**A PRUNING HIERARCHICAL TREE ALGORITHM (PHTA) USING IN A\* ALGORITHM IN VIETNAMESE PARSING TECHNIQUE**

*Le Quang Thang1, Tran Do Dat1,2, Nguyen Thi Thu Trang2*

1. International Research Institute MICA CNRS UMI 2954 – Hanoi University of Science and Technology, INPG

2. School of Information and Communication Technology – Hanoi University of Science and

Technology

**ABSTRACT**

This paper presents our research on A-star (A\*) algorithm using pruning hierarchical tree in Vietnamese parsing technique in order to improve the quality of Vietnamese text-to-speech system. Based on the virtual node method proposed in [ref], we will describe our replace method: pruning hierarchical tree algorithm. This algorithm

**Categories**

Knowledge-based and information systems

**General Terms**

Algorithm, languages

**Key words**

A\*, parsing, PHTA

# Introduction

There are some parsing algorithms which have been researched and developed in Vietnamese Language. However, in our knowledge, there is not any parser using A\* algorithm.

# A\* algorithm for parsing

A\* algorithm operates on basically items called “node”. A *node* includes three attributes: *name, start,* and *end.* *Name* attribute indicates the name of node (also known as lexical tag or POS [ref]). And the attribute couple (*start, end*) is the start and end position of the text which is covered by *node* in the sentence.

Based on this basically items, the parser maintains two data structures: a chart (note as CHART) which records *nodes* for which (best) parses have already been found, and an agenda of newly-formed *nodes* needs to be processed (note as AGENDA).

## A\* parsing process

## Candidate estimating for A\* parsing

# Pruning hierarchical tree algorithm

## The context

## Fundamental hierarchical tree algorithm

### Basic idea

### The proposed model

#### Element classification phase

#### Generator the combinable chains