Extended Abstract

Tree transducers are finite-state devices computing relations on trees. Their study was initiated by Thatcher (1970) and Rounds (1970), who established the classical top-down tree transducers that process the input tree from the root towards the leaves. Shortly afterwards, Baker (1973) introduced the bottom-up tree transducers that process the input tree from the leaves towards the root in analogy to the top-down and bottom-up tree automata (Thatcher, 1973). Due to applications in syntax-directed semantics (Fülöp and Vogler, 1998), tree transducers were extensively studied in the following years as detailed in (Gécseg and Steinby, 1984) and (Gécseg and Steinby, 1997). Notable extensions to the original top-down tree transducers include the top-down tree transducers with regular look-ahead by Engelfriet (1977), the attributed tree transducers of Fülöp (1981), and the macro tree transducers by Courcelle and Franchi-Zannettacci (1982) and Engelfriet and Vogler (1985).

In statistical machine translation (Koehn, 2009), syntax-based models (Chiang, 2010) [i.e., models that translate from or to syntax trees] have recently seen a lot of progress. It was identified already by Eisner (2003) that the classical linear top-down and bottom-up tree transducers cannot properly handle phenomena (such as rotation) that occur during the translation between natural languages. This result was presented for synchronous context-free grammars [SCFG] (Chiang, 2006), which is a formalism similar in spirit to (and essentially equally expressive as) the syntax-directed translation schemata by Aho and Ullman (1969), which were later refined to the more general bimorphism approach by Arnold and Dauchet (1982). Instead Eisner (2003) proposes synchronous tree substitution grammars [STSG], which are a restriction of the synchronous tree adjoining grammars [STAG] by Shieber and Schabes (1990). Indeed, STAG have a natural finite-state correspondence with the linear and complete bimorphisms (Arnold and Dauchet, 1982), which did not receive much interest from the theoretical computer science community in the sequel. Only with the advent of related models proposed by Galley et al. (2004) and corresponding translation systems following Graehl and Knight (2004), a systematic study of those finite-state tree transformers, now called extended [top-down] tree transducers, was initiated by Maletti et al. (2009).

Initial results for extended tree transducers showed that several results for classical tree transducers are no longer valid in the extended setting. For example, the composition results for classical tree transducers, derived by Engelfriet (1975) and Baker (1979), no longer hold for extended tree transducers, which was actually already confirmed in the seminal paper by Arnold and Dauchet (1982). However, essentially all reasonable extended tree transducers can capture important linguistic transformations (such as rotation) without the help of copying. Knight and Graehl (2005) provide a nice overview and establish requirements for an ideal syntax-based translation model. Following this call, extended tree transducers were thoroughly investigated and their basic properties have been established. In this invited talk, we recall the extended tree transducer model and summarize the results reported by Maletti et al. (2009) and Maletti (2011a) enriched with the recent composition closure results of Fülöp and Maletti (2013).

However, none of the extended tree transducer models that can handle rotation is closed under composition (Maletti et al., 2009). Consequently, May et al. (2010) developed an on-the-fly composition approach that can efficiently evaluate chains.
of multiple extended tree transducers. With the hope that the composition closure is achieved at a low level, the composition hierarchy of most relevant extended tree transducer classes has been investigated in (Fülöp and Maletti, 2013), where we confirmed that several important composition hierarchies are actually infinite. As an alternative to extended tree transducers, we proposed another model originally proposed already by Arnold and Dauchet (1982). This model, nowadays known as multi bottom-up tree transducers [MBOT] following the nomenclature of Fülöp et al. (2004) and Fülöp et al. (2005), offers closure under composition (Engelfriet et al., 2009), which removes the need to evaluate chains of models. Maletti (2010) summarizes the advantages following the requirements set by Knight and Graehl (2005), and Maletti (2011b) provides an automatic extraction of an MBOT translation model from the usual training data for syntax-based translation models. This tree-to-tree model was subsequently implemented by Braune et al. (2013) in the syntax-based component (Hoang et al., 2009) of the statistical machine translation framework Moses (Koehn et al., 2007). In addition, a string-to-tree MBOT model was developed and evaluated by Seemann et al. (2015a). Finally, several related MBOT models were evaluated by Seemann et al. (2015b) and compared to state-of-the-art models in statistical machine translation. The invited talk will also recall those efforts and provide the latest developments and evaluations.

References


