Compact Representations of Large Tree Automata for Music Notation Processing

Mots clés :
- Directeur de thèse : Florent Jacquemard
- Co-encadrant(s) :
- Unité de recherche : INRIA-Paris
- École doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal: Sciences et technologies de l'information et de la communication

Résumé du projet de recherche (Langue 1)
Several extensions of automata and tree automata have been studied for different purposes, for instance: - attribute grammars [1], for the translation of expression during compilation, - symbolic (tree) automata [2], to deal with large or infinite input alphabets, - weighted tree automata [3] for ranking words of trees in natural language processing, - some combination of the two latter for stream processing [4]. We are currently using extended tree automata models in the development of a framework for automated music transcription, the problem of converting the performance of a musician into structured music notation [5]. This old challenge in computer music has important applications in the context of music score edition and preservation, and for corpus captation in musicology. Our approach is based on an a priori language model of preferred music notations. It gives accurate transcription results, usually better than widely distributed score editing software, and in an efficient way. One key to efficiency is the use of a compact representation of tree automata with very large state sets. The representation uses attributes which, roughly, store the input data to transcribe. We implemented procedures for the enumeration of transcription solutions (by Dynamic Programming [6]) based on such compact representations, without uncompacting. This thesis proposes a fundamental study, in these settings, of classes of compact representations of tree automata with very large, and possibly infinite, state set. The representation will be based on attributes over finite or infinite domains, and involve different settings for attribute propagation and update. The class may embed some of the above models, for instance one attribute can be use to store weight values. The main goal is to study conditions enabling composition properties, decidability results, and tree series enumeration, with procedures which do not need uncompacting, or as few as possible. The PhD may include developments activities and experiments in the context of the transcription system described above, but this is not mandatory and depends only on the affinities of the candidate.

Informations complémentaires (Langue 1)