

CoNLL Dependency Parser: Extrinsic Evaluation through the Open Information Extraction task

Jardel Baia², Arley Prates¹, Daniela Barreiro Claro¹

¹ FORMAS - Formalismos e Aplicações Semânticas Research Group
Departamento de Ciência da Computação - IME - Universidade Federal da Bahia

² FORMAS - Formalismos e Aplicações Semânticas Research Group
Departamento de Estatística - IME - Universidade Federal da Bahia
{jardel.baia, arley.prates, dclaro}@ufba.br

Abstract. Dependency Parsers (DP) are parsers that analyze dependencies between words in a sentence. Currently, dependency parser evaluation is a problem whose solutions are not well defined in the scientific community. Although the DP intrinsic metrics are the foremost choice of evaluation, extrinsic evaluation enables a different evaluation based on a downstream. Different results of DP can vary according to the domain task. Thus, this work applies an Open Information Extraction (OIE) method in Portuguese to provide an extrinsic evaluation of a set of CoNLL Dependency Parsers. Our results demonstrate that there is a difference in the evaluation of Dependency Parsers considering a particular task.

CCS Concepts: • **Computing methodologies** → **Natural language processing**.

Keywords: dependency parser, extrinsic evaluation, open information extraction

1. INTRODUÇÃO

A Análise de Dependência permite analisar sintaticamente uma sentença através da relação entre as estruturas sintáticas de uma sentença. Com o avanço de recursos e técnicas nessa tarefa ao longo dos anos [Collins et al. 1999, Nivre et al. 2005, McDonald et al. 2005], Analisadores de Dependência (AD) obtiveram um crescente espaço nas diversas tarefas empregadas pelo Processamento de Linguagem Natural (PLN). Com a aplicação de novos métodos [Dozat et al. 2017, Zeman et al. 2018] e a proposição e utilização de diferentes métricas [Gamallo and Garcia 2018, Wanderley et al. 2019], inicia-se uma análise sobre as diferentes abordagens para avaliar um Analisador de Dependência.

As avaliações de Analisadores de Dependência podem ser intrínsecas ou extrínsecas. As avaliações intrínsecas correspondem a um conjunto de métricas de qualidade enquanto que a avaliação extrínseca ocorre por meio de uma aplicação em um contexto de uso. Segundo os autores em [Jones 1994], ambas as formas de avaliação são necessárias para compreender e analisar um sistema aplicável em contextos diferentes. A avaliação extrínseca, na qual o analisador é avaliado pelo seu desempenho em uma determinada tarefa, é menos popular do que a avaliação intrínseca, que utiliza dados pré-annotados, e aplica métricas como a Labeled Attachment Accuracy (LAS), que anota a porcentagem de tokens que apontam corretamente para o nó pai e possuem a relação de dependência rotulados corretamente [Zeman et al. 2018].

Como inicialmente proposto por [Gamallo and Garcia 2017], este trabalho tem o objetivo de utilizar a tarefa de Extração de Informação Aberta (EIA) para a avaliação extrínseca de ADs. O desempenho de um sistema do estado da arte de EIA para a língua Portuguesa, denominado dptOIE [Glauber et al.

Copyright©2020 Permission to copy without fee all or part of the material printed in KDMiLe is granted provided that the copies are not made or distributed for commercial advantage, and that notice is given that copying is by permission of the Sociedade Brasileira de Computação.

2019], foi comparado ao receber como entrada as anotações realizadas por diferentes ADs disponíveis na *shared task* da CoNLL para a língua portuguesa. Assim, é possível comparar de forma contextualizada o desempenho dos analisadores de dependência com suas diferentes implementações e variadas escolhas linguísticas, tais como *tokenização*, *Pos tagger*, etc.

Este trabalho está organizado da seguinte forma: a seção 2 apresenta os trabalhos relacionados. A seção 3 introduz a tarefa EIA utilizada no método de avaliação extrínseca. A seção 4 descreve os materiais e métodos. A seção 5 descreve os experimentos e as ferramentas utilizadas, além dos resultados obtidos. E, por fim, a seção 6 apresenta as conclusões e os trabalhos futuros.

2. TRABALHOS RELACIONADOS

A avaliação de Analisadores de Dependência (AD) tem se mostrado bem relevante para as tarefas de Processamento de Linguagem Natural visto que muitas abordagens se utilizam desses tipos de analisadores para alcançar os resultados. Um conjunto de métricas são necessárias para comparar diferentes analisadores de dependência. Diversas métricas tem sido propostas, sendo adotadas para avaliação destes ADs [Carroll et al. 1998; Marcus et al. 1993].

A avaliação intrínseca de um AD estabeleceu um padrão a partir de sucessivas edições da *SIGNLL Conference on Computational Natural Language Learning - CoNLL*¹, relacionada a tarefa de avaliação de Analisadores de Dependência iniciada em 2006 [Buchholz and Marsi 2006]. As principais métricas de avaliação intrínseca, como *Labeled Attachment Accuracy* (LAS) e *Unlabeled Attachment Accuracy* (UAS), foram adotadas nas tarefas da CoNLL ao longo dos anos. No entanto, uma avaliação intrínseca de um AD não é adequada para avaliar sistemas com diferentes arcabouços linguísticos. Dessa forma, a partir de 2017 e de 2018, vem sendo adotado uma tarefa conjunta de avaliação extrínseca dos Analisadores de Dependência proposta por [Oepen et al. 2017; Fares et al. 2018]. Nessa tarefa, as avaliações extrínsecas foram realizadas nos seguintes contextos: extração de evento biomédico, resolução de negação e análise de opinião.

A primeira avaliação extrínseca para a língua portuguesa baseada na tarefa de Extração de Informação Aberta foi proposta por [Gamallo and Garcia 2018]. Porém, o método de EIA empregado dificultou a avaliação pois muitas triplas foram inválidas ou não foram extraídas. Ampliando a abordagem, os autores em [Wanderley et al. 2019] propuseram uma avaliação extrínseca baseado nos ADs da *ConLL Shared Task 2017*. Porém, os ADs utilizados já não mais correspondiam ao estado da arte para a língua portuguesa. Diferentemente dos demais trabalhos, este artigo propõe a utilização dos mais recentes Analisadores de Dependência da *ConLL Shared Task 2018*, além da padronização já adotada nos dados de entrada dos ADs que permitem mais ainda generalizar as tarefas que possuem as mesmas características de um EIA. Além disso, ajustes no método de extração de triplas foram realizadas com o intuito de melhorar a qualidade das extrações.

3. ANALISADORES DE DEPENDÊNCIA

Analisadores de Dependência ou *Dependency Parser* no inglês, realizam uma análise sintática na qual a estrutura da sentença é descrita não pela forma de suas componentes (ou constituintes), mas pelas relações gramaticais/funcionais entre as palavras da sentença [Wanderley et al. 2019]. Essa análise quando automatizada geralmente resulta em uma árvore de dependência sintática.

Uma sentença é analisada na Figura 1 através da ferramenta denominada *Dependency Viewer* [Tang 2012]. As setas representadas indicam as relações entre os tokens enquanto que as palavras entre as setas representam a análise sintática.² A sentença "Fulana é professora da Universidade" analisada

¹<https://www.conll.org/>

²O significado da análise sintática de cada palavra está disponível em: <https://universaldependencies.org/pt/dep/>

pelo AD na Figura 1, apresenta como raiz (*root*) o substantivo (NOUN) *professora*. Este substantivo conecta o sujeito *Fulana* (NSUBJ) ao modificador nominal *Universidade* (NMOD).

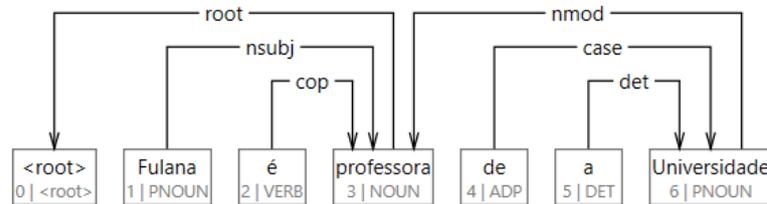


Fig. 1. Exemplo de uma sentença analisada por um AD através da ferramenta *Dependency Viewer*[Tang 2012].

Analisadores de Dependência não processam diretamente uma sentença para obter a árvore sintática, mas utilizam outras técnicas de Processamento de Linguagem Natural (PLN), que transformam os dados de entrada em um formato aceito pelo analisador. Da mesma forma, a saída do analisador ou seja, a árvore sintática servirá para outro método de PLN.

Um desses métodos da PLN são as técnicas de extração de informação. Segundo [Glauber et al. 2019], a utilização de Analisadores de Dependência apresentou resultados superiores a outras técnicas de *shallow parsers*. Assim, os sistemas de Extração de Informação Aberta têm utilizado ADs para ampliar a qualidade das triplas extraídas e consequentemente a sua reutilização em outras tarefas da PLN.

3.1 A tarefa de Extração de Informação Aberta (EIA)

A Extração de Informação Aberta é uma tarefa da PLN que permite extrair informação estruturada de textos não estruturados. A EIA apresenta algumas vantagens com relação a Extração de Informação tradicional ou também denominada EI Fechada. A Extração de Informação Aberta não possui a necessidade de dados pré-annotados e possibilita grande escalabilidade [Banko et al. 2007].

Sistemas de EIA podem ser divididos em duas categorias amplas: sistemas baseados em regras pré-definidas e baseados em aprendizagem automática [Gamallo et al. 2012]. Os sistemas que fazem análise rasa da sentença, geralmente através de POS taggers ou chunkers, normalmente obtêm precisão alta, mas possuem baixa cobertura. As abordagens que utilizam análise de dependência ou análise profunda, normalmente tem custo maior de processamento, no entanto, apresentam um melhor desempenho tanto na precisão quanto na cobertura. Sendo assim, autores em [Glauber et al. 2019] afirmam que o desempenho dos analisadores de dependência podem exercer grande influência nos sistemas de EIA.

4. MATERIAIS E MÉTODOS

A utilização dos Analisadores de Dependência na tarefa de Extração de Informação Aberta (EIA) permite realizar uma avaliação extrínseca desses analisadores. Esse tipo de abordagem, permite a avaliação de sistemas que utilizaram técnicas completamente diferentes sob as mesmas condições. Os materiais e os conjuntos de dados utilizados neste trabalho são apontados, além das métricas de avaliação apresentadas.

4.1 Conjunto de dados

O conjunto de dados utilizado para treino dos Analisadores de Dependência foi o *UD_Portuguese-Bosque*³. Esse *treebank* possui sentenças com as árvores de dependência anotadas. O conjunto de dados cujas árvores de dependência foram anotadas pelos ADs consistem em um *corpus* de 100 sentenças aleatoriamente selecionadas do Corpus⁴ de Extratos de Textos Eletrônicos NILC/Folha de S. Paulo (*CETENFolha*) e da *Wikipedia*.

4.2 Analisadores de Dependência selecionados

Para a avaliação, 4 Analisadores de Dependência de equipes participantes da CoNLL Shared Task 2018 foram selecionados. Essa seleção foi realizada com base nos melhores resultados de LAS para português brasileiro e, também, seguindo os critérios: (i) código-fonte disponibilizado e (ii) possibilidade de replicar o treino do modelo com os dados da tarefa. Após a aplicação desses critérios, as equipes e seus respectivos Analisadores de Dependência escolhidos para a tarefa foram: *Stanford* [Qi et al. 2018], *CEA LIST* [Duthoo and Mesnard 2018], *ICS PAS* [Rybak and Wróblewska 2018] e *TurkuNLP* [Kanerva et al. 2018]. A tabela I apresenta o desempenho com a métrica de LAS de cada Analisador de Dependência na CoNLL Shared Task 2018, além das respectivas posições no ranking de LAS para o português.

Parsers	LAS	Posição
Stanford	87.81	1
CEA LIST	87.72	2
ICS PAS	87.57	4
TurkuNLP	87.40	5

Table I. Desempenho dos Analisadores de Dependência selecionados da CoNLL *Shared Task* 2018.

4.3 Método de Extração de Informação Aberta

O sistema de Extração de Informação Aberta (EIA) utilizado foi o DptOIE [Glauber et al. 2019], um sistema de EIA que recebe como entrada textos sintaticamente anotados no padrão CoNLL-U para árvores de dependência. A partir do conjunto de rótulos das Universal Dependencies, o DptOIE extrai relações abertas para a língua portuguesa aplicando um algoritmo de busca em profundidade (*depth-first search*), através de regras de construções gramaticais. Além disso, o DptOIE realiza o tratamento de casos como conjunções coordenadas, orações subordinadas e apostos.

Neste trabalho, todas as extrações de triplas foram realizadas através do DptOIE. Como é usual em sistemas de EIA, o DptOIE extrai fatos no formato de triplas. Essas triplas possuem o formato $t = (e1, r, e2)$, os quais $e1$ e $e2$ representam entidades e r , a relação entre elas. Uma extração possível, na sentença "Alan Turing nasceu em Londres", é (*Alan Turing* ($e1$), *nasceu* (r), *em Londres* ($e2$)).

4.4 Métricas

Os fatos extraídos pelo DptOIE foram julgados como *válido* ou *inválido* por 2 avaliadores com experiência na área de EIA. As extrações foram consideradas válidas somente com a concordância de ambos os avaliadores. Na sentença "Fulano é aluno de mestrado da Universidade" por meio da EIA, poderia ter as seguintes extrações: ("Fulano", "é aluno", "da Universidade") e ("Fulano", "é aluno", "de mestrado"). Nessa extrações "é aluno" é responsável por relacionar os argumentos, "Fulano" com "da Universidade", ou "Fulano" com "de mestrado". Em ambos os casos as extrações seriam triplas

³https://github.com/UniversalDependencies/UD_Portuguese-Bosque

⁴https://www.linguateca.pt/cetenfolha/index_info.html

válidas. Já na sentença "Se ele tirar 10 na prova, Fulano será aprovado na disciplina", alguns sistemas de EIA, poderiam extrair a tripla ("Fulano", "será aprovado", "na disciplina"). Esta extração é considerada inválida, pois ela está condicionada a outra informação, (Se ele tirar 10 na prova).

Com a finalidade de avaliar o desempenho dos ADs na tarefa de Extração de Informação Aberta, diversas métricas foram adotadas. Essas métricas foram aplicadas após a obtenção dos fatos em que os avaliadores concordaram, sendo esses fatos válidos ou inválidos. As métricas foram:

- (1) quantidade de fatos extraídos;
- (2) *yield* - a quantidade de fatos extraídos considerados válidos;
- (3) *yield* negativo - a quantidade de fatos que ambos os avaliadores consideraram inválidos;
- (4) quantidade de sentenças sem extração de fatos;
- (5) precisão - a representação dos fatos válidos da extração de um AD entre todas os fatos extraídos desse AD. É determinada pela fração de fatos válidos entre todas os fatos;
- (6) cobertura - a representatividade das extrações válidas de um AD se comparadas ao total de extrações válidas, entre todos os ADs. É calculada através da fração da quantidade de fatos válidos de um AD e o somatório dos fatos válidos de todos os ADs;
- (7) medida-f - a relação entre precisão e cobertura.

5. EXPERIMENTOS E RESULTADOS

Considerando a finalidade de realizar a avaliação extrínseca de Analisadores de Dependência, a tarefa de Extração de Informação Aberta foi adotada.

5.1 Experimento A

Depois de realizada a seleção dos Analisadores de Dependência, baseando-se no ranking da CoNLL *Shared Task* 2018, os ADs foram treinados novamente com o mesmo *treebank* utilizado na tarefa para o português. Após este treinamento, os mesmos foram utilizados para fornecer as árvores de dependência para um *corpus* de 100 sentenças.

5.2 Experimento B

Após o experimento A, em que 4 árvores de dependência foram geradas para o mesmo corpus, o *dptOIE* foi utilizado para a tarefa de Extração de Informação Aberta. Com os fatos extraídos, 2 avaliadores julgaram as sentenças em válidas ou inválidas. Com os resultados adquiridos, algumas métricas foram aplicadas para avaliar o desempenho de cada Analisador de Dependência na tarefa.

5.3 Resultados

Um total de 667 fatos foram extraídos do corpus de 100 sentenças, entre extrações válidas e inválidas. Na Figura 2, é possível observar que o *CEA-LIST* obteve a maior quantidade de extrações, além do maior *yield* e maior quantidade de fatos inválidos. A menor quantidade de extrações e menor *yield* pertencem ao analisador *Stanford*. No total, 103 fatos foram considerados válidos no conjunto de todos os fatos extraídos.

Como mostra a tabela II, apesar do AD *TurkuNLP* obter uma quantidade maior de extrações que o analisador *Stanford*, os seus *yields* foram praticamente iguais, além da precisão e a medida-f do *Stanford* terem sido superior. É observável também, na tabela II e na figura 3 que, com menos extrações e menor *yield*, o AD de *Stanford* obteve quase a mesma precisão que o *ICS-PAS*. O *CEA-LIST* apresentou o melhor desempenho geral, sendo que a única métrica que não obteve o melhor resultado foi a precisão.

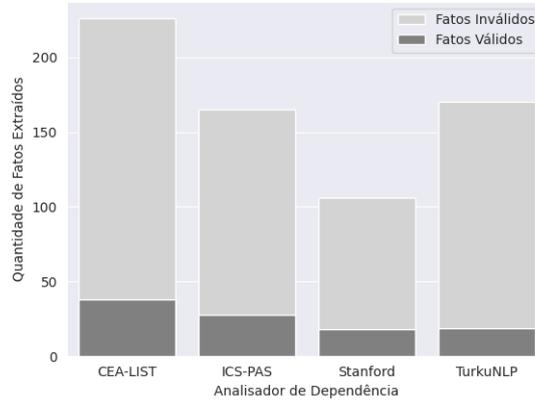


Fig. 2. Quantidade de fatos extraídos de cada Analisador de Dependência.

AD	Extrações	Yield	Precisão	Cobertura	Medida-F
CEA-LIST	226	38	0.168	0.161	0.164
ICS-PAS	165	28	0.169	0.119	0.139
Stanford	106	18	0.169	0.076	0.105
TurkuNLP	170	19	0.111	0.080	0.093

Table II. Resultados das métricas aplicadas aos Analisadores de Dependência.

Na figura 3 é perceptível a diferença no desempenho do *CEA-LIST*, mantendo uma precisão alta com uma quantidade de fatos extraídos superior.

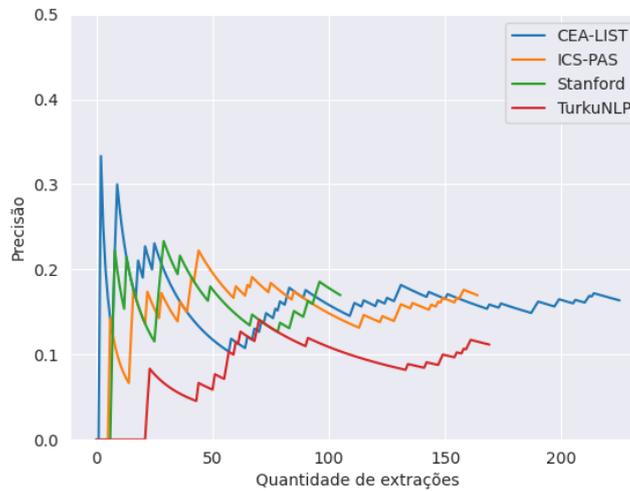


Fig. 3. Relação entre precisão e quantidade de fatos extraídos dos Analisadores de Dependência.

Observando a figura 4, o AD *Stanford* foi o que obteve um maior número de sentenças sem extração de nenhum fato (61 sentenças). Os ADs *TurkuNLP* e *ICS-PAS* também obtiveram um número relevante de sentenças sem nenhuma extração: 52 e 46, respectivamente. O *CEA-LIST* apresentou o menor número de sentenças sem extrações (30).

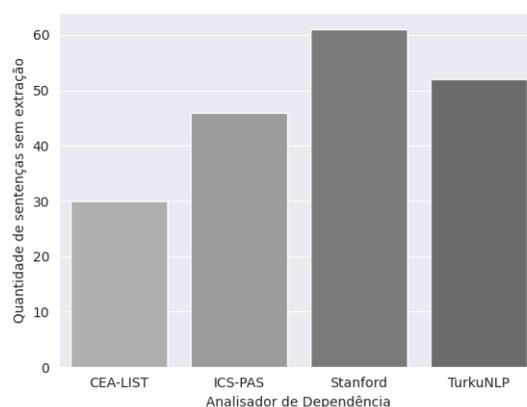


Fig. 4. Quantidade de sentenças sem extração em cada Analisador de Dependência.

5.4 Discussões

Dentre todas as sentenças, 5 delas não obtiveram nenhuma extração dentre as saídas de todos os ADs avaliados, ou seja, nessas 5 sentenças o *dptOIE* não conseguiu extrair fatos de nenhuma das árvores de dependência geradas por cada AD. Essas sentenças foram:

Id	Sentença
1	Passa por muitas praias da Costa da Prata, entre elas: Valadares, Espinho, Esmoriz, Ovar, Mira, Figueira da Foz e Termas de Monte Real.
2	Exerceu várias vezes o cargo de deputado provincial de Pernambuco
3	Músicos: Edgard Scandurra, Paulo Tatit, Zaba Moreau, Pedro Ito e Peter Price.
4	As embalagens, 33%.
5	Com isso, só falta marcar a data exata para a convocação extra dos parlamentares

Table III. Sentenças sem extração em todos os Analisadores de Dependência.

O Analisador de Dependência responsável pela anotação da árvore de dependências que obteve o maior número de extrações no experimento foi o *CEA-LIST*. Essas extrações (226) representam uma grande parcela do total (667).

O Analisador de Dependência responsável pelo menor número de extrações no experimento foi o *Stanford*. Esta é uma observação importante, pois o *Stanford* foi o primeiro colocado no *ranking* de LAS para o português.

6. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho, realizou-se uma avaliação extrínseca dos Analisadores de Dependência da CoNLL Shared Task 2018 que tiveram melhores desempenho na tarefa de avaliação intrínseca. Observando que a melhor medida-f apresentada, além da maior quantidade de extrações e maior *yield*, pertence ao *CEA-LIST*, é notável a importância da avaliação extrínseca de analisadores, levando em conta que o *CEA-LIST*, na segunda posição do *ranking* de LAS da tarefa da CoNLL, obteve um resultado superior ao analisador *Stanford*, que foi o primeiro no *ranking*.

A avaliação intrínseca de analisadores apresenta métricas importantes para o desempenho de ADs. No entanto, uma abordagem mista, avaliando o desempenho do AD intrinsecamente e também em uma determinada tarefa (extrinsecamente), pode resultar em melhores escolhas de ADs no contexto em que se deseja aplicá-lo.

Como trabalhos futuros, pretende-se explorar sentenças que não apresentaram nenhuma extração, ou não apresentaram extrações válidas, gerando alternativas às atuais métricas utilizadas na tarefa de EIA. Além disso, essa mesma observação pode resultar em importantes melhorias na ferramenta utilizada para a realização da extração de fatos.⁵

REFERENCES

- BANKO, M., CAFARELLA, M. J., SODERLAND, S., BROADHEAD, M., AND ETZIONI, O. Open information extraction from the web. In *IJCAI*. Vol. 7. pp. 2670–2676, 2007.
- BUCHHOLZ, S. AND MARSI, E. CoNLL-x shared task on multilingual dependency parsing. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*. Association for Computational Linguistics, New York City, pp. 149–164, 2006.
- CARROLL, J., BRISCOE, T., AND SANFILIPPO, A. Parser evaluation: a survey and a new proposal, 1998.
- DUTHOO, E. AND MESNARD, O. CEA LIST: Processing low-resource languages for CoNLL 2018. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Brussels, Belgium, pp. 34–44, 2018.
- FARES, M., OEPEEN, S., ØVRELID, L., BJÖRNE, J., AND JOHANSSON, R. The 2018 shared task on extrinsic parser evaluation: On the downstream utility of English universal dependency parsers. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Brussels, Belgium, pp. 22–33, 2018.
- GAMALLO, P. AND GARCIA, M. Linguakit: uma ferramenta multilingue para a análise linguística e a extração de informação. *Linguamática* 9 (1): 19–28, 2017.
- GAMALLO, P. AND GARCIA, M. Task-oriented evaluation of dependency parsing with open information extraction. In *Computational Processing of the Portuguese Language*. Springer International Publishing, Cham, pp. 77–82, 2018.
- GAMALLO, P., GARCIA, M., AND FERNÁNDEZ-LANZA, S. Dependency-based open information extraction. In *Proceedings of the Joint Workshop on Unsupervised and Semi-Supervised Learning in NLP*. Association for Computational Linguistics, pp. 10–18, 2012.
- GLAUBER, R., CLARO, D., AND OLIVEIRA, L. Dependency parser on open information extraction for portuguese texts -dptoie and dependentie on iberlef. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2019) co-located with 35th Conference of the Spanish Society for Natural Language Processing (SEPLN 2019)*. Vol. 2421. CEUR Workshops, pp. 442–448, 2019.
- JONES, K. S. Towards better nlp system evaluation. In *HUMAN LANGUAGE TECHNOLOGY: Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994*, 1994.
- KANERVA, J., GINTER, F., MIEKKA, N., LEINO, A., AND SALAKOSKI, T. Turku neural parser pipeline: An end-to-end system for the CoNLL 2018 shared task. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Brussels, Belgium, pp. 133–142, 2018.
- MARCUS, M. P., SANTORINI, B., AND MARCINKIEWICZ, M. A. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* 19 (2): 313–330, 1993.
- OEPEEN, S., ØVRELID, L., BJÖRNE, J., JOHANSSON, R., LAPPONI, E., GINTER, F., AND VELLDAL, E. The 2017 shared task on extrinsic parser evaluation. towards a reusable community infrastructure. In *Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation at the Fourth International Conference on Dependency Linguistics and the 15th International Conference on Parsing Technologies*. Pisa, Italy. pp. 1–16, 2017.
- QI, P., DOZAT, T., ZHANG, Y., AND MANNING, C. D. Universal dependency parsing from scratch. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Brussels, Belgium, pp. 160–170, 2018.
- RYBAK, P. AND WRÓBLEWSKA, A. Semi-supervised neural system for tagging, parsing and lematization. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Brussels, Belgium, pp. 45–54, 2018.
- TANG, G. Dependency viewer, 2012.
- WANDERLEY, M., CLARO, D. B., SOUZA, M., AND DE OLIVEIRA, L. Avaliação extrínseca de analisadores de dependência através da extração de informação aberta. In *Proceedings of the Symposium in Information and Human Language Technology*. Symposium in Information and Human Language Technology, pp. 171–180, 2019.
- ZEMAN, D., HAJIĆ, J., POPEL, M., POTTHAST, M., STRAKA, M., GINTER, F., NIVRE, J., AND PETROV, S. Conll 2018 shared task: Multilingual parsing from raw text to universal dependencies. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. pp. 1–21, 2018.

⁵Os autores agradecem ao PIBIC-CNPQ e a CAPES pelo apoio financeiro.