

Explorando a Geração Automática de Adivinhas em Português

Exploring the Automatic Generation of Riddles in Portuguese

Hugo Gonçalo Oliveira

CISUC, Departamento de Engenharia Informática, Universidade de Coimbra, Portugal

hroliv@dei.uc.pt

Ricardo Rodrigues

CISUC, Instituto Politécnico de Coimbra, Portugal

rmanuel@dei.uc.pt

Resumo

Neste artigo descrevemos um conjunto de experiências realizadas com o objectivo de gerar, de forma automática, adivinhas em português, tendo por base características conhecidas de um conceito. Para além de fazerem sentido, um dos objectivos seria a geração de adivinhas inéditas e, idealmente, com potencial humorístico, nem que por comparação às chamadas “piadas secas”. Parte do desafio passou pela identificação de recursos linguísticos adequados ao nosso objectivo, a que se seguiu a definição de um conjunto de regras para os explorar, tendo em conta um conjunto de hipóteses que poderiam potenciar a originalidade e o valor humorístico. Por fim, as adivinhas foram apresentadas sob a forma de pares pergunta-resposta. O artigo foca-se principalmente num tipo de adivinha, cujo modelo de geração é dissecado e resultados são apresentados a título de exemplo. Uma amostra de adivinhas deste tipo foi classificada manualmente por avaliadores humanos que as consideraram geralmente coerentes, moderadamente originais, mas, na sua maioria, com um baixo potencial humorístico. De forma a mostrar que muito pode ainda ser feito para melhorar os resultados anteriores, nomeadamente ao nível do humor, são também apresentadas, de forma breve, outras experiências num estado ainda incipiente.

Palavras chave

geração de adivinhas, humor computacional, geração de linguagem natural, criatividade linguística, criatividade computacional, relações semânticas

Abstract

This article describes several experiments towards the automatic generation of riddles in Portuguese, based on two known features of given concepts. In addition to making sense, one of our goals was to produce novel riddles, ideally, with some humour potential, at

least when compared to the so called “piadas secas” (dry jokes). Part of the challenge involved the identification of suitable linguistic resources for this goal, followed by setting rules for their exploitation, given some hypothesis that could increase their novelty and humouristic value. Finally, riddles were rendered as question-answer pairs. The article is mainly focused on a kind of riddles, for which the generation model is dissected, some examples are presented, and a sample was manually classified by human judges, who, overall, rated them as coherent, moderately novel, but still with a low humor potential. In order to show how much can still be done, namely about the humor value, the article also presents, briefly, two other experiments, still in an early stage.

Keywords

riddle generation, computational humor, natural language generation, linguistic creativity, computational creativity, semantic relations

1 Introdução

As adivinhas são uma espécie de jogo ou *puzzle* linguístico, para ser resolvido como forma de entretenimento. Georges & Dundes (1963) definem *riddle*, que aqui traduzimos para adivinha, como “*uma forma tradicional de expressão verbal que contém um ou mais elementos descritivos, e onde alguns podem estar em oposição; aquilo a que os elementos se referem deve ser adivinhado*”. As adivinhas estão presentes na maioria das culturas e idiomas, e frequentemente em contextos humorísticos, onde podem ser colocadas como uma pergunta, seguida de uma pausa —de forma a permitir que a audiência pense um pouco— para finalmente revelar a resposta, que pode funcionar como um remate, em inglês, vulgarmente chamado de *punchline*. A resolução de adivinhas requer normalmente criatividade, mas a criativi-



dade também é necessária na produção de novas adivinhas, diferentes daquelas conhecidas por todos, algo fundamental num acto humorístico original.

Em Portugal, é comum chamar-se “piada seca” a um tipo de piada normalmente curta e nem sempre com muita piada, por ser óbvia ou, pelo menos à primeira vista, não fazer muito sentido. Contudo, este tipo de piadas procura tirar partido do anti-clímax para fazer as pessoas rir. Em 2017, as piadas secas, que podem ser apresentadas sob a forma de adivinha, parecem ter regressado à popularidade, com a sua crescente utilização em programas de televisão, vídeos no YouTube¹, e compilação não só em sítios na web, mas também em livros editados (Pinto et al., 2017). Aliás, actualmente, as pessoas tendem mesmo a generalizar e chamar “piada seca” à maioria das piadas do tipo pergunta-reposta curtas e não muito elaboradas.

O presente trabalho inspira-se fortemente noutros trabalhos científicos que visam a geração automática de adivinhas e de humor verbal, para outras línguas (ver secção 2), mas acaba por ser também motivado pelo já mencionado regresso das piadas secas. O trabalho tem como objectivo a produção automática de adivinhas novas, em português, idealmente com valor humorístico, ainda que, tal como o é para a maioria das piadas, especialmente as secas, esse seja um objectivo difícil de avaliar. A tentativa de gerar humor passará pela introdução de incongruência ou ambiguidade na interpretação das perguntas e/ou respostas das nossas adivinhas, que se baseiam sempre num conceito previamente adquirido. Apesar de tudo, neste fase do trabalho, esta tentativa limita-se a um conjunto de hipóteses que considera conceitos com determinadas características, nomeadamente compostos, onde um novo sentido pode ultrapassar a soma do sentido dos seus constituintes, para além de pequenas alterações na ortografia dos anteriores, com impacto no seu som. Tanto quanto sabemos, esta é a primeira vez que um trabalho deste tipo é feito na língua portuguesa.

Apesar de estar nos nossos planos futuros abranger outros tipos de adivinha ou humor verbal, de forma a circunscrever o nosso objectivo, este trabalho começou por explorar alguns recursos linguísticos computacionais para a geração de adivinhas inéditas, tendo por base um pequeno conjunto de modelos de exploração dos recursos e de apresentação das adivinhas. O sistema resultante foi apelidado de SECO.

¹Veja-se, por exemplo, a série *Batalha das Piadas Secas* em <https://www.youtube.com/user/NaoQueresNada>

Este artigo descreve a exploração dos recursos usados, apresenta alguns resultados, e discute a sua validação manual. Mais especificamente, depois desta introdução, a secção 2 faz uma breve revisão de trabalho relacionado, nomeadamente na geração de adivinhas e de humor verbal. De seguida, na secção 3, descrevemos a abordagem actualmente seguida para produzir adivinhas a partir de um conceito inicial, com duas partes —palavras ou sub-sequências—, e duas características desse conceito. Na secção 4 passamos à enumeração e justificação dos recursos linguísticos utilizados, incluindo a origem dos conceitos iniciais, hipóteses e intuições relacionadas com a sua utilização, e ainda a base de conhecimento semântico usada. A secção 5 é dedicada aos modelos escolhidos para apresentar as adivinhas, ilustrados com alguns exemplos. Das adivinhas geradas, uma amostra foi classificada por colaboradores humanos, através de um serviço de *crowdsourcing*. Os resultados dessa classificação são discutidos na secção 6. Estes sugerem que as adivinhas geradas são coerentes ao nível sintáctico e semântico, são moderadamente originais, mas não têm grande potencial humorístico. Apesar de o humor ser um aspecto subjectivo, isto também mostra que, especialmente quando pensamos em “piadas secas”, de forma a acrescentar valor humorístico, é necessário trabalhar este aspecto de forma mais direccionada, e não nos basearmos apenas na escolha dos conceitos, como aconteceu. Para além de analisar as pontuações globais, as adivinhas mais bem pontuadas são reveladas, e os aspectos pontuados são ainda mostrados de acordo com a origem dos conceitos, apresentação, tipo de características usadas, e ainda país origem dos avaliadores. Antes de concluir, na secção 7 apresentamos alguns resultados gerados através de métodos semelhantes aos anteriores, onde procuramos aumentar o potencial humorístico. Estes resultados são ainda muito incipientes e com eles pretendemos mostrar essencialmente que, nesta linha, ainda muito pode ser feito.

2 Trabalho Relacionado

O foco nas adivinhas como tema de investigação não é propriamente recente. Veja-se, por exemplo, o trabalho de Georges & Dundes (1963), que procura precisamente definir o conceito de adivinha (*riddle*). Sobre os fenómenos importantes na criação de uma adivinha, Palma & Weiner (1992) referem que a ambiguidade lexical, resultante, por exemplo, de polissemia ou homofonia, é tão ou mais importante para a criação de adivi-

nhas do que a associação de palavras a determinadas categorias, de acordo com os seus múltiplos sentidos.

Mesmo a geração de adivinhas por programas de computador já é estudada desde a década de 1990, com o trabalho seminal de [Binsted & Ritchie \(1994\)](#), de onde resultou o JAPE, um sistema que gera trocadilhos sob a forma de adivinhas, com base em: um léxico com informação sintática e semântica acerca das palavras e seus possíveis sentidos; um conjunto de esquemas para combinar duas palavras baseadas na sua relação lexical ou fonética; e um conjunto de modelos para apresentar as adivinhas. Para a geração de trocadilhos, a versão inicial do JAPE seguia uma de três estratégias: (i) substituição de sílabas por outras com sons próximos; (ii) substituição de palavras por outras com uma sonoridade próxima; (iii) meta-tese, onde a inversão de sons e palavras sugere uma similaridade de sentido em frases que, de outra forma, seriam entendidas com um significado diferente. Um das adivinhas dadas como exemplo é “*What do you call a murderer that has fibre? A cereal killer*”, onde o sistema pode tirar partido das semelhanças entre as palavras *serial* e *cereal*. O STANDUP ([Manurung et al., 2008](#)) é um outro sistema que adopta uma abordagem semelhante ao JAPE, mas tem mais cuidado na escolha de palavras, restringidas a vários níveis para uma melhor apresentação e utilização por determinadas audiências, tais como crianças. Para além do inglês, a geração de adivinhas com piada foi também tentada para o japonês ([Sjöbergh & Araki, 2007](#)).

Trabalhos mais recentes na geração de adivinhas incluem o TheRiddlerBot ([Guerrero et al., 2015](#)), que gera adivinhas acerca de personagens famosas. Depois de seleccionar o nome de uma personagem conhecida de uma base de conhecimento, a geração passa pelas seguintes fases: (i) recuperação de características associadas à personagem; (ii) identificação de personagens análogas, por terem características em comum; (iii) selecção de um modelo de apresentação, com base em algumas das características ou na analogia; (iv) publicação da adivinha numa conta da rede social Twitter; (v) recuperação de nomes alternativos, a partir da Wikipédia. Depois de publicada a adivinha, os utilizadores do Twitter podem tentar responder com o nome da personagem ou um dos seus nomes alternativos. Um exemplo de adivinha para o *Joker*, personagem do filme *Batman*, seria: “*Tell me the name of a person that is the Morpheus of The Dark Knight Rises, is criminal, playful yet cruel, has been seen wearing a purple topcoat.*”

[Galvan et al. \(2016\)](#) também trabalharam na geração de adivinhas, com base na associação de palavras. Dado um conceito, a abordagem proposta passa pelas seguintes fases: (i) recuperação das suas possíveis categorias num tesouro criativo; (ii) recuperação de modificadores associados, e selecção aleatória de um deles; (iii) recuperação de novas categorias, a que o modificador seleccionado também esteja associado; (iv) composição de uma categoria final através da combinação do modificador com uma das novas categorias; (v) utilização de um conceito da categoria final para preencher um modelo textual. Um exemplo de adivinha com resposta “sol” (*sun*) seria: “*What is as hot as soup?*”

Ao contrário do JAPE e do STANDUP, nos dois trabalhos anteriores não há uma preocupação especial com o aspecto do humor. Já um outro trabalho que procura gerar humor verbal ([Labutov & Lipson, 2012](#)), incluindo adivinhas, explora a estrutura da rede semântica do ConceptNet. Neste caso, a produção de textos humorísticos tem por base a exploração de caminhos de relações e, para maximizar a incongruência, procura-se alinhar caminhos entre dois conceitos, diferentes mas com alguma sobreposição. Mais precisamente, para adivinhas apresentadas como pergunta-resposta, a pergunta menciona dois conceitos de caminhos diferentes, mas do mesmo domínio (identificados através de *clustering*), enquanto que o conceito na resposta está em um dos caminhos mas pertence a um domínio diferente. Um exemplo de uma adivinha gerada será “*Why is the computer in hospital? Because the computer has virus.*”

Para além de adivinhas, existem trabalhos com vista à geração de outros tipos de humor verbal, incluindo acrónimos com piada ([Stock & Strapparava, 2006](#)) —e.g., FBI: *Fantastic Bureau of Intimidation*—, ou mensagens curtas ([Valitutti et al., 2013](#)) —e.g., *I’ve sent you my fart.. I mean ‘part’ not ‘fart’...*. Ambos exploram a substituição de palavras para potenciar o humor. As palavras seleccionadas para substituir outras devem começar pela mesma letra ou ter um som próximo e respeitar outras características, como por exemplo serem linguagem tabu (e.g., calão).

Tanto na modalidade verbal como oral, o humor baseia-se geralmente em quatro fenómenos linguísticos ([Tagnin, 2005](#)): (i) homonímia — palavras com a mesma escrita e som (e.g., ‘banco’, de jardim ou instituição financeira); (ii) homofonia — palavras com o mesmo som, mas diferentes grafias (e.g., ‘cesta’ e ‘sexta’); (iii) polissemia — palavras com a mesma grafia e som, mas múltiplos sentidos relacionados (e.g., ‘banco’, ins-

tuição financeira ou de dados); (iv) paronímia — palavras com grafia e sons próximos (e.g., ‘tráfego’ e ‘tráfico’). Assim, para uma tentativa de humor funcionar, é importante que a audiência seja nativa ou fluente na língua em que uma piada é colocada.

De acordo com Attardo (2008), os textos humorísticos podem classificar-se em três tipos: (i) enredo cómico com uma *punchline*, que é o caso das piadas típicas; (ii) enredo cómico com uma ruptura na narrativa que introduz referências humorísticas acerca do autor ou de que a audiência tenha conhecimento; (iii) enredo cómico com uma complicação central humorística, isto é, uma situação banal onde alguns elementos provocam o riso devido às suas consequências.

Na maior parte dos trabalhos anteriores, as adivinhas são geradas com base em alguma teoria ou através da exploração de bases de conhecimento, que são recursos computacionais essenciais para este fim. Apesar de várias semelhanças com o nosso trabalho, salvo uma excepção para o japonês, todos eles produzem texto em inglês. No nosso caso, o objectivo é gerar adivinhas em português, língua para a qual, tanto quanto sabemos, não existe nenhum trabalho deste tipo. Há, contudo, trabalho anterior, também da nossa autoria, na geração automática de um tipo de humor característico da Internet (Gonçalo Oliveira et al., 2016), baseado numa imagem macro e uma linha de texto, em português, mas que não envolve adivinhas.

3 Abordagem

O principal objectivo do SECO é explorar recursos de conhecimento e linguísticos, em português, para gerar novas adivinhas, com a finalidade de entreter uma audiência, e que poderão, em alguns casos, ter valor do ponto de vista humorístico, devido a um trocadilho inerente. Especificamente, o trabalho descrito neste artigo foca-se em adivinhas baseadas em um conceito e duas características dele, apresentadas como pares pergunta-resposta. Se considerarmos os artefactos produzidos pelo sistema como piadas, de acordo com a classificação de Attardo (2008), o SECO produz humor do primeiro tipo, ou seja, apresenta um enredo (pergunta) e depois remata (resposta). Contudo, estamos ainda a dar os primeiros passos em direcção ao nosso objectivo final e, conforme iremos mostrar mais à frente, o potencial humorístico das adivinhas geradas actualmente ainda é baixo.

De certa forma, a abordagem actual do SECO é fortemente inspirada em trabalhos anteriores que procuram gerar adivinhas. A principal diferença será mesmo a geração de texto em português, o que acaba por implicar a utilização de outros recursos linguísticos (ver secção 4). A principal inspiração será o sistema JAPE (Binsted & Ritchie, 1994), apresentado na secção 2, nomeadamente algumas das adivinhas analisadas pelos seus autores e geradas por este sistema, tais como:

- *What do you get if you cross a zebra with a kangaroo? A striped jumper.*
- *What do you get if you cross a car with a vile substance? Crude oil.*
- *What do you get when you cross a chicken and a power pack? A battery hen.*

Cada uma das anteriores pergunta o resultado do cruzamento entre dois conceitos iniciais e a resposta é um outro conceito, lexicalizado por duas palavras. A principal nuance é que cada uma dessas palavras (características) tem um significado relacionado com um dos conceitos iniciais. Em português também é possível encontrar adivinhas que, de certa forma, se encaixam no modelo anterior, tais como:

- *Qual o resultado do cruzamento de uma galinha com uma cobra? Um pinto longo.*
- *Que resulta do cruzamento entre uma cobra e um ouriço? Arame farpado.*
- *Que resulta do cruzamento entre uma girafa e um papagaio? Um alto-falante.*

De forma a gerar adivinhas deste tipo, o procedimento actual segue uma abordagem de baixo para cima, também ele com algumas semelhanças ao procedimento de geração adoptado pelo JAPE. Mais propriamente, partindo de um conceito inicial, passa-se pelas seguintes fases, até se apresentar a adivinha sob a forma de pergunta-resposta:

1. Extracção de características;
2. Emparelhamento de características;
3. Pontuação e filtragem de características;
4. Apresentação como pergunta-resposta.

No nosso caso, o conceito inicial pode ser uma expressão com duas palavras ou uma palavra única que se pode dividir artificialmente em duas palavras válidas, isto é, que terá duas partes (cw_1 e cw_2), uma para cada palavra. Na primeira fase, cada parte do conceito é considerada individualmente, e as características são recuperadas

a partir de uma base de conhecimento, algumas envolvendo a primeira parte (T_1) e outras a segunda (T_2). As características são representadas através de palavras (fw_1 e fw_2) e obtidas a partir de triplos do tipo *a relacionadoCom b*, onde *a* e *b* são palavras e *relacionadoCom* é o nome de uma relação semântica entre significados de *a* e *b* (e.g., *animal hiperónimoDe cão*). Assim, interessam-nos os triplos $t_1 \in T_1$ e $t_2 \in T_2$ que envolvem, respectivamente, cw_1 e cw_2 :

- $T_1 : \forall(x, relacionadoCom, y) \in T_1$
 $\rightarrow (x = cw_1 \wedge y = fw_1) \vee (x = fw_1 \wedge y = cw_1)$
- $T_2 : \forall(x, relacionadoCom, y) \in T_2$
 $\rightarrow (x = cw_2 \wedge y = fw_2) \vee (x = fw_2 \wedge y = cw_2)$

Na segunda fase, as características $fw_1 \in T_1$ são emparelhadas com as características $fw_2 \in T_2$. Juntamente com o conceito inicial, cada par $\{fw_1, fw_2\}$ é suficiente para produzir uma adivinha. A figura 1 serve para ilustrar as duas primeiras fases. O conceito tem duas partes, cada uma relacionada com uma característica, sob a forma de uma palavra (fw_1 e fw_2), recuperada da base de conhecimento.

Na terceira fase, as adivinhas são pontuadas automaticamente, onde se procura considerar a comunalidade e representatividade das suas características. Finalmente, na quarta fase, as adivinhas são apresentadas num formato textual, mais propriamente como uma pergunta e uma resposta, de acordo com um modelo pré-definido, onde cada característica é inserida.

4 Recursos Linguísticos

Um dos desafios deste trabalho foi a identificação de recursos computacionais linguísticos que nos pudessem ajudar na geração automática de adivinhas em português, idealmente com algum valor humorístico, tais como as apresentadas na secção 3. No desenvolvimento da versão actual do SECO, foram identificados e explorados recursos para obter os conceitos iniciais (listas), recuperar características (base de conhecimento semântica), pontuar as adivinhas (corpo), e tratar das flexões em diferentes géneros e números para a apresentação da adivinha (léxico morfológico). As secções seguintes descrevem os recursos usados, juntamente com exemplos do seu conteúdo, e motivam a sua seleção.

4.1 Conceitos Iniciais

A geração de adivinhas tem por base um conceito inicial, e a criação de um conjunto de adivinhas deve ser apoiado por uma lista de conceitos

que sigam as especificações desejadas. Tendo em conta que o procedimento seguido tira partido de conceitos que se podem dividir em duas partes, expressões com duas palavras pareceram-nos adequadas para este fim. Sendo assim, começamos por explorar os seguintes recursos:

- **Compostos:** uma lista de 180 expressões compostas em português (Ramisch et al., 2016), com instâncias tais como *água doce*, *mau-humor*, ou *primeira mão*.
- **N-Adj:** todos os 289 pares substantivo-adjectivo que ocorrem pelo menos 750 vezes no corpo jornalístico CETEMPúblico (Rocha & Santos, 2000), com instâncias tais como *comunicação social*, *ensino superior* ou *prisão preventiva*.

Os conceitos de ambas as listas encontram-se lexicalizados em duas palavras: um substantivo modificado por um adjectivo. A nossa intuição é que o significado dos compostos disponíveis como tal (Compostos) será mais do que a simples soma dos significados de ambas as palavras, enquanto que a maioria dos pares nome-adjectivo extraídos do corpo terão um sentido mais literal. Mais propriamente, a nossa hipótese é que, se as características estiverem associadas aos significados mais literais, os conceitos da primeira lista irão, potencialmente, resultar em associações mais surpreendentes, podendo mesmo sugerir alguma incongruência e, por isso, serem mais susceptíveis de produzir humor. É certo que, a este respeito, poderíamos ter ido mais longe e, por exemplo, aplicar uma medida para a não-composicionalidade de expressões multi-palavra (ver, e.g., Biemann & Giesbrecht (2011)). No entanto, optamos por não o fazer porque, nesta fase, o foco principal do trabalho foi mesmo averiguar o que conseguiríamos obter com uma abordagem simplista, e de que forma os resultados obtidos com os dois tipos de lista se equiparavam.

Tendo em conta que o humor pode resultar de duas palavras com sons próximos (homofonia, paronímia), criámos mais duas listas a partir das anteriores, Compostos-d1 e N-Adj-d1, com expressões em que ambas as palavras são válidas, mas em que uma tem um distância de edição igual 1 (remoção, adição ou substituição), respectivamente para uma expressão nas listas anteriores. Algumas das instâncias geradas incluem:

- *amido oculto*, obtido a partir de *amigo oculto*;
- *véu aberto*, obtido a partir de *céu aberto*;

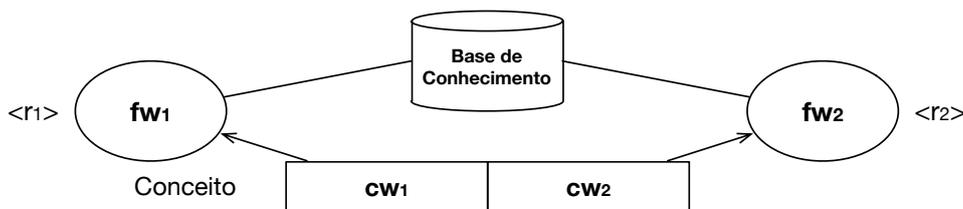


Figura 1: Procedimento para a geração de adivinhas com base em duas características de um conceito, obtidas através de uma base de conhecimento.

- *primeiro pano*, obtido a partir de *primeiro plano*.

Aqui, a nossa hipótese foi que, ainda que a interpretação possa ser literal e/ou o composto não seja comum, o paralelismo com um composto original e bem conhecido pode dar origem a algo novo e com algum potencial humorístico.

Uma última lista de conceitos ($W1+W2$) contém cerca de 900 mil palavras, ou formas, que se encontram no léxico LABEL-Lex (Ranchhod et al., 1999) e que podem ser interpretados como amálgamas. Referimo-nos a conceitos do tipo w_1w_2 , onde w_1 e w_2 são sequências de caracteres válidas, por também estarem presentes no léxico. Instâncias deste tipo incluem, por exemplo:

- *malabar* = *mala* + *bar*
- *centralidade* = *central* + *idade*
- *restolho* = *resto* + *olho*

A nossa hipótese aqui é que um significado inesperado e, de certa forma, criativo, poderia ser atribuído à palavra, mais uma vez percebido como incongruência e ampliando o potencial humorístico.

4.2 Características Consideradas

As características são recuperadas a partir de uma base de conhecimento léxico-semântico obtida a partir de dez recursos lexicais para o português, incluindo redes extraídas de dicionários, wordnets e da ConceptNet (Gonçalo Oliveira, 2018). Mais concretamente, neste trabalho foram usados 45.510 triplos relacionais (*a relacionadoCom b*), correspondentes àqueles que se encontram em pelo menos três dos dez recursos². Cada triplo liga duas palavras, de acordo com os seus significados, mas diferentes sentidos da mesma palavra não são identificados de forma explícita. Ainda que de forma

²A partir de http://ontopt.dei.uc.pt/index.php?sec=download_outros, Triplos relacionais (10 recursos), é possível obter uma lista com todos os triplos nos dez recursos, seguidos do número de recursos em que ocorrem (por exemplo, *fruto HIPERONIMO DE tomate* 3).

não intencional, isto permite a utilização de significados mais inesperados. A nossa expectativa é que a probabilidade de seleccionar exactamente o mesmo significado que uma palavra tem num termo composto seja baixa, especialmente em compostos onde o significado não resulta da soma dos significados das suas palavras.

Como as relações usadas estão presentes em pelo menos três recursos, elas serão, geralmente, conhecidas e consensuais. Ao levantar essa restrição, seria possível obter mais características, mas não tão imediatas para o público em geral. Haveria também um maior risco de utilizar características incorrectas, porque a maior parte dos dez recursos base foi criada de forma automática ou semi-automática, com um tratamento manual inexistente ou limitado.

Entre as relações cobertas pela base de conhecimento, identificámos um subconjunto de tipos de relação que poderiam ser usados para obter características, dada a sua frequência e adequação à nossa tarefa, considerando também que a maior parte dos conceitos das nossas listas são lexicalizados através de um substantivo e um adjetivo. A tabela 1 mostra os tipos de relação usados para as características, juntamente com a sua frequência na base de conhecimento e com o texto usado para as apresentar nas adivinhas³.

4.3 Pontuação de Adivinhas

Recorrendo somente ao procedimento anterior, seria possível seleccionar palavras de utilização pouco comum, não muito conhecidas pelo público em geral, ou mesmo características pouco representativas das partes alvo do conceito. Para minimizar estas situações, cada par de características foi pontuado, positivamente e, em alguns casos, negativamente, de acordo com a equação 1, justificada de seguida:

³Na verdade, no ficheiro disponível, algumas das relações têm um nome diferente, todo em maiúsculas, também com a categoria gramatical esperada para os argumentos identificados, mas optamos por representar os nomes de relações de uma forma em que essa identificação fosse mais imediata.

#	Relação	arg ₁	arg ₂	Apresenação (r _i)
7,538	adj-sinónimoDe	fw/cw	cw/fw	<i>o que é fw</i>
353	adj-antónimoDe	fw/cw	cw/fw	<i>o que não é/está fw</i>
4,035	hiperonimoDe	fw	cw	<i>fw</i>
590	adj-dizSeSobre-n	cw	fw	<i>fw</i>
	adj-dizSeSobre-n	fw	cw	<i>o que é fw</i>
100	n-parteDe-adj	cw	fw	<i>o que é fw</i>
		fw	cw	<i>fw</i>
58	n-parteDe-n	cw	fw	<i>uma parte de fw</i>
		fw	cw	<i>o que tem fw</i>
49	n-membroDe-n	cw	fw	<i>um membro de fw</i>
		fw	cw	<i>o que tem fw</i>
46	adj-temQualidade-n	cw	fw	<i>o que tem fw</i>
		fw	cw	<i>o que é fw</i>
45	n-fazSeCom-n	cw	fw	<i>a finalidade de fw</i>
		fw	cw	<i>o que serve para fw</i>
110	v-finalidadeDe-n	cw	fw	<i>a finalidade de fw</i>
		fw	cw	<i>o que serve para fw</i>
1,572	v-causador-n	cw	fw	<i>fw</i>
		fw	cw	<i>o efeito de fw</i>
101	n-localDe-n	cw	fw	<i>o que tem fw</i>
		fw	cw	<i>o que vem de fw</i>

Tabela 1: Características extraídas e sua apresentação textual.

- Com o objectivo de favorecer palavras conhecidas pelo público em geral, a pontuação de cada par é proporcional à frequência de cada palavra que representa a característica (fw_1 e fw_2) no corpo CETEMPúblico (α , na equação 2).

$$\alpha = \frac{\log(freq(w_1)) + \log(freq(w_2))}{\log(\#maxFreq)} \quad (2)$$

- Quando uma das características é uma relação de hiperonímia, o par é penalizado de acordo com o número de hipónimos da característica (β , na equação 3).

$$\beta = \frac{\log(nHiponimos(w_1))}{\log(\#maxHiponimos)} \quad (3)$$

A última opção é tomada porque, mesmo com a restrição de usar triplos em pelo menos três recursos, algumas palavras têm muitos hipónimos (e.g., pessoa, 383; planta, 115; instrumento, 94), o que faz aumentar também o número de respostas possíveis e diminuir a resolubilidade da adivinha. Isto acontece principalmente em redes semânticas que não têm uma taxonomia equilibrada e, por isso, algumas relações de hiperonímia ligam directamente conceitos de uma ordem elevada a conceitos muito específicos (e.g., em vez de *ave hiperonimoDe animal aquático hiperonimoDe animal marinho hiperonimoDe pinguim*, existe uma ligação directa *ave hiperonimoDe pinguim*).

$$Pontos(par) = \alpha - \beta \quad (1)$$

5 Apresentação de Adivinhas

As adivinhas são apresentadas como pares pergunta-resposta, baseados em um conceito e um par de características extraídas a partir desse conceito, fw_1 e fw_2 , por sua vez apresentadas como r_1 e r_2 , respectivamente. Aqui, a principal diferença entre fw_1 e fw_2 e r_1 e r_2 é que as últimas podem incluir um artigo antes das características. Esta secção mostra os modelos de apresentação actualmente utilizados, com exemplos para cada um.

A forma de apresentação base para uma adivinha usa o seguinte modelo, motivado pela sua utilização recorrente em adivinhas e trocadilhos:

- *Que resulta do cruzamento entre $\langle r_1 \rangle$ e $\langle r_2 \rangle$? $\langle c \rangle$.*

Considere-se, por exemplo, o conceito *direitos humanos*, com o seguinte conjunto de características extraídas:

- *direito* sinónimoDe *liso*
- *direito* sinónimoDe *plano*
- *humano* dizSeSobre *homem*

Ou as seguintes, para o conceito *diapositivo*, dividido em duas partes — *dia* + *positivo*:

- *hora* parteDe *dia*
- *momento* hiperónimoDe *dia*
- *positivo* sinónimoDe *real*
- *positivo* sinónimoDe *confiante*

Os conceitos e características anteriores permitem gerar as seguintes adivinhas:

- *Que resulta do cruzamento entre o que é liso e um homem?*
direitos humanos.
- *Que resulta do cruzamento entre o que é plano e um homem?*
direitos humanos.
- *Que resulta do cruzamento entre o que tem uma hora e o que é real?*
diapositivo.
- *Que resulta do cruzamento entre um momento e o que é real?*
diapositivo.

O procedimento de geração para a produção da primeira adivinha da lista é ilustrado na figura 2.

Sempre que necessário, recorremos ao LABEL-Lex para identificar o género das características e seleccionar o artigo mais adequado. No seguinte exemplo, é usado o artigo feminino *uma* para a primeira característica e o artigo masculino *um* para a segunda:

- *Que resulta do cruzamento entre uma linguagem e um militar?* *língua oficial.*

Embora não nos pareça tão crítico, no futuro, o mesmo léxico poderá ser usado para tratar também o número.

Para além da apresentação base, identificámos uma situação específica onde se justifica uma pergunta diferente. Mais propriamente, quando uma das características é uma relação de antonímia, o seguinte modelo é utilizado:

- *Qual é o contrário de $\langle r_1 \rangle$ $\langle r_2 \rangle$? $\langle c \rangle$.*

Este modelo funciona para todos os conceitos compostos por um substantivo e um adjectivo, que é o caso da maioria dos conceitos usados.

Por exemplo, para *direitos humanos* e *diapositivo*, as seguintes características também foram extraídas:

- *direito* antónimoDe *esquerdo*
- *direito* antónimoDe *torto*
- *humano* antónimoDe *desumano*
- *positivo* antónimoDe *negativo*

Essas características são usadas para gerar as seguintes adivinhas, em que, por ser mais natural, na pergunta, o antónimo é sempre colocado em segundo lugar:

- *Qual é o contrário de homem esquerdo?*
direitos humanos.
- *Qual é o contrário de homem torto?*
direitos humanos.
- *Qual é o contrário de plano desumano?*
direitos humanos.
- *Qual é o contrário de hora negativa?*
diapositivo.

O procedimento de geração é ilustrado na figura 3 para a última adivinha da lista.

Para além dos modelos anteriores, durante as experiências realizadas, questionámo-nos se as adivinhas funcionariam melhor quando apresentadas tal como descrito — uma pergunta que menciona as duas características e uma resposta que é um conceito potencialmente conhecido — ou se seria preferível inverter a apresentação — uma pergunta simples acerca do significado do conceito inicial, e uma resposta que o explica com recurso às características. De forma a explorar a segunda opção, todas as adivinhas foram também apresentadas através do seguinte modelo:

- *O que significa $\langle c \rangle$? $\langle r_1 \rangle$ e $\langle r_2 \rangle$.*

Aplicando este modelo nos conceitos *direitos humanos* e *diapositivo*, com as características anteriores, é possível obter as seguintes adivinhas:

- *O que significa direitos humanos?*
O que é plano e um homem.
- *O que significa direitos humanos?*
O que é liso e um homem.
- *O que significa direitos humanos?*
Um homem que não é/está esquerdo.
- *O que significa direitos humanos?*
Um homem que não é/está torto.
- *O que significa direitos humanos?*
Um plano que não é/está desumano.

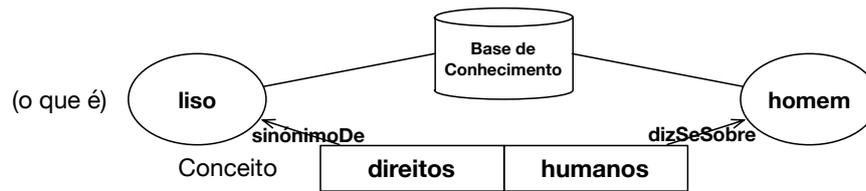


Figura 2: Instanciação do procedimento de geração para o conceito *direitos humanos*, usando o modelo base.

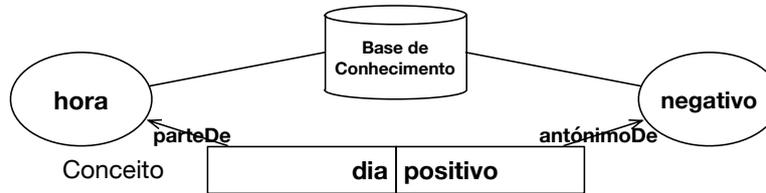


Figura 3: Instanciação do procedimento de geração para o conceito *diapositivo*, usando o modelo para antónimos.

- *O que significa diapositivo?*
O que tem hora e é/está real.
- *O que significa diapositivo?*
Um momento que é/está real.
- *O que significa diapositivo?*
Uma hora que não é/está negativa.

Em busca de algumas conclusões relativamente à apresentação preferível, decidimos gerar as adivinhas seguindo ambos os modelos (Características→Conceito ou Contrário, e Conceito→Características) e, mais à frente, olhar para os resultados de uma avaliação manual independente.

6 Validação

De forma a obter uma opinião próxima daquela do público em geral, uma amostra de adivinhas foi gerada e enviada para a plataforma de crowdsourcing Figure Eight⁴ (antigo Crowdfunder), onde foi criada uma tarefa (job) a que se deu o nome de *Adivinhas em Português*. Esta tarefa pedia a colaboradores humanos, de Portugal ou Brasil, que classificassem adivinhas de acordo com três aspectos: coerência, novidade e potencial humorístico. Por cada conjunto de respostas completo, cada colaborador teria direito a uma pequena recompensa monetária e, numa tentativa de generalizar as respostas, definiu-se que cada colaborador não poderia classificar mais de 15 adivinhas.

Nesta secção descrevemos os aspectos classificados por cada colaborador, explicamos como a

amostra de avaliação utilizada foi gerada, e apresentamos os resultados globais. No entanto, para além da opinião global, esta tarefa foi também desenhada para nos fornecer informação que pudesse ser útil em escolhas a fazer no futuro, com vista a melhorias do sistema. Por isso, também apresentamos e discutimos os resultados organizados de acordo com quatro parâmetros: origem dos conceitos, apresentação das adivinhas, características usadas, e ainda origem dos avaliadores.

6.1 Aspectos classificados

Com o objectivo de manter a tarefa curta, foram colocadas apenas três questões por adivinha, cada uma focada num aspecto diferente, e respondida através de uma escala de Likert de 5 pontos. Com a primeira pergunta, pretendíamos saber se o texto gerado era sintacticamente e semanticamente coerente, e podia ser considerado como uma adivinha. Pedíamos inclusivamente aos colaboradores para, em caso de dúvida, recorrerem a um dicionário. A escala da resposta, explicada aos colaboradores, tinha os seguintes extremos:

- Incoerente (1): a pergunta está mal estruturada, é de difícil interpretação, e a resposta não tem qualquer relação (nem mesmo com um sentido pouco óbvio e depois de pensar um pouco).
- Perfeitamente coerente (5): a pergunta está bem estruturada, tem uma interpretação clara, e a resposta responde efectivamente à pergunta (mesmo que com um sentido menos óbvio e/ou que obrigue a pensar um pouco).

⁴<https://www.figure-eight.com/>

De seguida, pretendíamos perceber até que ponto a adivinha era nova, surpreendente, e tinha, por isso, valor criativo, o que já envolve algum nível de subjectividade. A pergunta pedia para classificar a originalidade da adivinha numa escala com os seguintes extremos:

- Nada original (1): na sua opinião, a adivinha é demasiado óbvia e nada inovadora.
- Muito original (5): na sua opinião, a adivinha é autêntica / surpreendente e demonstra criatividade.

Por fim, pretendíamos apurar se a adivinha tinha a capacidade de fazer rir e poderia ser usada num contexto humorístico. A pergunta colocada pedia para a classificar precisamente o seu potencial humorístico, numa escala com os seguintes extremos:

- Nenhum potencial (1): na sua opinião, será impossível fazer alguém rir com esta adivinha, por exemplo, por não fazer qualquer sentido ou por ter uma interpretação completamente literal.
- Muita piada (5): na sua opinião, a adivinha tem um grande potencial humorístico e dá vontade de rir.

A classificação por parte dos colaboradores baseou-se apenas no senso comum. Ou seja, para além da descrição das escalas a usar, não foi realizado qualquer tipo de treino nem apresentados exemplos do que seriam adivinhas coerentes, originais e com piada. Ainda que fosse possível escolher um conjunto de adivinhas coerentes e bem formadas, por ser o aspecto menos subjectivo, a novidade e, principalmente, o potencial humorístico têm um elevado nível de subjectividade, para além de que a apresentação de bons exemplos de adivinhas poderia condicionar a classificação da originalidade.

6.2 Geração de amostras

Inicialmente foram geradas adivinhas para cada lista de conceitos referida na secção 4.1, depois ordenadas de acordo com a sua pontuação automática (da maior para a menor). Das anteriores, apenas se mantiveram as 300 adivinhas com maior pontuação para cada lista de conceitos. Garantimos ainda que não havia mais de três adivinhas por conceito inicial. Depois, cada adivinha das anteriores foi apresentada com o modelo Conceito→Características/Contrário e também Características→Conceito, duplicando

assim o número de adivinhas usadas. Das anteriores, foi feita uma selecção aleatória de 320 adivinhas, enviada para o Figure Eight, onde cada uma foi classificada por três colaboradores diferentes.

6.3 Resultados globais

A tabela 2 apresenta os resultados globais da validação humana de adivinhas. Para cada aspecto classificado, é mostrado o número de adivinhas para cada resposta possível (entre 1 e 5), a moda, a mediana, e a concordância entre colaboradores. A concordância fica apenas ligeiramente acima dos 50%, o que confirma a subjectividade envolvida. Globalmente, esta validação mostra que a maioria das adivinhas geradas são sintacticamente e semanticamente coerentes — mais de metade foram classificadas com 4 ou 5 —, e pode-se dizer que há uma proporção interessante de adivinhas originais. Por outro lado, os colaboradores não acharam muita piada às adivinhas e consideraram que tinham um baixo potencial humorístico, ao classificarem quase dois terços das adivinhas com 1 ou 2 neste aspecto.

Class.	Coerência	Original.	Humor
1	137	206	444
2	137	231	216
3	216	270	144
4	266	158	86
5	204	95	70
Moda (Mo)	4	3	1
Mediana (Md)	3	3	2
Concord.	51%	59%	58%

Tabela 2: Resultados globais da validação humana.

A correlação de Pearson entre cada aspecto e a pontuação automática de cada adivinha também foi calculada mas estava, para cada aspecto, próxima de zero. Isto pode significar que essa pontuação não reflecte nem a coerência, nem a originalidade, nem o potencial humorístico das adivinhas. Pode também significar que, ao termos criado uma amostra a partir das adivinhas mais bem pontuadas, invalidámos a possibilidade de tirar conclusões acerca desta pontuação. Ou seja, ainda que seja nossa intenção melhorar esta pontuação, terá de ser algo a estudar com mais cuidado, no futuro. A tabela 3 mostra as adivinhas com melhores classificações ao nível da originalidade e do humor.

<i>Original</i>	<i>Humor</i>	Adivinha	Pontuação
4.67	4.67	<i>Que resulta do cruzamento entre a finalidade de lixa e um treinador? politécnico.</i>	0.56
4.67	4.00	<i>Que resulta do cruzamento entre um julgamento e uma sociedade? justiça social.</i>	0.61
4.33	4.33	<i>Que resulta do cruzamento entre o que é plano e um homem? direitos humanos.</i>	0.74
4.33	4.00	<i>O que significa lei orgânica? uma norma que é um ser.</i>	0.69
4.33	3.67	<i>O que significa porto forte? um vinho que é violento.</i>	0.62
4.33	3.67	<i>O que significa novo mudo? o que é calado e não é/está antigo.</i>	0.62
4.33	3.67	<i>Que resulta do cruzamento entre uma área e o que é marginal? zona ribeirinha.</i>	0.59
4.33	3.33	<i>O que significa junção pública? o efeito de juntar que não é/está privado.</i>	0.68
4.33	3.00	<i>Que resulta do cruzamento entre o efeito de discutir e o que é manifesto? discussão pública.</i>	0.62
4.33	2.67	<i>Que resulta do cruzamento entre o que é chato e uma parte de minuto? segundo plano.</i>	0.57
4.33	1.00	<i>O que significa crise económica? o que é crítico e é uma economia.</i>	0.67
3.00	4.67	<i>O que significa quadro-negro? uma obra que não é/está branca.</i>	0.60
4.00	4.00	<i>Que resulta do cruzamento entre o que é calado e o que é comum? mudo geral.</i>	0.59
4.00	4.00	<i>O que significa pronto forte? o que é rápido e é uma força.</i>	0.71
4.00	4.00	<i>Que resulta do cruzamento entre o que é mágico e o que é preto? magia negra.</i>	0.56
3.67	4.00	<i>O que significa solícito? uma estrela que não é/está ilegal.</i>	0.54
3.67	4.00	<i>O que significa politécnico? a finalidade de lixa que é um treinador.</i>	0.56
3.67	3.67	<i>Qual é o contrário de pobre velho? novo-rico.</i>	0.66

Tabela 3: Adivinhas com maior classificação humana na originalidade e humor.

6.4 Resultados parciais

Para cada lista de conceitos iniciais, a tabela 4 mostra o número de adivinhas classificadas juntamente com a moda (Mo) e mediana (Md) dos aspectos avaliados. Verifica-se que a coerência é alta para os termos compostos originais, seguida da lista N-Adj, mas baixa para os novos compostos criados e para as palavras onde a divisão é essencial para forçar um novo significado. A originalidade é comparável para todas as listas, com moda e mediana sempre 3, excepto para a lista N-Adj, com moda 2, o que segue a nossa intuição inicial: o significado dos compostos extraídos do corpo é demasiado literal quando comparado com os termos compostos disponíveis na lista, onde o significado é mais do que uma mera soma dos significados das palavras constituintes.

Para cada modelo de apresentação, a tabela 5 mostra o número de adivinhas classificadas e os respectivos resultados. A coerência é ligeiramente mais alta para o modelo Características→Conceito, mas a principal diferença

está na originalidade. Aparentemente, a sensação de novidade é substancialmente menor quando o conceito é referido na pergunta. Isto é uma constatação relevante, a ser considerada no futuro. Contudo, após uma revisão manual dos resultados, não descartamos o impacto negativo de haver demasiados “*que é*” a ocorrer na resposta, por vezes de forma pouco natural. Assim, para já, a nossa decisão será alterar o modelo de apresentação da seguinte forma: quando a primeira característica é um adjectivo e a segunda um substantivo, o modelo passará a considerar que o substantivo é modificado pelo adjectivo. Veja por exemplo esta alteração aplicada a duas das adivinhas anteriores:

- *O que significa direitos humanos? Um homem plano.*
- *O que significa direitos humanos? Um homem liso.*

Outra alteração que acabamos por realizar, neste caso, no modelo Características→Conceito, foi,

Lista	#	Coerência		Originalidade		Humor	
		Mo	Md	Mo	Md	Mo	Md
Compostos	165	5	4	3	3	1	2
N-Adj	273	4	4	2	3	1	2
Compostos-d1	144	4	3	3	3	1	2
N-Adj-d1	165	4	3	3	3	1	2
W1+W2	213	4	3	3	3	1	1

Tabela 4: Classificação humana de acordo com as listas de conceitos iniciais.

sempre que possível, utilizar os adjetivos como substantivos na pergunta. Aplicado a um dos exemplos anteriores, como a palavra “plano” pode ser utilizada como adjetivo mas também como substantivo, o resultado é:

- *Que resulta do cruzamento entre um plano e um homem? direitos humanos.*

Contudo, as alterações anteriores não estão ainda reflectidas nos resultados apresentados da validação.

A tabela 6 mostra, para cada tipo de características extraídas (relações), o número de adivinhas e as suas classificações. Tipos de relação com menos de 10 adivinhas na amostra foram deixados de fora desta tabela. A maioria das adivinhas foi produzida através de sinónimos e hiperónimos, que são também os tipos de relação com mais instâncias na base de conhecimento. Quando agrupados por tipo de relação, os resultados não são conclusivos.

Tal como nos resultados globais, o potencial humorístico é sempre baixo. A originalidade é especialmente baixa para relações de antonímia combinadas com relações do tipo *dizSe-Sobre*. Muitas das adivinhas geradas com essa combinação de características serão demasiado literais, como no seguinte exemplo:

- *Qual é o contrário de vivo público? vida privada.*

A combinação entre características do tipo *dizSe-Sobre* com hiperonímia destacou-se com a moda mais alta para a coerência. Segue-se um exemplo de uma das adivinhas obtidas com essa combinação:

- *Que resulta do cruzamento entre uma norma e um ser? lei orgânica.*

A originalidade mais elevada foi obtida pela combinação de relações de antonímia e hiperonímia. A seguinte é um exemplo dessa combinação:

- *Qual é o contrário de organismo alto? planta baixa.*

Por fim, atendendo às diferenças culturais e linguísticas entre Portugal e Brasil, a tabela 7 apresenta os resultados separados por país. Há algumas diferenças mas, tendo em conta o tamanho da amostra utilizada, o número de avaliadores por adivinha (três) e ainda que os colaboradores de Portugal não terão avaliado exactamente as mesmas adivinhas que os do Brasil, não será possível tirar grandes conclusões. O mais evidente é a menor originalidade atribuída pelos avaliadores de Portugal às adivinhas que têm por base as listas N-Adj (obtida de um jornal português) e Compostos-d1 (que resulta de alterações a compostos recolhidos por uma equipa brasileira), o que tem reflexo no mesmo aspecto a nível global.

7 À procura do humor: exploração de outros modelos de adivinha

Depois das experiências relatadas nas secções anteriores, decidimos explorar outras estratégias para a geração de adivinhas, com algumas semelhanças, mas pequenas diferenças com vista ao aumento do potencial humorístico. Nesta secção descrevemos duas dessas estratégias e revelamos alguns exemplos de adivinhas obtidos em experiências iniciais.

Uma das estratégias passa por criar uma lista semelhante à lista W1+W2, isto é, em que cada entrada tem uma palavra que pode ser dividida em duas partes. Contudo, desta vez, a palavra original sofre pequenas alterações de letras ou sequências que têm um som igual ou com alguma semelhança com a original. Por exemplo, permitindo trocas de sequências como *ce* para *se*, *i* para *e*, *n* para *m*, ou até *on* para *ão* e *ção* para *som*. Apresentam-se alguns exemplos, criados com os mesmos modelos que as adivinhas anteriores, mas depois de gerar palavras com as trocas referidas:

- *Qual é o contrário de orgânico imaginário? sereal.*

Modelo	#	Coerência		Originalidade		Humor	
		Mo	Md	Mo	Md	Mo	Md
Características→Conceito	120	4	4	3	3	1	2
Contrário	36	4	3	3	3	1	2
Conceito→Características	166	4	3	1	2	1	1

Tabela 5: Classificação humana de acordo com a apresentação da adivinha.

Características	#	Coerência		Originalidade		Humor	
		Mo	Md	Mo	Md	Mo	Md
sinónimoDe, hiperónimoDe	53	4	3	2	3	1	2
sinónimoDe, sinónimoDe	50	3	3	3	3	1	2
sinónimoDe, dizSeSobre	37	4	4	1, 3	3	1	2
antónimoDe, sinónimoDe	24	4	3	1, 3	3	1	1
antónimoDe, hiperónimoDe	19	4	4	3, 4	3	1	2
causador, sinónimoDe	19	4	4	1, 3	2	1	1
dizSeSobre, hiperónimoDe	17	5	4	3	3	1	2
antónimoDe, dizSeSobre	14	4	3	1	2	1	1.5
partDe, sinónimoDe	12	3, 4	3	3	3	1	1

Tabela 6: Classificação humana de acordo com o tipo de características usado.

Lista	País	#	Coerência		Originalidade		Humor	
			Mo	Md	Mo	Md	Mo	Md
Total	Portugal	390	4	3	2	2	1	2
	Brasil	570	4	4	3	3	1	2
Compostos	Portugal	65	4	4	2	3	1	2
	Brasil	99	5	4	3	3	1	2
N-Adj	Portugal	115	4	4	2	2	1	2
	Brasil	157	5	4	3	3	1	1
Compostos-d1	Portugal	57	2	3	2	2	1	1
	Brasil	87	4	3	3	3	1	2
N-Adj-d1	Portugal	71	3	3	2	3	1	2
	Brasil	94	4	4	3	3	1	2
W1+W2	Portugal	81	1	3	1	2	1	2
	Brasil	132	4	3	3	3	1	1

Tabela 7: Classificação humana de acordo com as listas de conceitos iniciais.

- *Que resulta do cruzamento entre uma sequência e o que é ecológico? genecologia.*
 - *Qual é o contrário de trabalho doente? deversão.*
 - *Qual é o contrário de criação imaginária? artereal.*
 - *Qual é o contrário de história má? bomito.*
 - *Que resulta do cruzamento entre um imposto e um membro de melodia? obrigasom.*
 - *Que resulta do cruzamento entre o que é canino e o que é tradicional? cãotradição.*
 - *Que resulta do cruzamento entre o que é canino e um destino? cãosorte.*
- Numa outra estratégia, inspiramo-nos numa brincadeira de crianças, onde se procuram gerar falsos antónimos, exemplificada pela adivinha:
- *Qual é o contrário de paixão? mãe-tecto*

Para gerar este tipo de adivinhas, recorreremos diretamente à lista de conceitos W_1+W_2 , mas com um procedimento focado apenas nos antónimos das duas partes. Seguem-se alguns exemplos de adivinhas deste tipo geradas automaticamente:

- *Qual é o contrário de somali?*
silêncio-aqui.
- *Qual é o contrário de bombom?*
mau-mau.
- *Qual é o contrário de malcheiroso?*
bem-fedorento.
- *Qual é o contrário de causador?*
efeito-prazer.
- *Qual é o contrário de reverter?*
esquecer-carecer.
- *Qual é o contrário de diapositivo?*
noite-negativo.
- *Qual é o contrário de atrocidade?*
branco-mato.

De forma a aplicar a ideia original, e sem nenhuma restrição, a norma será gerar palavras inválidas. No futuro, uma forma de melhorar a selecção pode passar por considerar na pontuação a proximidade com palavras do léxico, por exemplo, calculando a distância de edição.

Ambas as estratégias aqui apresentadas requerem, certamente, mais trabalho, tanto ao nível das regras a aplicar e pontuação automática, como também ao nível da validação.

8 Conclusão

Descrevemos neste artigo um conjunto de abordagens baseadas em regras para a geração automática de adivinhas em português, onde se exploram recursos linguísticos disponíveis para esta língua, também eles descritos, juntamente com uma justificação da sua escolha, que teve em vista o aumento do potencial humorístico. Foram ainda apresentados exemplos de adivinhas e os resultados da validação humana de uma amostra, focada na coerência, originalidade e potencial humorístico das adivinhas.

Tanto quando sabemos, este é o primeiro trabalho deste tipo em português. Por se tratar de uma abordagem inicial ao tema, decidimos restringir-nos à geração de um tipo específico de adivinhas — baseadas num conceito conhecido e duas características extraídas, recuperadas de uma base de conhecimento, e apresentadas como um par pergunta-resposta — cuja geração já havia sido realizada por um sistema automático,

mas noutras línguas, nomeadamente inglês (Binsted & Ritchie, 1994).

O sistema que envolve as abordagens descritas foi baptizado de SECO. No futuro, pretende-se que produza mais adivinhas deste ou de outros tipos, através da definição de novas estratégias; variações do modelo base de apresentação através de pergunta-resposta; algum tratamento de pequenos aspectos linguísticos, incluindo alguns que, entretanto, já foram tratados com o objetivo de tornar o texto mais natural, ou avaliar a inserção de hífens entre conceitos do tipo w_1w_2 , para forçar a interpretação como duas palavras (e.g., *dia-positivo*); e, claro, melhorar o seu potencial humorístico. De forma a mostrar que ainda há muita coisa a experimentar, a penúltima secção deste artigo mostra algumas explorações iniciais, precisamente no sentido de gerar adivinhas com mais piada. Outros tipos a explorar incluem adivinhas do tipo: *Qual é a diferença entre X e Y?*; *O que têm X e Y em comum?*; ou *Qual é o cúmulo de X?*.

Em teoria, seria também possível gerar mais adivinhas do tipo em que nos focámos se fossem exploradas listas alternativas de conceitos iniciais (por exemplo, com verbos), outras bases de conhecimento, ou se fosse considerada a fonética das palavras. O último ponto poderia ser feito com recurso à representação fonética dos grafemas no CETEMPúblico (Veiga et al., 2011). No entanto, não colocando essa possibilidade de parte, acreditamos que em português é possível fazer muita coisa considerando apenas a grafia. No entanto, a partir do momento que tivermos mais adivinhas, torna-se ainda mais crucial conseguir identificar as mais interessantes. Apesar de, actualmente, as adivinhas serem pontuadas de forma automática, os resultados da validação manual sugerem que esta pontuação não está correlacionada com os aspectos avaliados. Algumas ideias a explorar na pontuação automática da originalidade passam por considerar o relacionamento das características usadas — obtido, por exemplo a partir de um modelo distribucional de palavras —, isto é, quanto menor o relacionamento, maior a novidade, sendo que alguma relação terá sempre de existir com partes do conceito inicial. Pontuar o potencial humorístico será mais desafiante, mas poderão ser exploradas características normalmente utilizadas no reconhecimento automático de humor (e.g., Mihalcea & Strapparava (2006)), tais como ambiguidade ou mesmo a utilização de calão. A utilização desta última em sistemas de geração de humor também não é novidade (Valitutti et al., 2013) e, no caso do português, poderíamos recor-

rer a um dicionário focado precisamente neste registo (Almeida, consultado em 2018). Por exemplo, quando utilizamos a base de conhecimento que inclui todos os triplos, muito maior, o nosso sistema gera a seguinte adivinha, com a presença de calão: “*Que resulta do cruzamento entre um fundo e o que é claro? cu aberto*”, obtida com recurso às características *fundo* hiperónimoDe *cu* e *claro* sinónimoDe *aberto*.

Para além dos anteriores, seria importante considerar outros aspectos na pontuação, tais como a resolubilidade, onde se poderia penalizar adivinhas que usam características demasiado genéricas e que, por isso, não restringem suficientemente o número de respostas possíveis. Este aspecto também foi considerado por outros autores (e.g., Labutov & Lipson (2012)) e, no nosso caso, poderia passar por ampliar para outras relações o que já é feito com a penalização de características com um número elevado de hiperónimos.

Por fim, e tal como outros fizeram (e.g., Guerrero et al. (2015), Gonçalo Oliveira et al. (2016)), pretendemos, num futuro próximo, desenvolver um *bot* na rede social Twitter que procurará gerar adivinhas inspiradas na tendências actuais. Como conceitos iniciais, um agente deste tipo poderá usar as tendências, conceitos frequentes em publicações relacionadas com as tendências, ou variações das mesmas.

Referências

- Almeida, José João. consultado em 2018. Dicionário aberto de calão e expressões idiomáticas. <http://natura.di.uminho.pt/jjbin/dac>.
- Attardo, Salvatore. 2008. A primer for the linguistics of humor. Em Victor Raskin (ed.), *The Primer of Humor Research*, chap. 3, 101–156. De Gruyter Mouton.
- Biemann, Chris & Eugenie Giesbrecht. 2011. Distributional semantics and compositionality 2011: Shared task description and results. Em *Workshop on Distributional Semantics and Compositionality*, 21–28.
- Binsted, Kim & Graeme Ritchie. 1994. An implemented model of punning riddles. Em *12th National Conference on Artificial Intelligence*, vol. 1, 633–638.
- Galvan, Paloma, Virginia Francisco, Raquel Hervás & Gonzalo Méndez. 2016. Riddle generation using word associations. Em *10th International Conference on Language Resources and Evaluation (LREC 2016)*, .
- Georges, Robert A. & Alan Dundes. 1963. Towards a structural definition of the riddle. *Journal of American Folklore* 76(300). 111–18. doi:10.2307/538610.
- Gonçalo Oliveira, Hugo. 2018. A survey on Portuguese lexical knowledge bases: Contents, comparison and combination. *Information* 9(2). 34. doi:10.3390/info9020034.
- Gonçalo Oliveira, Hugo, Diogo Costa & Alexandre Pinto. 2016. One does not simply produce funny memes! – explorations on the automatic generation of internet humor. Em *7th International Conference on Computational Creativity*, 238–245.
- Guerrero, Ivan, Ben Verhoeven, Francesco Barbieri, Pedro Martins & Rafael Perez y Perez. 2015. TheRiddlerBot: A next step on the ladder towards creative Twitter bots. Em *6th International Conference on Computational Creativity*, 315–322.
- Labutov, Igor & Hod Lipson. 2012. Humor as circuits in semantic networks. Em *50th Annual Meeting of the Association for Computational Linguistics: Short Papers - Volume 2*, 150–155.
- Manurung, Ruli, Graeme Ritchie, Helen Pain, Annalu Waller, Dave O’Mara & Rolf Black. 2008. The construction of a pun generator for language skills development. *Applied AI* 22(9). 841–869.
- Mihalcea, Rada & Carlo Strapparava. 2006. Learning to laugh (automatically): Computational models for humor recognition. *Computational Intelligence* 22(2). 126–142.
- Palma, Paul de & E. Judith Weiner. 1992. Riddles: Accessibility and knowledge representation. Em *COLING 1992 Volume 4: The 15th International Conference on Computational Linguistics*, 1121–1125.
- Pinto, Pedro, João Ramalhinho & Gonçalo Castro. 2017. *O Caderno das Piadas Secas – 500 Tentativas de ter graça*. Manuscrito Editora.
- Ramisch, Carlos, Silvio Cordeiro, Leonardo Zilio, Marco Idiart & Aline Villavicencio. 2016. How naked is the naked truth? a multilingual lexicon of nominal compound compositionality. Em *54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 156–161.
- Ranchhod, Elisabete, Cristina Mota & Jorge Baptista. 1999. A computational lexicon of Portuguese for automatic text parsing. Em *SIGLEX99 Workshop: Standardizing Lexical Resources*, 74–80.

- Rocha, Paulo Alexandre & Diana Santos. 2000. CETEMPúblico: Um corpus de grandes dimensões de linguagem jornalística portuguesa. Em Maria das Graças Volpe Nunes (ed.), *V Encontro para o processamento computacional da língua portuguesa escrita e falada (PRO-POR 2000)*, 131–140.
- Sjöbergh, Jonas & Kenji Araki. 2007. Automatically creating word-play jokes in Japanese. Em *Procs. of NL-178*, 91–95.
- Stock, Oliviero & Carlo Strapparava. 2006. Laughing with HAHAcronym, a computational humor system. Em *21st National Conference on AI - Volume 2*, 1675–1678.
- Tagnin, Stella E. O. 2005. O humor como quebra da convencionalidade. *Revista Brasileira de Lingüística Aplicada* 5(1). 247–257.
- Valitutti, Alessandro, Hannu Toivonen, Antoine Doucet & Jukka M. Toivanen. 2013. "Let everything turn well in your wife": Generation of adult humor using lexical constraints. Em *Proceedings 51st Annual Meeting of the Assoc. for Computational Linguistics*, vol. 2, 243–248.
- Veiga, Arlindo, Sara Candeias & Fernando Perdigão. 2011. Conversão de grafemas para fonemas em Português Europeu – abordagem híbrida com modelos probabilísticos e regras fonológicas. *Linguamática* 3(2). 39–51.