MONADIC SECOND-ORDER DEFINABLE GRAPH TRANSDUCTIONS(*)

Bruno COURCELLE

Université BORDEAUX-1,
Laboratoire d'Informatique
351, Cours de la Libération
33405 TALENCE, FRANCE

ABSTRACT

Formulas of monadic second-order logic can be used to specify graph transductions, i.e., multivalued functions from graphs to graphs. We obtain in this way classes of graph transductions, called monadic second-order definable graph transductions (or more simply definable transductions) that are closed under composition and preserve the two known classes of context-free sets of graphs, namely the class of Hyperedge Replacement (HR) and the class of Vertex Replacement (VR) sets. These two classes can be characterized in terms of definable transductions and recognizable sets of finite trees. These characterizations are independent of the rewriting mechanisms used to define the HR and VR grammars. When restricted to words, the definable transductions are strictly more powerful than the rational transductions with finite image; they do not preserve context-free languages. We also describe the sets of discrete (edgeless) labeled graphs that are the images of HR and VR sets under definable transductions: this gives a version of Parikh’s Theorem (i.e., the characterization of the commutative images of context-free languages) which extends the classical one and applies to HR and VR sets of graphs.

INTRODUCTION

The Theory of Formal Languages investigates finite devices defining sets of finite and countably infinite words and trees, compares their expressive powers, and investigates the solvability of the associated decision problems. These investigations make an essential use of transformations from words or trees to words or trees usually called transductions. Of special importance are rational transductions; they are closed under composition and inverse, and they preserve the families of recognizable and context-free languages. Tree transductions are more complicated, and there is no unique notion that can be considered as the analogue of that of a rational transduction. For each class of tree transductions, the

(*) Supported by the ESPRIT Basic-Research project 3299 ("Computing by graph transformations") and by the "Programme de Recherches Coordonnées: Mathématiques et Informatique".

(+) Unité associée au CNRS n° 1304, email : courcel@geocub.greco-prog.fr
closure under composition is a major concern, and so is the preservation of recognizability; we refer the reader to the survey by Raoult [Rao]. Another important transduction is yield that maps derivation trees of context-free grammars to the corresponding words. The context-free languages can be characterized as the images of the recognizable sets of (finite) trees under yield mappings.

The study of sets of finite and countably infinite graphs (and hypergraphs) by tools like grammars, systems of equations and logical formulas is a relatively recent development of the Theory of Formal Languages. The need for a manageable and powerful notion of graph transduction appears in constructions dealing with graph grammars and is of interest on its own.

This paper is a survey presenting the notion of a monadic second-order definable graph transduction (a definable transduction for short), which has been introduced more or less explicitly and sometimes in restricted forms in several papers [ALS, Cou5, Cou7-10, CE, Eng2]. We collect the main results and give references to their proofs.

We now introduce informally these transductions. The term "monadic second-order" refers to a logical language, the monadic second-order logic. We recall the rôle of logic for defining sets of graphs (or hypergraphs; all what we shall say concerning graphs applies to hypergraphs as well). Graphs can be described by relational structures, i.e., by logical structures with no function symbols. The domain of the structure representing a graph is the set of its vertices and edges put together; basic relations describe the incidence of vertices and edges and possible labellings. (This is actually not the only way to represent a graph; see Sections 1 and 4 for more details.) Hence, formulas of appropriate logical languages define properties of this graph. Monadic second-order logic is popular among logicians because of its expressive power and its decidability properties. (See Gurevich [Gu] for a survey.) For dealing with graphs, it is very useful because it can express many fundamental properties (like planarity, connectivity, k-colorability for fixed k) whereas several general decidability results hold. The sets of words (resp. of trees) characterized by a property expressible in monadic second-order logic are exactly the recognizable sets by the results of Büchi and Elgot [Bü, Elg], see [Tho, Thm 3.2] (resp. of Doner [Don], see [Tho, Thm 11.1]). Sets of graphs defined similarly by characteristic monadic second-order