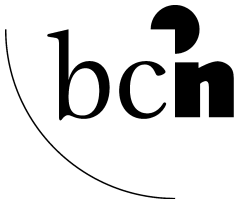
Finite-state pre-processing for natural language analysis

Prins, Robbert

Finite-State Pre-Processing for Natural Language Analysis

Robbert Prins



The work in this thesis has been carried out under the auspices of the Behavioral and Cognitive Neurosciences (BCN) research school, Groningen, and has been part of the PIONIER project *Algorithms for Linguistic Processing*, supported by grant number 220-70-001 from the Netherlands Organization for Scientific Research (NWO).

Groningen Dissertations in Linguistics 53
ISSN 0928-0030
Document prepared with LaTeX
Printed by PrintPartners Ipskamp, Enschede

Rijksuniversiteit Groningen

Finite-State Pre-Processing for Natural Language Analysis

Proefschrift

ter verkrijging van het doctoraat in de
Letteren
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
donderdag 12 mei 2005
om 14.45 uur

door

Robbert Paul Prins

geboren op 18 december 1977
te Ten Boer

Promotor:	Prof. dr. ir. J. Nerbonne
Copromotor:	Dr. G.J.M. van Noord
Beoordelingscommissie:	Prof. dr. E.W. Hinrichs Prof. dr. M. Johnson Prof. dr. G. Satta

Acknowledgements

At the end of my research and the beginning of this thesis I would like to thank a number of people for helping me.

To start with, I would like to thank my supervisor Gertjan van Noord and my second supervisor and promotor John Nerbonne. Discussing topics with Gertjan always helped when I was stuck or in need of some inspiration. John also helped me out by reading and commenting on the chapters as I was writing them. Thanks go out to the members of my reading committee for taking upon them this task of reading, judging, and commenting on the whole of the thesis.

Then I would like to thank my other colleagues at Alfa-Informatica and especially my roommates Gerlof, Francisco, and Tamas. Francisco deserves a special mentioning in any Alfa-Informatica thesis for being the printer guy. Then there were the cooking club evenings with Leonoor, Tanja, and Menno, which I really enjoyed. Menno and I also took part in several running events, and to much critical acclaim! I want to thank Francisco en Tamas a second time for being so kind as to accept the role of "paranimfs" at my defence. Thanks go out to Stasinos for the stylefile used to format this dissertation.

I want to thank Jan Daciuk for writing and making available his finite-state automata package of which my own programs make heavy use, Mark-Jan Nederhof for taking the time to read and suggest improvements on part of my thesis, and Hans van Halteren for supplying me with data used in his own work, allowing me to perform additional comparative experiments.

Finally I thank my parents for their support during those four years.

Contents