“Beijing Frowns and Washington Pays Close Attention”
Computer Processing of Relations between Geographical Proper Names in Foreign Affairs

Odile Piton¹ and Denis Maurel²

¹ CERMSEM, Université Paris 1, 12, place du Panthéon, 75005 Paris, France, piton@univ-paris1.fr
² LI/E3i-University of Tours, 64 avenue Jean-Portalis, F37200 Tours, France, maurel@univ-tours.fr

Abstract. We present results of the project Prolex. The aim of the project is the automated analysis of proper names, especially a description of relations between different proper names in a text. The system currently works with geographical proper names (place names, derived adjectives and names of inhabitants) in French. It consists of a database containing specific types of proper names and relations between the different names. Using these names and relations, the program can group the proper names appearing in a text that may belong together (such as Beijing-Chinese-Pekinese-China; American-United States-Washington). This is done by constructing an association matrix between them and by computing the transitive closure of this Boolean matrix. The method is explained with an example.

Key words: Proper Names, Place Name, Inhabitant Name, Natural Language Processing (NLP), Electronic Dictionary, Relational Data Base, Geographical Proper Names, Computing Method, Transitive closure.

1 Introduction

"Beijing frowns and Washington pays close attention"

"Pékín fonce les sourcils et Washington ouvre l’œil" (Belgium Newspaper Le Soir, 02/25/00)

The Prolex project is a university research project whose purpose is the automatic processing of proper names. In particular, it studies the relations between the various forms of the proper names of a text. This relational aspect is original. Current research about proper names is based on heuristics. Some simply declare as "proper names" all the unknown words starting with a capital letter [14]. More sophisticated ones use local grammars [57,3] and propose to classify the identified proper names [44]. Today, the whole of this research often merges with more general research on named entities. This research was the occasion...
for various evaluations, especially within the framework of Muc conferences [13].
In NLDB’99, we presented our database of toponyms (place names), inhabitant
names and derived adjectives that we constructed for the Project Prolex [12].
We have introduced that we register the links between these words, in the pur-
pose of the automated analysis of geographical proper names. We shall present
linguistic and computational aspects of a use of our Geographical Proper Names
Database for the Natural Language Processing of a text dealing with interna-
tional politics.
We have acquired some experiment with the first database made by Claude
Belleil [1], for France. We had noticed that the use of a database is much too
slow, as concerns NLP. So we have used transducers automatically build from
it. We call it ”electronic dictionary” . It is in fact a minimal transducer, within
the meaning of [9]. It is built directly by the algorithm of [8]. This transducer
includes only a part of the direct links of the DB.
We will explain the method, not the tool that we will make. We study the prop-
erties of texts dealing with international affairs (Section 2), then we explain the
different steps of our method (Section 3). First we collect information from our
database (Section 3.3.1), then we sum it up (Section 3.3.2). Next we build a
Boolean matrix of the links between the collected words (Section 3.3.3). Fourth
we compute the transitive closure of the matrix (Section 3.3.4). Lastly, we obtain
tags to add to the geographical proper names of the text (Section 3.3.5).

2 Properties of the Texts Dealing with International
Politics
Our purpose is to automatically compute texts of newspapers. They have some
peculiarities. Let us look at a text of an example, from the French newspaper Le
Monde (02/23/00):

"Pékín menace Taipêh d’intervention militaire en cas d’enlisement du
processus de réunification... C’est un glissement d’importance dans la
position de Pékín à l’égard de Taiwan ...on en trouve trace dans de
multiples déclarations de dirigeants pékinos... Le raidissement de sa
position sur une intervention militaire traduit l’irritation croissante de
Pékín devant les obstacles placés par les Taiwanais sur le chemin de
la réunification alors que la récupération de Hong-Kong et de Macao
a été bouclée sans heurts.... Les Chinois font certes un geste en accep-
tant d’ouvrir les discussions sur un pied d’égalité ... Les Américains
intensifient leurs manœuvres diplomatiques afin d’éviter la réédition de
la crise des missiles de la présidentielle du printemps 1996. L’objectif de
Washington est d’ouvrir des canaux de communication au plus haut
niveau[1]..."

1 Beijing has threatened to use force against Taipei in the event of the sinking of the
reunification process. It is a modification of importance in the position of Beijing