



KATHOLIEKE UNIVERSITEIT  
**LEUVEN**

**FACULTEIT**  
INGENIEURSWETENSCHAPPEN

Master  
Computer-  
wetenschappen

Masterproef  
*Jens Claes*

Promotor  
*Marc Denecker*

Begeleiders  
*Bart Bogaerts*  
*Laurent Janssens*

Academiejaar  
2016-2017



# Automatisch vertalen van logigrammen naar logica

## Motivatie

In het *Knowledge Base*-paradigma wordt een probleem gereduceerd tot een specificatie waarop verschillende inferenties worden uitgevoerd. Een formele specificatie is echter moeilijk te schrijven. Het automatisch vertalen van natuurlijke taal naar logica lost dit probleem op.

Deze thesis onderzoekt het vertalen van een logigram (een puzzel met een aantal constraints in natuurlijke taal) naar een formele specificatie in logica

		Persoon				Kleur				Land		
		An	Bob	Charles	Doris	Blauw	Groen	Rood	Paars	Belgie	Frankrijk	Italië
Dier	Kat											
	Hond											
	Paard											
	Varken											
Land	Belgie											
	Frankrijk											
	Italië											
	Spanje											
Kleur	Blauw											
	Groen											
	Rood											
	Paars											

## Voorbeeld - The contestant from Mount Union threw the black darts

### Lexicon

Woord	Categorie	Betekenis
the	Lidwoord (DET)	$\lambda R \cdot \lambda S \cdot \exists x \cdot R(x) \wedge S(x)$
contestant	Substantief (NOUN)	$\lambda x \cdot \text{contestant}(x)$
from	Voorzetsel (PREP)	$\lambda N \cdot \lambda P \cdot \lambda x \cdot N(\lambda y \cdot \text{from}(x, y) \wedge P(x))$
Mount Union	Eigenaam (PN)	$\lambda P \cdot P(\text{MountUnion})$
threw	Overgankelijk werkwoord (TV)	$\lambda N1 \cdot \lambda N2 \cdot N2(\lambda x2 \cdot N1(\lambda x1 \cdot \text{threw}(x2, x1)))$
the black darts	Eigenaam (PN)	$\lambda P \cdot P(\text{TheBlackDarts})$

### Grammatica

Grammaticale regel	Betekenis
$S \rightarrow NP VP$	$\llbracket VP \rrbracket(\llbracket NP \rrbracket)$
$NP \rightarrow DET N$	$\llbracket DET \rrbracket(\llbracket N \rrbracket)$
$NP \rightarrow PN$	$\llbracket PN \rrbracket$
$N \rightarrow NOUN NMOD$	$\llbracket NMOD \rrbracket(\llbracket NOUN \rrbracket)$
$NMOD \rightarrow PREP NP$	$\llbracket PREP \rrbracket(\llbracket NP \rrbracket)$
$VP \rightarrow TV NP$	$\llbracket TV \rrbracket(\llbracket NP \rrbracket)$

### Een semantisch framework

(Blackburn et. Al 2005, 2006)

Het lexicon is verschillend per logigram. De grammatica is gedeeld voor alle logigrammen.

De betekenis van een woord is een functie van de lexicale categorie.

Compositionality: de betekenis van een woordgroep is een combinatie van de betekenissen van de woorden waaruit ze bestaat. Zo wordt de betekenis van de woorden naar boven toe gepropageerd.

## Types en het formeel vocabularium

Substantieven en eigennamen introduceren een *basistype*  
Overgankelijke werkwoorden en voorzetsels introduceren een *afgeleid type* (met 2 basistypes als argument).

Veronderstelling: elk woord heeft 1 type per logigram. Bij meerdere constraints, unificeren de woorden die meerdere keren voorkomen zo de types. Verdere unificatie verloopt via vragen aan de gebruiker i.v.m. synonymie van woorden.

**Eigennamen** worden vertaald naar constanten van constructed types. Door unificatie van de basistypes worden deze eigennamen gegroepeerd.

**Substantieven** geven een naam aan het type.

**Voorzetsels** en **overgankelijke werkwoorden** introduceren een predicaat.

**Extra axioma's** (toegevoegd aan de theorie, specifiek voor logigrammen)

- Twee predicaten met dezelfde signatuur zijn gelijk
- Elk predicaat is een bijectie

## Resultaten

**Gegeven:** Aantal types, de constraints (in het Engels) en het logigram-specifiek lexicon

**Extra vragen aan de gebruiker:**

- Unificatie types (op basis van synonymie van woorden)
- Domein voor numerieke types

**Resultaat:** Vocabularium + Theorie in IDP

Met een grammatica op basis van ... puzzels toegepast op ... nieuwe puzzels

Problemen	Aantal
Geen	
Datums	
...	

Conclusie: ...