaliação de um parágrafo de texto extraído de livro, obtendo-se avaliação melhor que as outras métricas utilizadas. Na aplicação da métrica ABEF em frases isoladas, foram obtidas avaliações mais precisas do que as métricas BLEU e METEOR. O estudo de criação de tradução automática para o processo de otimização possibilitou a utilização do método proposto como ferramenta auxiliar de obtenção de traduções.

A vantagem do método proposto é a quantidade de dados envolvida na criação do dicionário de estruturas frasais correspondentes. Pela estrutura proposta, o método pode ser aplicado em outros idiomas, desde que se tenha as estruturas frasais correspondentes e dicionário entre os idiomas. A métrica proposta pode ser utilizada em vários segmentos como: i) avaliação de fluência escrita em programas de imigração na qual é exigida proficiência em determinados idiomas, ii) avaliação eletrônica em processos seletivos, iii) desenvolvimento de ferramentas de tradução melhoradas e iv) outras. Conclui-se que a métrica proposta apresentou avaliação próxima da avaliação de tradutores humanos, além de prover parâmetros para construção de mecanismo de tradução automática baseado em processo de otimização.

## 7.1 Contribuições do trabalho

### **Metodologia e técnica de tradução automática:**

O trabalho gerou nova metodologia de avaliação de tradução bilíngue que pode ser utilizada em múltiplos pares de idioma. A metodologia proposta pode servir como ferramenta para aprimoramento de treinamento de redes neurais. Além da metodologia, foi gerada nova técnica de tradução automática através de processo de otimização

### **Artigo submetido em revista:**

SILVA, ALAN H. F. ; MAGALHAES, ALANA S. ; WAINER, GABRIEL A. ; CALIXTO, WESLEY P. . BEPHS: bilingual evaluation by phrasal-structure for machine translation. Applied Soft Computing. The Official Journal of the World Federation on Soft Computing (WFSC), Elsevier.

### **Artigos publicados em revista:**

SILVA, ALAN H. F. ; MAGALHAES, ALANA S. ; BULHOES, JUNIO S. ; WAINER, GABRIEL A. ; FURRIEL, GEVANNE P. ; CALIXTO, WESLEY P. . Parametric Regression Applied for Determination of Electrical Parameters of Synchronous and Induction Generators Operating in Parallel on the Electrical Energy Repowering System. Energies, v. 14, p. 3875, 2021.

MAGALHAES, ALANA S. ; BULHOES, JUNIO S. ; REIS, MARCIO R. C. ; GOMES, VIVIANE M. ; SILVA, ALAN H. F. ; ALVES, AYLTON J. ; WAINER, GABRIEL A. ; GANZAROLI, CLEBER A. ; ARAUJO, WANDERSON R. H. ; CALIXTO, WESLEY P. . Experimental study of induction generator as a repowering solution. International Transactions on Electrical Energy Systems, v. 1, p. 1-25, 2020.

SILVA, UYARA FERREIRA ; LEMOS, SANTIAGO ; DE QUEIROZ JÚNIOR, MANOEL IVANY ; CALIXTO, WESLEY PACHECO ; SEGATTI, NIKHOLAS ; FURRIEL, GEOVANNE PEREIRA ; SILVA, ALAN HENRIQUE FERREIRA . Estimate of Geopressures using Conformal Mapping in Eccentric Wellbores. Transactions on Environment and Electrical Engineering, v. 2, p. 1, 2017.

MAGALHÃES, ALANA DA SILVA ; MORAES, PEDRO HENRIQUE FRANCO ; SILVA, ALAN HENRIQUE FERREIRA ; GOMES, PEDRO HENRIQUE GARCIA

; ALVES, AYLTON JOSÉ ; CALIXTO, WESLEY PACHECO . Reconditioning in synchronous operation with one parallel induction generator. Transactions on Environment and Electrical Engineering, v. 1, p. 66, 2016.

**Artigos publicados em congresso:**

CARARO, JOSE A. G. ; SILVA, ALAN H. F. ; NETO, JOAO CAETANO ; REIS, MARCIO R. C. ; CALIXTO, WESLEY P. . Methodology for voltage adequacy using photovoltaic distributed generation. In: 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), 2019, Genova. 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), 2019.

SILVA, A. H. F.; MAGALHAES, A. S. ; SILVA, U. F. ; ALVES, A. J. ; CALIXTO, W. P. . Conformal Mapping Applied to Encoding and Decoding of Images. In: IEEE CHILECON 2017, 2017, Pucón. IEEE CHILECON 2017, 2017.

MAGALHAES, ALANA S. ; BULHOES, JUNIO S. ; FURRIEL, GEOVANNE P. ; REIS, MARCIO R. C. ; ALVES, AYLTON J. ; SILVA, ALAN H. F. ; CALIXTO, WESLEY P. . Parametric regression in synchronous and induction generators. In: 2017 18th International Scientific Conference on Electric Power Engineering (EPE), 2017, Kouty nad Desnou. 2017 18th International Scientific Conference on Electric Power Engineering (EPE), 2017.

ALELUIA, LEOVIR CARDOSO ; DA SILVA MAGALHAES, ALANA ; SILVA, ALAN HENRIQUE FERREIRA ; CALIXTO, WESLEY P. ; ALVES, AYLTON JOSE ; MATIAS, CALEBE ABREHOSA ; DOMINGUES, ELDER GERALDO ; DOMINGOS, JOSE LUIZ . Attenuation methodology for torque oscillations in synchronous machine. In: 2017 IEEE URUCON, 2017, Montevideo. 2017 IEEE URUCON, 2017.

SILVA, ALAN H. F.; SILVA, UYARA F. ; CALIXTO, WESLEY P. ; MAGALHAES, ALANA S. ; ALVES, AYLTON J. . Conformal mapping applied to encoding and decoding of images. In: 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017, Pucon. 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017.

CARARO, JOSE A. G. ; SILVA, ALAN H. F. ; ANICETO, BRUNO C. ; REIS, MARCIO R. C. ; RODRIGUES, BRUNO U. ; GALVAO, NATALIA K. M. ; VILELA, WAGNER A. ; NETO, JOAO CAETANO ; CALIXTO, WESLEY P. . Optimizing of the insertion of distributed generation into a power distribution network. In: 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017, Pucon. 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017.

MAGALHAES, A. S. ; MORAES, P. H. F. ; SILVA, A. H. F. ; GOMES, P. H. G. ; ALVES, A. J. ; CALIXTO, W. P. . Parallel Operation Repowering of Synchronous and Induction Generator. In: 16 IEEE International Conference on Environment and Electrical Engineering (EEEIC 2016), 2016, Florença-Itália. 16 IEEE International Conference on Environment and Electrical Engineering (EEEIC 2016), 2016.

## 7.2 Sugestões para trabalhos futuros

- Aplicação da métrica em outros pares de idiomas;
- Refinamento da métrica na conferência de porções invertidas das frases;
- Implementação de palavras compostas;
- Aplicação de outros métodos de otimização;
- Aplicação de outras métricas no processo de tradução por otimização;
- Hibridização do processo de otimização;
- Hibridização da métrica com outro método utilizando inteligência artificial;
- Estudo da possibilidade da aplicação da rede BLEU no processo de avaliação de traduções, no processo de tradução e a possibilidade de utilizar a métrica proposta como mecanismo auxiliar na rede BLEU na expectativa de gerar resultados de classificação mais eficientes;
- Aplicação da tradução automática por processo de otimização utilizando inteligência artificial.

## REFERÊNCIAS BIBLIOGRÁFICAS

AC03561797. **Scientific and technical translating: and other aspects of the language problem.** [S.l.]: Unesco, 1957. 37

ALTIPARMAK, F.; GEN, M.; LIN, L.; PAKSOY, T. A genetic algorithm approach for multi-objective optimization of supply chain networks. **Computers & industrial engineering**, Elsevier, v. 51, n. 1, p. 196–215, 2006. 53, 58

ARAÚJO, W. R. H. d. **Potencialização do rendimento do gerador a relutância chaveado empregando técnica de rastreamento associada a controle de tensão otimizado.** Tese (Doutorado) — Universidade Federal de Goiás, 2020. 59

BABYCH, B. Automated mt evaluation metrics and their limitations. **Tradumática**, n. 12, p. 0464–470, 2014. 48

BANERJEE, S.; LAVIE, A. Meteor: An automatic metric for mt evaluation with improved correlation with human judgments. In: **Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization.** [S.l.: s.n.], 2005. p. 65–72. 50

BHATTACHARYYA, P. **Machine translation.** [S.l.]: Chapman and Hall/CRC, 2015. 40, 43

BROWN, P. F.; COCKE, J.; PIETRA, S. A. D.; PIETRA, V. J. D.; JELINEK, F.; LAFFERTY, J. D.; MERCER, R. L.; ROOSSIN, P. S. A statistical approach to machine translation. **Computational linguistics**, v. 16, n. 2, p. 79–85, 1990. 45, 46

CALISKAN, A.; GREENSTADT, R. Translate once, translate twice, translate thrice and attribute: Identifying authors and machine translation tools in translated text. In: IEEE. **Semantic Computing (ICSC), 2012 IEEE Sixth International Conference on.** [S.l.], 2012. p. 121–125. 25

CALIXTO, W. P.; ALVARENGA, B.; MOTA, J. C. da; BRITO, L. d. C.; WU, M.; ALVES, A. J.; NETO, L. M.; ANTUNES, C. F. Electromagnetic problems solving by conformal mapping: a mathematical operator for optimization. **mathematical problems in engineering**, Hindawi, v. 2010, 2010. 58

CALIXTO, W. P.; COIMBRA, A. P.; MOTA, J. C. d.; WU, M.; SILVA, W. G.; ALVARENGA, B.; BRITO, L. d. C.; ALVES, A. J.; DOMINGUES, E. G.; NETO, D. P. Troubleshooting in geoelectrical prospecting using real-coded genetic algorithm with chromosomal extrapolation. **International Journal of Numerical Modelling: Electronic Networks, Devices and Fields**, Wiley Online Library, v. 28, n. 1, p. 78–95, 2015. [54](#), [59](#)

CALIXTO, W. P.; MOTA, J. C. da; ALVARENGA, B. P. de. Methodology for the reduction of parameters in the inverse transformation of schwarz–christoffel applied to electromagnetic devices with axial geometry. **International Journal of Numerical Modelling: Electronic Networks, Devices and Fields**, Wiley Online Library, v. 24, n. 6, p. 568–582, 2011. [57](#)

CALIXTO, W. P.; NETO, L. M.; WU, M.; KLIEMANN, H. J.; CASTRO, S. S. de; YAMANAKA, K. Calculation of soil electrical conductivity using a genetic algorithm. **Computers and Electronics in Agriculture**, Elsevier, v. 71, n. 1, p. 1–6, 2010. [58](#)

CALIXTO, W. P.; NETO, L. M.; WU, M.; YAMANAKA, K.; MOREIRA, E. da P. Parameters estimation of a horizontal multilayer soil using genetic algorithm. **IEEE Transactions on Power Delivery**, IEEE, v. 25, n. 3, p. 1250–1257, 2010. [57](#), [60](#)

CALLISON-BURCH, C.; OSBORNE, M.; KOEHN, P. Re-evaluation the role of bleu in machine translation research. In: **11th Conference of the European Chapter of the Association for Computational Linguistics**. [S.l.: s.n.], 2006. [28](#), [29](#)

CHO, K.; MERRIËNBOER, B. V.; GULCEHRE, C.; BAHDANAU, D.; BOUGARES, F.; SCHWENK, H.; BENGIO, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. **arXiv preprint arXiv:1406.1078**, 2014. [27](#)

CLASSON, G. S. **The richest man in Babylon**. [S.l.]: Sristhi Publishers & Distributors, 2020. [77](#)

DANTZIG, G. B. **Origins of the simplex method**. [S.l.: s.n.], 1987. [56](#)

DAVE, S.; PARIKH, J.; BHATTACHARYYA, P. Interlingua-based english–hindi machine translation and language divergence. **Machine Translation**, Springer, v. 16, n. 4, p. 251–304, 2001. [41](#)

DODDINGTON, G. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In: MORGAN KAUFMANN PUBLISHERS INC. **Proceedings of the second international conference on Human Language Technology Research**. [S.l.], 2002. p. 138–145. 49, 50

DONG, T.; CREMERS, A. B. A novel machine translation method for learning chinese as a foreign language. In: SPRINGER. **International Conference on Intelligent Text Processing and Computational Linguistics**. [S.l.], 2014. p. 343–354. 26

DORR, B.; SNOVER, M.; MADNANI, N. Part 5: Machine translation evaluation. Citeseer, p. 801–894, 2018. 29, 48, 49

DOUGLAS, M. J. **Trading in the zone: Master the market with confidence, discipline and a winning attitude**. [S.l.]: Penguin, 2000. 69

FINCH, A. M.; AKIBA, Y.; SUMITA, E. How does automatic machine translation evaluation correlate with human scoring as the number of reference translations increases? In: **LREC**. [S.l.: s.n.], 2004. 49, 50

FLOUDAS, C. A. Deterministic global optimization in design, control, and computational chemistry. In: **Large-Scale Optimization with Applications**. [S.l.]: Springer, 1997. p. 129–184. 55

FORCADA, M. L.; GINESTÍ-ROSELL, M.; NORDFALK, J.; O'REGAN, J.; ORTIZ-ROJAS, S.; PÉREZ-ORTIZ, J. A.; SÁNCHEZ-MARTÍNEZ, F.; RAMÍREZ-SÁNCHEZ, G.; TYERS, F. M. Apertium: a free/open-source platform for rule-based machine translation. **Machine translation**, Springer, v. 25, n. 2, p. 127–144, 2011. 43

FOURNIER, B. **Preprocessing on bilingual data for Statistical Machine Translation**. Dissertação (Mestrado) — University of Twente, 2008. 45

GIL, J. R. B.; PYM, A. Technology and translation (a pedagogical overview). **Translation technology and its teaching**, Intercultural Studies Group, universitat rovíra i Virgili Tarragona, Spain, p. 5–20, 2006. 25

GUZMÁN, R. Advanced automatic mt post-editing. **Multiling Comput**, v. 19, n. 3, p. 52–57, 2008. 43, 44

HAFTKA, R. T.; WALSH, J. L. Stacking-sequence optimization for buckling of laminated plates by integer programming. **AIAA journal**, v. 30, n. 3, p. 814–819, 1992. 55

- HALLIDAY, M. A. K.; MATTHIESSEN, C. M. **Halliday's introduction to functional grammar**. [S.l.]: Routledge, 2013. 34
- HATIM, B.; MUNDAY, J. **Translation: An advanced resource book**. [S.l.]: Psychology Press, 2004. 31, 34
- HAUENSCHILD, C.; HEIZMANN, S. **Machine translation and translation theory**. [S.l.]: Walter de Gruyter, 2011. 40
- HILL, N. **Think and grow rich**. [S.l.]: Hachette UK, 2011. 77
- HOANG, V. C. D.; KOEHN, P.; HAFFARI, G.; COHN, T. Iterative back-translation for neural machine translation. In: **Proceedings of the 2nd Workshop on Neural Machine Translation and Generation**. [S.l.: s.n.], 2018. p. 18–24. 42
- HOLLAND, J. H. et al. **Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence**. [S.l.]: MIT press, 1992. 57
- HUTCHINS, J. Retrospect and prospect in computer-based translation. **Machine Translation Summit VII, 13th-17th September**, p. 30–34, 1999. 25
- \_\_\_\_\_. Multiple uses of machine translation and computerised translation tools. **Machine Translation**, p. 13–20, 2009. 25, 26
- HUTCHINS, W. J. **Machine translation: past, present, future**. [S.l.]: Ellis Horwood Chichester, 1986. 38, 42
- HUTCHINS, W. J.; SOMERS, H. L. **An introduction to machine translation**. [S.l.]: Academic Press London, 1992. 37, 39
- KALCHBRENNER, N.; BLUNSOM, P. Recurrent continuous translation models. In: **Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing**. [S.l.: s.n.], 2013. p. 1700–1709. 47
- KOEHN, P. A parallel corpus for statistical machine translation. **Proceedings of MT Summit X, September**, p. 12–16, 2009. 29
- KOEHN, P.; MONZ, C. Manual and automatic evaluation of machine translation between european languages. In: **Proceedings on the Workshop on Statistical Machine Translation**. [S.l.: s.n.], 2006. p. 102–121. 49

KOYAMA, N.; SATA, I.; FUKUMOCHI, Y.; SUZUKI, H. **Translation machine having capability of registering idioms**. [S.l.]: Google Patents, jul. 30 1996. US Patent 5,541,838. 26

KOZA, J. R. Genetic programming. Citeseer, 1997. 58, 59

KOZIEL, S.; YANG, X.-S. **Computational optimization, methods and algorithms**. [S.l.]: Springer, 2011. 53, 55

LAGARIAS, J. C.; REEDS, J. A.; WRIGHT, M. H.; WRIGHT, P. E. Convergence properties of the nelder–mead simplex method in low dimensions. **SIAM Journal on optimization**, SIAM, v. 9, n. 1, p. 112–147, 1998. 56

LAVIE, A.; DENKOWSKI, M. J. The meteor metric for automatic evaluation of machine translation. **Machine translation**, Springer, v. 23, n. 2-3, p. 105–115, 2009. 51

LEAVITT, J. R.; LONSDALE, D. W.; FRANZ, A. M. A reasoned interlingua for knowledge-based machine translation. In: CANADIAN INFORMATION PROCESSING SOCIETY. **Proceedings of the Biennial Conference-Canadian Society for Computational Studies of Intelligence**. [S.l.], 1994. p. 17–22. 41

LEE, A.; PRZYBOCKI, M. **NIST 2005 machine translation evaluation official results. official release of automatic evaluation scores for all submissions**. [S.l.]: August, 2005. 29

LEFFA, V. J. Making foreign language texts comprehensible for beginners: An experiment with an electronic glossary. **System**, Elsevier, v. 20, n. 1, p. 63–73, 1992. 26

LU, Y.; HUANG, J.; LIU, Q. Improving statistical machine translation performance by training data selection and optimization. In: **Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)**. [S.l.: s.n.], 2007. 27

LUONG, M.-T.; PHAM, H.; MANNING, C. D. Effective approaches to attention-based neural machine translation. **arXiv preprint arXiv:1508.04025**, 2015. 46, 47

LUONG, M.-T.; SUTSKEVER, I.; LE, Q. V.; VINYALS, O.; ZAREMBA, W. Addressing the rare word problem in neural machine translation. **arXiv preprint arXiv:1410.8206**, 2014. 28

MAGALHAES, A. S.; BULHOES, J. S.; REIS, M. R.; GOMES, V. M.; SILVA, A. H.; ALVES, A. J.; WAINER, G. A.; GANZAROLI, C. A.; ARAUJO, W. R.; CALIXTO, W. P. Experimental study of induction generator as a repowering solution. **International Transactions on Electrical Energy Systems**, Wiley Online Library, v. 30, n. 6, p. e12365, 2020. 54

MARCU, D.; WONG, D. A phrase-based, joint probability model for statistical machine translation. In: **Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)**. [S.l.: s.n.], 2002. 26

MATHUR, N.; BALDWIN, T.; COHN, T. Tangled up in bleu: Reevaluating the evaluation of automatic machine translation evaluation metrics. **arXiv preprint arXiv:2006.06264**, 2020. 29

NAGAO, M. A framework of a mechanical translation between japanese and english by analogy principle. **Artificial and human intelligence**, p. 351–354, 1984. 45

NELDER, J. A.; MEAD, R. A simplex method for function minimization. **The computer journal**, Oxford University Press, v. 7, n. 4, p. 308–313, 1965. 56

NESI, H. **Dictionaries in electronic form**. London: [s.n.], 2009. 25

NEWMARK, P. **A textbook of translation**. [S.l.]: Prentice hall New York, 1988. 34, 35

NIRENBURG, S.; DOMASHNEV, C.; GRANNES, D. J. Two approaches to matching in example-based machine translation. In: **In Proc. of the 5th International Conference on Theoretical and Methodological Issues in Machine Translation**. [S.l.: s.n.], 1993. 38, 45

OKPOR, M. Machine translation approaches: issues and challenges. **International Journal of Computer Science Issues (IJCSI)**, International Journal of Computer Science Issues (IJCSI), v. 11, n. 5, p. 159, 2014. 38

OLIVERI, G.; ZHONG, Y.; CHEN, X.; MASSA, A. Multiresolution subspace-based optimization method for inverse scattering problems. **JOSA A**, Optical Society of America, v. 28, n. 10, p. 2057–2069, 2011. 55

- PAPINENI, K.; ROUKOS, S.; WARD, T.; ZHU, W.-J. Bleu: a method for automatic evaluation of machine translation. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. **Proceedings of the 40th annual meeting on association for computational linguistics**. [S.l.], 2002. p. 311–318. 28, 48
- PRASAD, T. V.; MUTHUKUMARAN, G. M. Telugu to english translation using direct machine translation approach. **International Journal of Science and Engineering Investigations**, v. 2, n. 12, p. 25–32, 2013. 38
- RAO, J. R.; ROHATGI, P. et al. Can pseudonymity really guarantee privacy? In: **USENIX Security Symposium**. [S.l.: s.n.], 2000. p. 85–96. 25
- SHUTTLEWORTH, M. **Dictionary of translation studies**. [S.l.]: Routledge, 2014. 31
- SHUTTLEWORTH, M.; COWIE, M. **Dictionary of translation studies**. manchester: St. **Jerome Publishing**, v. 192, p. 193, 1997. 33
- SIMARD, M.; UEFFING, N.; ISABELLE, P.; KUHN, R. Rule-based translation with statistical phrase-based post-editing. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. **Proceedings of the Second Workshop on Statistical Machine Translation**. [S.l.], 2007. p. 203–206. 44
- SLOCUM, J. A survey of machine translation: its history, current status, and future prospects. **Computational linguistics**, MIT Press, v. 11, n. 1, p. 1–17, 1985. 37
- SOANES, C.; STEVENSON, A. **Concise oxford English dictionary**. [S.l.]: Oxford University Press Oxford, 2004. 31
- SPINKS, N.; WELLS, B. Intercultural communication: a key element in global strategies. **Career Development International**, MCB UP Ltd, v. 2, n. 6, p. 287–292, 1997. 37
- VINAY, J.-P.; DARBELNET, J. **Comparative stylistics of French and English: A methodology for translation**. [S.l.]: John Benjamins Publishing, 1995. 33
- WAGNER, R. A.; FISCHER, M. J. The string-to-string correction problem. **Journal of the ACM (JACM)**, ACM New York, NY, USA, v. 21, n. 1, p. 168–173, 1974. 63

WALL, R. E. Some of the engineering aspects of the machine translation of language. **Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics**, IEEE, v. 75, n. 5, p. 580–585, 1956. 25

WANG, X.-X.; ZHU, C.-H.; LI, S.; ZHAO, T.-J.; ZHENG, D.-Q. Neural machine translation research based on the semantic vector of the tri-lingual parallel corpus. In: **IEEE. 2016 International Conference on Machine Learning and Cybernetics (ICMLC)**. [S.l.], 2016. v. 1, p. 69–74. 27

WEAVER, W. Translation. **Machine translation of languages**, Cambridge: Technology Press, MIT, v. 14, p. 15–23, 1955. 25, 37

WILKS, Y. **Machine translation: its scope and limits**. [S.l.]: Springer Science & Business Media, 2008. 41

WRIGHT, M. H. et al. Nelder, mead, and the other simplex method. **Documenta Mathematica**, v. 7, p. 271–276, 2010. 56

YANG, S.; WANG, Y.; CHU, X. A survey of deep learning techniques for neural machine translation. **arXiv preprint arXiv:2002.07526**, 2020. 27

ZHAO, W.; GLAVAŠ, G.; PEYRARD, M.; GAO, Y.; WEST, R.; EGER, S. On the limitations of cross-lingual encoders as exposed by reference-free machine translation evaluation. **arXiv preprint arXiv:2005.01196**, 2020. 29