Sobre a decomposição de um transdutor bidirecional finitamente valorado

Rodrigo de Souza[1]

[1]Departamento de Computação – Universidade Federal Rural de Pernambuco (UFRPE)
Rua Dom Manuel de Medeiros, Dois Irmãos – 52171-900 – Recife – PE – Brazil

rodrigo.npmsouza@ufrpe.br

Abstract. *The possibility of decomposing a k-valued two-way transducer $\mathcal{T}$ on k functional two-way transducers is an open problem. We consider two particular cases of this problem: the one where $\mathcal{T}$ is input k-ambiguous, and the case where distinct images of every input word, by the relation realised by $\mathcal{T}$, have different lengths. We discuss how these decompositions can be obtained using constructions that we have presented to solve other problems: the decomposition of a k-valued transducer, and the construction of an uniformisation of a two-way transducer.*

Resumo. *A possibilidade de se decompor um transdutor bidirecional k-valorado $\mathcal{T}$ em k transdutores bidirecionais funcionais é um problema em aberto. Consideramos dois casos particulares desse problema: aquele em que $\mathcal{T}$ é k-ambíguo na entrada, e aquele em que imagens distintas de uma mesma palavra pela relação realizada por $\mathcal{T}$ têm comprimentos distintos. Discutimos como essas decomposições podem ser obtidas utilizando construções que apresentamos para resolver outros problemas: decompor um transdutor unidirecional k-valorado, e construir uma uniformização de um transdutor bidirecional.*

1. Introdução

Transdutores unidirecionais são máquinas de Turing com duas fitas, nas quais ambas as cabeças só se movem para a direita. Uma maneira equivalente de definí-los é como autômatos nos quais o rótulo de toda transição consiste de um par, formado por uma letra, entendida como a leitura de uma *entrada*, e uma palavra, a ser concatenada à direita de uma *saída* sendo construída. Dessa forma, transdutores realizam relações entre linguagens, as chamadas *relações racionais*. Transdutores vêm sendo estudados desde os primórdios da Teoria dos Autômatos e têm várias aplicações por exemplo em processamento de linguagem natural e Biologia Computacional [Mihov and Schulz 2019]. Consulte [Sakarovitch 2009] e [Muscholl and Puppis 2019] para definições desse e de outros conceitos discutidos neste texto, bem como um panorama de resultados fundamentais.

Recentemente, transdutores *bidirecionais* ganharam muita atenção. Em um transdutor bidirecional, a leitura da palavra de entrada pode acontecer nas duas direções; mas, toda palavra de saída é construída da esquerda para a direita, através da concatenação das saídas das transições que compõem uma computação. É um fato clássico da Teoria dos Autômatos que autômatos (sem saída) bidirecionais e unidirecionais reconhecem a mesma família de linguagens, resultado estabelecido no artigo fundador de Rabin e Scott [Rabin and Scott 1959] e de forma independente por Shepherdson, através de uma construção que chamamos aqui de *dobradura de um autômato* (ou transdutor) bidirecional [Shepherdson 1959]. Por outro lado, não é difícil de encontrar relações que não são racionais, mas que podem ser realizadas por um transdutor bidirecional (por exemplo, a função reverso). Ou seja, os transdutores bidirecionais realizam uma classe mais ampla de relações, e uma das propriedades mais importantes desses objetos é que as funções realizadas pelos determinísticos coincidem com aquelas definíveis na Lógica Monádica de Primeira Ordem [Engelfriet and Hoogeboom 2001].

Dessa forma, não é de se espantar que frequentemente a generalização de propriedades de transdutores unidirecionais para os bidirecionais seja bastante complexa. É o que se verifica com o problema da decomposição de uma relação de valoração limitada. Dizemos que uma relação entre monóides livres A^* e B^* é k-valorada, para um inteiro $k > 0$, se, para toda palavra $u \in A^*$, sua imagem pela relação é um conjunto contendo no máximo k palavras. Para as relações racionais, Weber mostrou o seguinte [Weber 1996]:

Teorema 1 (Weber 1996) *Toda relação racional k-valorada pode ser decomposta em uma união de k funções racionais.*

(Por *ser decomposto em uma união* entenda-se: a união das funções realizadas pelos k transdutores obtidos é igual à relação realizada pelo transdutor original.) Todavia, a transcrição desse enunciado para os transdutores bidirecionais é um problema em aberto:

Problema 1 *Todo transdutor bidirecional k-valorado pode ser decomposto em uma união de k transdutores bidirecionais funcionais?*

Nosso objetivo é explicitar como construções que utilizamos para obter uma nova prova para o Teorema 1 [Sakarovitch and de Souza 2010] e para construir uma uniformização de um transdutor bidirecional [De Souza 2013] podem ser utilizadas em conjunto para demonstrar casos particulares do Problema 1. Nossa contribuição reside nos dois enunciados seguintes:

Teorema 2 *Todo transdutor bidirecional k-ambíguo na entrada $\mathcal{T}$ pode ser decomposto em uma união de k transdutores bidirecionais funcionais.*

(Um transdutor é k-ambíguo na entrada se no autômato obtido apagando-se as saídas das transições, toda palavra reconhecida é o rótulo de no máximo k passeios bem-sucedidos.)

Teorema 3 *Seja $\mathcal{T}$ um transdutor bidirecional k-valorado com a propriedade de que imagens distintas de uma mesma palavra têm comprimentos distintos. Então $\mathcal{T}$ pode ser decomposto em uma união de k transdutores bidirecionais funcionais.*

Em razão de restrições de espaço, apresentamos na sequência uma discussão geral sobre as provas, e as referências das construções utilizadas, para mais detalhes.

2. Discussão sobre as provas

Iniciamos comentando nossas provas para o Teorema 1 [Sakarovitch and de Souza 2010] e a uniformização de um transdutor bidirecional [De Souza 2013].

Nossa prova para o Teorema 1 baseia-se no conceito de *revestimento de autômatos*. Intuitivamente, um revestimento de um autômato $\mathcal{A}$ é uma expansão de $\mathcal{A}$, um novo autômato $\mathcal{B}$ acompanhado de um morfismo de estados, que estabelece uma bijeção entre os passeios bem-sucedidos de $\mathcal{A}$ e de $\mathcal{B}$. Ambos os autômatos são assim equivalentes, mas em $\mathcal{B}$ os passeios bem-sucedidos podem estar mais "desenrolados", de forma que a propriedade almejada (sobre o comportamento de $\mathcal{A}$) corresponderá a uma escolha desses passeios, através de uma escolha de determinado sub-autômato de $\mathcal{B}$.

Com esse ferramental, abordamos o Teorema 1 em duas etapas, cada uma consistindo de um revestimento especial (que chamamos de revestimento lexicográfico): a primeira transforma um transdutor k-valorado $\mathcal{T}$ em um transdutor k-ambíguo na entrada; a segunda decompõe o autômato de entrada desse transdutor (que é k-ambíguo) em k autômatos não-ambíguos. Nessa segunda etapa, a união dos passeios bem-sucedidos dos k autômatos não-ambíguos obtidos projeta-se bijetivamente sobre os passeios bem-sucedidos de $\mathcal{A}$, o que escrevemos resumidamente da seguinte forma:

Lema 1 *Todo autômato k-ambíguo $\mathcal{A}$ pode ser decomposto em uma união de k autômatos não-ambíguos.*

Finalmente, o morfismo entre passeios permite recuperar, nas transições de cada um desses autômatos, as saídas do transdutor original.

Uma uniformização de um transdutor $\mathcal{T}$ é um transdutor funcional que associa toda palavra u no domínio de $\mathcal{T}$ a uma das palavras na imagem de u por $\mathcal{T}$. Em [De Souza 2013], provamos que todo transdutor bidirecional $\mathcal{T}$ pode ser uniformizado por um transdutor sequencial (= determinístico na entrada). Utilizamos para isso três construções. A primeira é a construção clássica de Shepherdson para transformar um autômato bidirecional $\mathcal{A}$ em um unidirecional $\mathcal{S}$, que chamamos de dobradura de $\mathcal{A}$. Intuitivamente, toda computação de $\mathcal{S}$ é a *dobradura* de uma computação *zigzag* de $\mathcal{A}$ em uma sequência de vetores coluna de estados. A segunda etapa utiliza um revestimento lexicográfico para obter, a partir de uma dobradura de $\mathcal{T}$, um autômato não-ambíguo, cujos passeios bem-sucedidos, se *desdobrados*, constituem uma uniformização de $\mathcal{T}$. Esse *desdobramento* é a terceira etapa de nossa construção: ela utiliza o que chamamos em [De Souza 2013] de *pathfinder*, uma adaptação de uma construção antiga de Hopcroft e Ullman (veja [Hopcroft and Ullman 1967]) que permite recuperar, a partir de uma dobradora não-ambígua, os passeios correspondentes do transdutor original.

Nossa prova para o Teorema 2 consiste no seguinte. Seja $\mathcal{A}$ o autômato de entrada de $\mathcal{T}$. Como $\mathcal{A}$ é k-ambíguo, então resulta da construção da dobradura $\mathcal{S}$ que esse autômato (unidirecional) também é k-ambíguo, e que há uma bijeção entre os passeios de $\mathcal{S}$ e os de $\mathcal{A}$. Em seguida, aplicamos em $\mathcal{A}$ um revestimento lexicográfico, obtemos uma decomposição desse autômato em k autômatos não-ambíguos – Lema 1. Finalmente, a construção *pathfinder* permite desdobrar cada um desses k autômatos em um transdutor bidirecional, de forma que a união dos mesmos consista na relação realizada por $\mathcal{T}$.

Nossa prova para o Teorema 2 é um pouco mais elaborada. Começamos substituindo as saídas de todas as transições de $\mathcal{T}$ por seu comprimento, ou seja, por uma palavra sobre um alfabeto unário. Obtemos então um transdutor bidirecional k-valorado. Transdutores com alfabeto de saída unário foram estudados em [Guillon 2016], e depreende-se dos resultados apresentados que os k-valorados realizam, de fato, relações racionais. Uma maneira explícita de demonstrar essa propriedade é construindo a dobradura $\mathcal{S}$ de $\mathcal{T}$, e adicionando em cada transição de $\mathcal{S}$ a soma das saídas na coluna de transições correspondentes. Isso transforma $\mathcal{S}$ em um transdutor k-valorado, que agora pode ser decomposto em uma união de k transdutores não-ambíguos através de nossa construção em [Sakarovitch and de Souza 2010]. Utilizando finalmente a construção *pathfinder*, obtemos k transdutores bidirecionais funcionais, e o fato de que a união das funções resultantes é a relação realizada por $\mathcal{T}$ vem da hipótese de que saídas distintas tem comprimentos distintos (e portanto aparecem em transdutores distintos na decomposição).

3. Comentários

Não fazemos nenhuma discussão sobre a complexidade (tamanho) das construções resultantes. Limitamo-nos a mencionar que cada etapa – dobradura, revestimento lexicográfico, *pathfinder* – introduz, em uma análise superficial, uma nova exponencial, pelo menos. Uma perspectiva de desenvolvimento é verificar se uma combinação mais fina dessas construções permite uma redução dessas estimativas.

Em [Muscholl 2017], o problema da decomposição é demonstrado para a família dos transdutores *streaming* com 1 registro. Esses transdutores correspondem a uma família particular de relações realizadas por transdutores bidirecionais. A prova é fortemente orientada a combinatória de palavras. Deixamos como trabalhos futuros uma comparação com nossos resultados.

Referências

De Souza, R. (2013). Uniformisation of two-way transducers. In *Language and Automata Theory and Applications. LATA 2013. Lecture Notes in Computer Science*, volume 7810, pages 547–558. Springer, Berlin.

Engelfriet, J. and Hoogeboom, H. (2001). MSO definable string transductions and two-way finite-state transducers. *ACM Trans. Comput. Logic*, 2(2):216–254.

Guillon, B. (2016). Input- or output-unary sweeping transducers are weaker than their 2-way counterparts. *RAIRO - Theoretical Informatics and Applications*, 50(4):275–294.

Hopcroft, J. E. and Ullman, J. D. (1967). An approach to a unified theory of automata. In *8th Annual Symposium on Switching and Automata Theory (SWAT 1967)*, pages 140–147. IEEE Computer Society.

Mihov, S. and Schulz, K. U. (2019). *Finite-state techniques : automata, transducers and bimachines*. Cambridge University Press.

Muscholl, A. (2017). A Tour of Recent Results on Word Transducers. In *Fundamentals of Computation Theory. FCT 2017*, pages 29–33.

Muscholl, A. and Puppis, G. (2019). The many facets of string transducers. In *Leibniz International Proceedings in Informatics, LIPIcs*, volume 126. Schloss Dagstuhl-Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing.

Rabin, M. O. and Scott, D. (1959). Finite Automata and Their Decision Problems. *IBM Journal of Research and Development*, 3(2):114–125.

Sakarovitch, J. (2009). *Elements of Automata Theory*. Cambdrige University Press.

Sakarovitch, J. and de Souza, R. (2010). Lexicographic Decomposition of k-Valued Transducers. *Theory of Computing Systems*, 47(3):758–785.

Shepherdson, J. C. (1959). The reduction of two-way automata to one-way automata. *IBM J. Res. Dev.*, 3:198–200.

Weber, A. (1996). Decomposing a k-valued transducer into k unambiguous ones. *RAIRO - Theoretical Informatics and Applications*, 30(5):379–413.