

Sobre as relações racionais intrinsecamente ambíguas

Rodrigo de Souza^{1*}

¹DEINFO – Universidade Federal Rural de Pernambuco (UFRPE)
Dois Irmãos – 52.171-900 – Recife – PE – Brazil

rodrigo.npmsouza@ufrpe.br

Abstract. *The family of rational relations, or the relations between words which can be realised by transducers (automata with inputs and outputs) is considered. Not every rational relation can be realised by an unambiguous transducer. Our main contribution is a characterisation of the intrinsically ambiguous k -valued rational relations based on the concept of lag between successful computations. We also show that it is undecidable whether the union of two rational functions is an intrinsically ambiguous relation. Finally, we show that the unambiguous k -valued relations are precisely the sums of k pairwise disjoint rational functions. Our proofs are built on structural constructions with automata based on the concept of covering of automata.*

Resumo. *A família das relações racionais, ou relações entre palavras que podem ser realizadas por transdutores (autômatos com entrada e saída) é considerada. Nem toda relação racional pode ser realizada por um transdutor não-ambíguo. Nossa principal contribuição é uma caracterização das relações k -valoradas intrinsecamente ambíguas baseada no conceito de deslocamento entre passeios. Também apresentamos uma prova simples de que é indecidível se a união de duas funções racionais é uma relação intrinsecamente ambígua. Finalmente, mostramos que as relações k -valoradas não-ambíguas são precisamente somas de k funções racionais duas a duas disjuntas. Nossas provas consistem de construções estruturais baseadas no conceito de revestimento de autômatos.*

1. Introdução

Dizemos que um autômato \mathcal{A} com transições rotuladas por elementos de um monóide M é não-ambíguo se passeios bem-sucedidos distintos de \mathcal{A} tem rótulos distintos; dizemos neste caso que o *comportamento* de \mathcal{A} (= o subconjunto de M dos rótulos dos passeios bem-sucedidos) é um *conjunto racional não-ambíguo*.¹

O conceito de ambiguidade está na raiz de alguns dos problemas mais desafiadores na área de Linguagens Formais e Autômatos. Tal é a situação mesmo no contexto mais elementar, onde M é um monóide livre A^* e \mathcal{A} é um autômato não-determinístico (transições rotuladas por letras de A). Neste caso, sempre é possível construir um autômato não-ambíguo equivalente (no limite, basta considerar um autômato determinístico equivalente, usando-se a construção dos subconjuntos). Em outras palavras,

*Este trabalho é apoiado pelo projeto *Problemas estruturais em modelos formais de Computação*, Edital MCTI/CNPQ/Universal 14/2014 (Processo 459957/2014-7).

¹Por limitação de espaço, não apresentamos neste resumo definições e notações básicas sobre autômatos e transdutores. Esse material pode ser consultado no livro de Jacques Sakarovitch [Sakarovitch 2009].

os *subconjuntos reconhecíveis*² de um monóide livre são não-ambíguos. Todavia, a compreensão desses objetos está longe de concluída: por exemplo, a estimativa da complexidade de operações clássicas, como complementação, envolvendo autômatos não-ambíguos esconde uma série de questões em aberto [Colcombet 2015].

Em se tratando de autômatos rotulados por pares de palavras, ou *transdutores*, onde o monóide M é um produto cartesiano $A^* \times B^*$ de dois monóides livres, os problemas correspondentes se tornam muito mais delicados. Antes de tudo, nem toda *relação racional* (= subconjunto racional de $A^* \times B^*$) é não-ambígua. Ou seja, certas relações racionais são *intrinsecamente ambíguas*: todos os transdutores que as realizam tem necessariamente pares de passeios bem-sucedidos distintos com o mesmo rótulo. Além disso, saber se uma relação racional (dada através de um transdutor) é intrinsecamente ambígua é um problema *indecidível*. Essa complicação se verifica mesmo entre relações mais elementares, como aquelas onde a imagem de toda palavra em A^* é um conjunto de cardinalidade no máximo k , para um k fixo – são as chamadas *relações k -valoradas*.³

A primeira propriedade é revisitada aqui sob um novo ponto de vista (Teorema 2.1), inscrito em um programa de abordagem de propriedades de autômatos através de *construções estruturais* com os mesmos. No centro dessas construções se destaca o conceito de *revestimento de autômatos*. Essa abordagem, desenvolvida na escola francesa da área – veja por exemplo discussão em [Sakarovitch 1998] – e adotada na tese de doutorado do autor [de Souza 2008], consiste em *explicitar* o resultado almejado em construções com os objetos envolvidos. Um exemplo característico é o *revestimento de Schützenberger*, definido em [Sakarovitch 1998]. Trata-se do produto cartesiano $\mathcal{A} \times \mathcal{A}_{\text{det}}$ de um autômato não-determinístico \mathcal{A} com o determinístico equivalente obtido com a construção dos subconjuntos. A projeção na primeira coordenada é um morfismo que estabelece uma bijeção entre os passeios bem-sucedidos de \mathcal{A} e $\mathcal{A} \times \mathcal{A}_{\text{det}}$, de forma que ambos os autômatos tem, intuitivamente, a mesma estrutura; ocorre que, através da eliminação de determinadas transições de $\mathcal{A} \times \mathcal{A}_{\text{det}}$, obtém-se um sub-autômato não-ambíguo equivalente a \mathcal{A} , que *seleciona* (através do morfismo de transições) exatamente um passeio para cada palavra no comportamento de \mathcal{A} . Esse objeto fornece por exemplo uma nova prova – uma prova estrutural – para propriedades clássicas de transdutores como o *Teorema da Uniformização* e o fato de que toda função racional é não-ambígua [Eilenberg 1974]. De forma mais geral, um revestimento de um autômato \mathcal{A} é uma expansão de \mathcal{A} , um novo e maior autômato \mathcal{B} , juntamente de um morfismo entre os passeios bem-sucedidos de \mathcal{A} e \mathcal{B} que estabelece uma bijeção entre esses passeios. Mas, a estrutura desses passeios aparece mais explícita em \mathcal{B} , de forma que a manipulação das transições e estados de \mathcal{B} leva à obtenção de sub-autômatos com propriedades notáveis.

A base de nossas demonstrações é o conceito de *revestimento lexicográfico*, cujas definição e propriedades desenvolvemos em [Sakarovitch and de Souza 2010] (ver também [de Souza 2008]) para demonstrar que toda relação racional k -valorada pode ser decomposta numa união de k funções racionais (melhorando um resultado de A. Weber,

²De um ponto de vista algébrico, os subconjuntos que se convencionou chamar de reconhecíveis e racionais são definidos de formas distintas; mas pelo Teorema de Kleene ambas as famílias coincidem quando M é um monóide livre, e usamos o termo “reconhecível” por ser mais comum no contexto brasileiro.

³Os transdutores onde a entrada de toda transição é uma letra (que são chamados de *tempo real* em parte da literatura) realizam as relações racionais de *imagem finita*: a imagem de toda palavra no domínio da relação é um conjunto finito. Doravante, os termos *transdutor* e *relação racional* referem-se a esse objeto.

cuja prova original é longa e difícil, e a complexidade envolvida na construção é maior). Em linhas gerais, um revestimento lexicográfico é baseado em uma ordem lexicográfica colocada entre os passeios bem-sucedidos do autômato de partida; do revestimento podemos então extrair sub-autômatos contendo precisamente os passeios que ocupam determinadas posições nessa ordem, como os menores (dentre aqueles com o mesmo rótulo), ou os k -ésimos menores, para um k fixo. É essa possibilidade de identificar posições específicas dessa ordem que torna o revestimento lexicográfico útil para problemas de ambiguidade. Frequentemente, essa abordagem leva a uma compreensão mais clara da propriedade em questão: por exemplo, nosso método permite concluir rapidamente que determinadas famílias de relações são intrinsecamente ambíguas (Teorema 2.1), enquanto que a abordagem comum na literatura consiste de um raciocínio combinatório *ad hoc* (no estilo do *lema do bombeamento*) aplicado a determinada relação particular [Sakarovitch 2009].

Também apresentamos uma caracterização das relações racionais intrinsecamente ambíguas com base no conceito de *deslocamento* entre passeios bem-sucedidos. O deslocamento entre passeios c_1 e c_2 com o mesmo rótulo de um transdutor \mathcal{T} é o máximo entre as diferenças de comprimento das saídas, considerando-se todos os prefixos de c_1 e c_2 com a mesma entrada; veja [Sakarovitch and de Souza 2010] para a definição formal. O máximo dentre todos os deslocamentos, considerando-se todos os pares de passeios bem-sucedidos de \mathcal{T} , é denotado por $\langle \mathcal{T} \rangle$, e pode ser limitado ou não. Em se tratando das relações racionais k -valoradas, o Teorema 3.1 diz que as intrinsecamente ambíguas são precisamente aquelas realizadas por transdutores onde $\langle \mathcal{T} \rangle$ não é limitado. Para essas relações, também mostramos um resultado estrutural: as não-ambíguas são precisamente as somas de funções racionais duas a duas disjuntas (Teorema 2.2). Completamos nossa exposição com um refinamento da indecidibilidade para a família das relações intrinsecamente ambíguas, em um enunciado envolvendo as relações 2-valoradas (Teorema 4.1).

Em razão do espaço reduzido, limitamo-nos na sequência a apresentar os enunciados de nossos resultados, e um comentário breve sobre a demonstração. Nossa intenção é exibir completamente nosso estudo de alguns aspectos da ambiguidade para as relações finitamente valoradas. Mais detalhes sobre as demonstrações podem ser encontrados na tese de doutorado do autor [de Souza 2008], que é o ponto de partida desta pesquisa.

2. Identificando relações intrinsecamente ambíguas

Teorema 2.1 *Se τ é uma relação racional não-ambígua, então, para todo inteiro $k > 0$, o conjunto $\{u \in \text{dom } \tau \mid \text{card}(u\tau) = k\}$ é reconectível.*

Esboço da prova Seja \mathcal{T} um transdutor não-ambíguo que realiza τ . Para cada u no domínio de τ , o número de passeios bem-sucedidos de \mathcal{T} com entrada u é exatamente $\text{card}(u\tau)$ (porque \mathcal{T} é não-ambíguo). Para cada inteiro $k > 0$ podemos construir um revestimento lexicográfico possuindo um sub-autômato que seleciona precisamente os passeios bem-sucedidos tais que há exatamente outros $k - 1$ com a mesma entrada. As saídas de τ restritas a esse sub-autômato formam o conjunto desejado.

O Teorema 2.1 permite identificar facilmente certas relações intrinsecamente ambíguas.⁴ Por exemplo, sejam σ_1, σ_2 os morfismos gerados por $a\sigma_1 = a, b\sigma_1 = 1, a\sigma_2 = 1, b\sigma_2 = a$ e σ a relação racional $\sigma_1 \cup \sigma_2$. Ou seja, $u\sigma = \{a^{|u|_a}, a^{|u|_b}\}$. Essa

⁴Mas não fornece um critério para decidir essa propriedade; como mencionamos, ela é indecidível.

relação racional 2-valorada é intrinsecamente ambígua. De fato, $\text{card}(u\sigma) = 1$ sse $|u|_a = |u|_b$, e assim $\{u \in \text{dom } \sigma : \text{card}(u\sigma) = 1\}$ não é reconhecível. De forma semelhante, podemos mostrar que toda relação racional contida em $(a, c)^*(b, 1)^* + (a, 1)^*(b, c)^*$ e contendo um número infinito de pares da forma $(a^i b^i, c^i)$ é intrinsecamente ambígua.

Teorema 2.2 *Uma relação racional k -valorada é não-ambígua se, e somente se, ela pode ser escrita como uma soma de k funções racionais duas a duas disjuntas.*

Esboço da prova A substância da prova consiste na decomposição de um transdutor k -valorado não-ambíguo \mathcal{T} em k transdutores funcionais não-ambíguos com comportamentos dois a dois disjuntos. O revestimento lexicográfico permite identificar, para cada $\ell \geq 0$, os passeios bem-sucedidos para os quais há ℓ menores na ordem lexicográfica. O sub-autômato correspondente dá origem a um transdutor funcional e não-ambíguo. Os k transdutores para $\ell = 0, 1, \dots, k - 1$ formam a decomposição desejada.

3. Uma caracterização

Teorema 3.1 *Seja $\tau : A^* \rightarrow B^*$ uma relação racional k -valorada. As condições seguintes são equivalentes:*

1. τ é intrinsecamente ambígua;
2. existe um transdutor \mathcal{T} realizando τ tal que $\langle \mathcal{T} \rangle = \infty$;
3. para todo transdutor \mathcal{T} realizando τ , $\langle \mathcal{T} \rangle = \infty$.

Comentário sobre a prova A substância da prova é a implicação (2) \Rightarrow (3), e é elaborada, dependendo de manipulações estruturais com passeios bem-sucedidos do revestimento lexicográfico. Os detalhes podem ser consultados nas referências citadas.

4. Indecidibilidade

Teorema 4.1 *É indecidível se a união de duas funções racionais é uma relação intrinsecamente ambígua.*

Esboço da prova A primeira etapa consiste em mostrar que não é decidível se a intersecção entre duas funções racionais é racional, o que estabelecemos com uma redução ao Problema da Correspondência de Post. Mostramos em seguida que a união de duas funções racionais é não-ambígua se, e somente se, sua intersecção é racional; essa segunda etapa faz uso do Teorema 2.1.

References

- Colcombet, T. (2015). Unambiguity in automata theory. In *Proceedings of Descriptive Complexity of Formal Systems - 17th International Workshop, DCFS 2015*, volume 9118 of *Lecture Notes in Computer Science*, pages 3–18. Springer-Verlag.
- de Souza, R. (2008). *Approche structurelle des transducteurs de norme bornée*. PhD thesis, TELECOM ParisTech.
- Eilenberg, S. (1974). *Automata, Languages, and Machines*. Academic Press, Inc.
- Sakarovitch, J. (1998). A construction on finite automata that has remained hidden. *Theoretical Computer Science*, 204(1-2):205–231.
- Sakarovitch, J. (2009). *Elements of Automata Theory*. Cambridge University Press.
- Sakarovitch, J. and de Souza, R. (2010). Lexicographic decomposition of k -valued transducers. *Theory of Computing Systems*, 47(3):758–785.