

## Estudos recentes sobre a detecção de contradição no Processamento Automático de Línguas Naturais

Denis Luiz Marcello Owa<sup>1</sup>

<sup>1</sup>Doutorando em Linguística pela Universidade Federal de Sao Carlos  
denismarcello@gmail.com

***Resumo.** Neste artigo, são apresentados dois estudos realizados sobre como o Processamento Automático de Línguas Naturais pode ser utilizado para se detectar contradição em textos em inglês, mas que servem de base para outras línguas. Os objetivos são entender dois sistemas já desenvolvidos para essa finalidade e alguns dos problemas enfrentados pelos pesquisadores. Podemos notar que ainda há muito o que se pesquisar nesse campo, uma vez que os sistemas de detecção de contradição devem possuir não só uma ampla base de dados semântica, mas também uma grande base de dados sobre conhecimento de mundo.*

*Palavras-chave:* Processamento Automático de Línguas Naturais, Linguística Computacional, Linguística Aplicada, Contradição

### 1. Introdução

Quando elaboramos textos orais ou escritos, possivelmente entraremos em contradição em pontos dispersos desse texto. Em alguns casos, por leve descuido, não percebemos que estamos entrando em contradição, mas pessoas atentas percebem quando ela existe. Mas um computador pode ser treinado para detectar a contradição?

Seja em discursos gerais de políticos, como debates pré-eleitorais, ou seja, em dissertações de mestrado, teses de doutorado ou mesmo depoimentos feitos à polícia, o computador pode ter grande importância para se detectar contradição nesses textos.

[MÜLLER 2004] apresentam que a contradição se localiza dentro da Semântica Formal. A contradição ocorre quando duas expressões apresentam sentidos incompatíveis com a mesma situação. Vejamos alguns exemplos:

- (1) Santos Dumont nasceu no estado brasileiro de Minas Gerais.
- (2) Santos Dumont não nasceu no Brasil.
- (3) João chegou de táxi à faculdade.
- (4) João chegou a pé à faculdade.

Esses exemplos servem para afirmar a noção de que duas sentenças são contraditórias quando ambas não podem ser simultaneamente verdadeiras. É o caso dos pares (1) e (2). Se na sentença (1) dizemos que Santos Dumont nasceu no estado brasileiro de Minas Gerais, a sentença (2) entra em contradição com a sentença (1). Afinal de contas, Minas Gerais pertence ao Brasil.

Pela sentença (3), sabemos que João chegou de táxi à faculdade. Mas a sentença (4) contradiz o que foi transmitido pela sentença (3). Se João chegou de táxi à faculdade, logo ele não chegou a pé. Ou seja, duas sentenças são contraditórias quando uma delas

é verdadeira e a outra é falsa. Por outro lado, existem casos em que duas sentenças não entram em contradição. Vejamos outros exemplos:

(5) Santos Dumont nasceu em Minas Gerais.

(6) Santos Dumont morreu em São Paulo.

(7) João comprou o celular.

(8) João vendeu o celular.

Os exemplos acima mostram itens lexicais opostos (nasceu/morreu, comprou/vendeu), mas não apresentam contradição. *Nascer* e *morrer* são pontos extremos no processo de viver. O par *comprar* e *vender*, por sua vez, são resultados obtidos por duas ações. Esses exemplos demonstram que a antonímia pode não envolver contradição. No entanto, a antonímia pode envolver contradição nos seguintes casos:

(9) Aquele computador é rápido.

(10) Aquele computador é lento.

(11) Minha casa é grande.

(12) Minha casa é pequena.

O par (9) e (10) é contraditório porque surge aqui a noção de acarretamento (*entailment*): *Aquele computador é rápido* acarreta *Aquele computador não é lento*. O mesmo ocorre com o par (11) e (12). *Minha casa é grande* acarreta *minha casa não é pequena*.

Apresentaremos os estudos de [DE MARNEFFE 2008] e [RITTER 2008] sobre a tarefa de se detectar contradições em textos por meio do Processamento Automático de Línguas Naturais. O artigo de [RITTER 2008] procura complementar alguns métodos de [DE MARNEFFE 2008], pois tenta trabalhar com relações de função.

## 2. O que é contradição?

[DE MARNEFFE 2008] definem que, por padrão, a contradição ocorre quando uma determinada frase A e uma determinada frase B não podem ser ambas verdadeiras.

Vejamos um exemplo dos próprios autores:

(13) A polícia especializada em bombas desarmou os explosivos. Por volta de 100 pessoas estavam trabalhando na fábrica.

(14) 100 pessoas ficaram feridas.

A contradição ocorre porque entendemos pela frase (13) que os explosivos foram desarmados e, conseqüentemente, ninguém ficou ferido. Dessa forma, a frase (14) entra em contradição com a frase (13), desde que essas duas sentenças se refiram ao mesmo evento.

Para [DE MARNEFFE 2008], para ocorrer uma contradição, deve-se ter como referência o mesmo evento. Vejamos outro exemplo:

(15) Sérgio vendeu um barco para João.

(16) João vendeu um barco para Sérgio.

Caso as frases (15) e (16) se refiram ao mesmo evento (o evento de uma venda de um determinado barco), elas serão contraditórias. No entanto, caso se refiram a eventos diferentes, deixam de ser contraditórias, uma vez que essas sentenças podem ser lidas como *Sérgio vendeu um barco para João e João vendeu um outro barco para Sérgio* (ou, João vendeu o mesmo barco em outra ocasião posterior).

### 3. O estudo apresentado por [DE MARNEFFE 2008]

#### 3.1. Tipologia das contradições, segundo [DE MARNEFFE 2008]

Os autores propõem uma tipologia de classes de contradição. Esses tipos de contradição são divididos em duas categorias: as de detecção fácil e as de detecção complexa. Vejamos a primeira categoria:

a) antonímia

(21) Quando a criança brinca com *videogame*, ela desenvolve o raciocínio lógico.

(22) Quando a criança brinca com *videogame*, ela prejudica o raciocínio lógico.

b) negação

(23) O tribunal, praticamente dividido, decidiu que o júri, e não os juízes, deve dar a sentença.

(24) O tribunal decidiu que somente os juízes podem dar a sentença.

c) numérico

(25) A tragédia causada pela explosão na cidade libanesa de Qana, que matou mais de 50 civis, colocou Israel num dilema.

(26) Uma investigação sobre o ataque em Qana confirmou 28 mortes até agora.

A segunda categoria compreende os tipos mais complexos de serem detectados. Vejamos:

d) factivo

(27) O primeiro-ministro australiano John Howard diz que não se deixará ser influenciado pela ameaça de sofrer mais ataques terroristas caso não retire suas tropas do Iraque.

(28) A Austrália retira-se do Iraque.

e) estrutural

(29) Jacques Santer sucedeu Jacques Delors como presidente da Comissão Europeia, em 1995.

(30) Delors sucedeu Santer na presidência da Comissão Europeia.

Outro exemplo de contradição estrutural ocorre em:

(31) O Eurotúnel liga a Inglaterra à França. É o segundo túnel ferroviário mais longo do mundo, sendo o primeiro um túnel do Japão.

(32) O Eurotúnel conecta a França ao Japão.

f) lexical

(33) A Comissão de Ética do Parlamento Canadense disse que a ex-Ministra da Imigração, Judy Sgro, nada fez de errado.

(34) A Comissão de Ética do Parlamento Canadense acusa Judy Sgro.

g) conhecimento de mundo

(37) A Microsoft Israel, uma das primeiras filiais da Microsoft fora dos EUA, foi fundada em 1989.

(38) A Microsoft foi criada em 1989.

Como dito acima, os autores consideram a primeira categoria (que compreende a antonímia, a negação e o numérico) como mais fácil de ser detectada pelo computador. A justificativa para isso é o fato de essa detecção automática ser possível sem a compreensão da sentença inteira. Para isso, as duas sentenças devem apresentar palavras razoavelmente similares, especialmente com relação à polaridade delas, para que a contradição seja detectada. Além disso, pouca informação externa é necessária.

A polaridade ocorre quando estão presentes marcadores linguísticos de negação (não), quantificadores de redução (poucos, raros) e preposições restritivas (sem, exceto). Sentenças como *nenhuma bala perfurou* ou *a bala não perfurou* apresentam a mesma polaridade.

Vejamos outros exemplos:

(39) O candidato a governador de estado afirmou que, caso seja eleito, reduzirá os impostos.

(40) O candidato a governador de estado negou que, caso seja eleito, reduzirá os impostos.

Nesses dois exemplos acima, as sentenças apresentam a mesma estrutura, mas utilizam palavras antônimas (afirmou e negou), o que é facilmente detectado pelo computador.

A segunda categoria possui tipos de contradição mais difíceis de serem detectados porque exigem modelos precisos de significação sentencial.

Consideremos o tipo “e” (estrutural, sentenças (31) e (32)), e as sentenças seguintes:

(41) Débora comprou um presente para seu pai na loja de eletrônicos.

(42) Débora adquiriu um bolo.

Os dois pares de sentenças apresentam uma entidade (Débora, Eurotúnel) com uma relação (comprar, ligar) com outras entidades. A segunda sentença de cada par apresenta uma relação similar que inclui uma das entidades envolvidas na relação original, assim como uma entidade que não foi envolvida. No entanto, obtemos resultados diferentes, porque um túnel conecta somente dois lugares únicos, ao passo que mais de uma entidade pode adquirir presentes. Essa estrutura sintática e esse tipo de remissão ao conhecimento de mundo dificultam assegurar que qualquer diferença estrutural é, de fato, uma contradição.

### 3.2. Visão geral do sistema proposto por [DE MARNEFFE 2008]

O sistema proposto por [DE MARNEFFE 2008] é baseado na arquitetura de estágios do sistema Stanford RTE.

No primeiro estágio, há a computação das representações linguísticas, contendo informações sobre o conteúdo semântico dos trechos. Tanto texto (a primeira sentença a ser analisada), como hipótese (a segunda sentença a ser analisada) são convertidos para gráficos de dependência tipificados, produzidos pelo analisador (parser) Stanford.

No segundo estágio, criam-se os gráficos de alinhamento entre texto e hipótese, consistindo em um mapeamento a partir de cada nó na hipótese para um único nó no texto, ou para nada (*null*).

No terceiro estágio, são extraídos os tipos de contradição baseados nas incoerências entre texto e hipótese. Para isso, deve-se primeiro remover os pares de sentenças que não descrevem o mesmo evento e que, portanto, não podem ser contraditórias uma à outra. Vejamos o seguinte par de sentenças:

(43) A lua de Plutão, que possui apenas 42 quilômetros de diâmetro, foi fotografada há 13 anos.

(44) A lua Titã possui um diâmetro de 5.100 quilômetros.

O sistema precisa reconhecer que lua de Plutão não é o mesmo que lua Titã e que, portanto, não podem ser consideradas um par de sentenças a ser analisado como contraditório.

### 3.3. Análise dos resultados do sistema proposto por [DE MARNEFFE 2008]

A detecção de contradição lexical e de conhecimento de mundo apresenta maior dificuldade, pois exigem múltiplas inferências e estão além das capacidades do sistema desenvolvido por [DE MARNEFFE 2008].

Vejamos um par de sentenças:

(45) O Colégio Vitória incluiu recentemente em sua lista de formandos a turma de 2014, que possui 10 alunos.

(46) A lista de formandos do Colégio Vitória possui 10 membros.

O sistema dos autores não consegue inferir que, se a lista de formandos do colégio recebeu recentemente uma turma de 2015 com 10 membros, então essa lista já possuía outros alunos antes.

No entanto, o sistema de [DE MARNEFFE 2008] foi capaz de melhor detectar contradições factivas e modais do que lexicais e de conhecimento de mundo. Intuitivamente, duas sentenças que possuem verbos alinhados com os mesmos sujeitos e diferentes objetos (ou vice-versa) são contraditórias.

## 4. O sistema proposto por [RITTER 2008]

[RITTER 2008] foram motivados em parte pelo trabalho de [DE MARNEFFE 2008], mas elaboraram um estudo com importantes diferenças.

Em primeiro lugar, a simples fundação lógica para a tarefa de detecção de contradição (chamado a partir daqui de DC) sugere que o conhecimento de mundo extensivo é essencial para a construção de um sistema de DC independente de domínio. Eles propõem o uso de conhecimento prévio.

Em segundo lugar, os autores geram um grande corpus de aparentes contradições encontradas em textos arbitrários da internet. Os autores mostram que a maioria destas contradições não são, de fato, contradições, porque são sentenças com meronímias (Alan Turing nasceu em Londres e na Inglaterra), sinônimos (George Bush é casado com ambas Sra. Bush e Laura Bush), hiperônimos (Mozart morreu por problemas de saúde e Mozart comreou por causa de uma doença no sistema urinário) e referência ambígua (um certo John Smith nasceu em 1997, e um outro John Smith nasceu em 1883).

Vejam os dois pares de sentenças apresentadas pelos próprios autores:

(49) Mozart nasceu em Salzburgo.

(50) Mozart nasceu em Viena.

(51) Mozart visitou Salzburgo.

(52) Mozart visitou Viena.

As sentenças (49) e (50) são contraditórias, mas as sentenças (51) e (52) não são. A distinção não é sintática. A relação expressa pelo sintagma *nasceu em* pode ser caracterizada como uma função de nomes de pessoas com seus locais de nascimento. Por outro lado, visitou não denota uma relação funcional.

O sistema de DC precisa reconhecer que:

- a) Mozart refere-se à mesma entidade nas sentenças (49) e (50).
- b) *Nascer em* denota uma relação referencial.
- c) Viena e Salzburgo são lugares inconsistentes.

#### 4.1. Um sistema de DC baseado em funções

Os autores trabalham com o sistema de DC denominado AUCONTRAIRE. Esse sistema trabalha em três etapas:

- 1) Identifica frases funcionais estatisticamente.
- 2) Usa essas frases para criar automaticamente um grande corpus de aparentes contradições.
- 3) Rastreia esse corpus para encontrar verdadeiras contradições usando o conhecimento sobre sinonímia, meronímia, tipos de argumentos e ambiguidade.

Em vez de analisar frases diretamente, AUCONTRAIRE conta com o sistema denominado Text Runner Open Information Extraction (disponível em <http://openie.cs.washington.edu>, acesso em 10 mar. 2015). Esse sistema mapeia cada frase com uma ou mais tuplas que representam as entidades nas sentenças e as relações entre elas, com, por exemplo, nasceu\_em(Mozart, Salzburg).

No entanto, tuplas extraídas são uma aproximação conveniente ao conteúdo da sentença, o que permite enfatizar a detecção de função e a DC baseada em funções.

As contribuições dos autores para a pesquisa de DC são:

a) um novo modelo de tarefa para a detecção de contradição, que oferece um fundamento lógico simples para a tarefa e enfatiza o papel central do conhecimento prévio.

b) a introdução e a avaliação de um novo algoritmo no estilo de EM (*expectation maximization*) para detectar se frases denotam relações funcionais e se substantivos (por exemplo, papai) são ambíguos, o que permite a um sistema de DC identificar as funções em domínios arbitrários.

c) a geração automática de um corpus de aparentes contradições em textos da internet, junto com as experiências sobre este corpus, que fornece uma base para futuros trabalhos de identificação de funções estatísticas e de DC.

Para se acessar o sistema AUCONTRAIRE na internet, faz-se necessária uma senha. Como não obtivemos acesso à senha, consequentemente não obtivemos acesso ao seu conteúdo (disponível em <http://www.cs.washington.edu/research/aucontraire>, acesso em 10 mar. 2015).

#### 4.2. Visão geral do sistema AUCONTRAIRE

AUCONTRAIRE identifica frases denotando relações funcionais e utiliza-as para encontrar afirmações contraditórias em corpora grandes e de domínio público.

Esse sistema começa por encontrar extrações da forma  $R(x, y)$  e identifica um conjunto de relações  $R$  que apresenta uma alta probabilidade de ser funcional. Em seguida, AUCONTRAIRE identifica conjuntos de contradição da forma  $R(x, \cdot)$ .

Os principais componentes de AUCONTRAIRE são: extrator, função de aprendizado, detector de contradição, sinônimos, meronímias, tipos de argumentos e ambiguidade.

Se um par de extrações  $(R(x, y1), R(x, y2))$  não se encontra em nenhuma das categorias acima e  $R$  é funcional, então é provável que as sentenças sejam realmente contraditórias.

#### 4.3. Resultados Experimentais

Foram realizados experimentos para avaliar até que ponto AUCONTRAIRE faz distinção entre contradição genuína e falsos positivos. O sistema trabalhou com um conjunto de dados artificialmente equilibrado construído pelos autores de modo a conter 50% de contradição genuína e 50% de aparentes contradições. Pesquisas anteriores em DC apresentaram resultados a partir de dados selecionados manualmente com um relativo equilíbrio de casos positivos e negativos. Os dados recolhidos a partir da internet são muito assimétricos, contendo apenas 1,2% de contradições genuínas.

De acordo com o levantamento feito pelos próprios autores, o sistema AUCONTRAIRE apresenta queda de desempenho se forem retiradas suas fontes de conhecimento, a saber: conhecimento de sinônimos, conhecimento de meronímias e, especialmente, tipos de argumentos. Por outro lado, se melhorias forem feitas nessas fontes de conhecimento, o desempenho de AUCONTRAIRE melhora.

#### 4.4. Análise de alguns erros do AUCONTRAIRE

[RITTER 2008] rotularam todos os falsos positivos com F-score máximo: 29% para revocação e 48% para precisão. As fontes de erros do AUCONTRAIRE são: ambiguidade (49%), falta de meronímias (34%), falta de sinônimos (14%) e erros de extração (3%). Por esses dados, vemos que argumentos ambíguos (o “x” no par R) dificultam a detecção de contradições genuínas, o que pode ser interpretado, segundo os autores, como uma evidência de que o conhecimento de mundo possui grande importância na tarefa de DC. A falta de meronímias e de sinônimos sugere que recursos lexicais com maior cobertura do que a WordNet e o Tipster Gazetteer melhorariam muito o desempenho do AUCONTRAIRE.

#### 5. Considerações finais

[DE MARNEFFE 2008] apresentaram um estudo sobre como o processamento automático de línguas naturais pode ser utilizado para se detectar contradição em textos em inglês. O foco desses autores está em o sistema detectar se as duas frases analisadas (texto e hipótese) referem-se ao mesmo evento. O sistema, de três etapas, apresentou bons resultados com as contradições que eles classificaram como primeira categoria, mas não apresentou bons resultados com as contradições da segunda categoria, principalmente as do tipo lexical e de conhecimento de mundo.

[RITTER 2008] apresentaram estudo de DC para ser trabalhado com textos também em inglês. A base do sistema consiste em detectar as relações funcionais, e que nos pareceu possuir melhor fundamentação para se detectar contradições. Neste contexto, eles introduziram e avaliaram o AUCONTRAIRE, um sistema com um algoritmo em estilo EM para determinar se uma frase é funcional. Para testar o sistema, os autores criaram um conjunto único de dados com contradições aparentes, com base em sentenças obtidas da internet. Os autores alegam ter tirado duas lições fundamentais: 1) muitas contradições aparentes (aproximadamente 99% nos testes) não são verdadeiras contradições. Assim, a tarefa de DC pode ser muito mais difícil em dados naturais do que nos dados RTE (utilizados por [DE MARNEFFE 2008]); 2) experiência e conhecimento em larga escala são necessários para separar aparentes contradições de genuínas contradições.

Os dois estudos apresentam excelentes bases para se compor um sistema de DC para textos em língua portuguesa.

#### Referências

- DE MARNEFFE, Marie-Catherine; RAFFERTY, A. M. C. (2008). Finding contradictions in text. In HLT, editor, *Proceedings of ACL 2008*, pages 1039–47. HLT.
- MÜLLER, Ana Lúcia de Paula; VIOTTI, E. d. C. (2004). Semântica formal. In FIORIN, J. L. o., editor, *Introdução à Linguística II*. Contexto.
- RITTER, Alan; DOWNEY, D. S. S. E. O. (2008). It's a contradiction - no, it's not: a case study using functional relations. In Center, T., editor, *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. Turing Center.