Ontological semantics in modified categorial grammar

Szymczak, Bartlomiej Antoni

Published in:
International Multiconference on Computer Science and Information Technology, 2009. IMCSIT ’09

Link to article, DOI:
10.1109/IMCSIT.2009.5352712

Publication date:
2009

Document Version
Publisher’s PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Ontological Semantics in Modified Categorial Grammar

Bartłomiej Antoni Szymczak
DTU Informatics, Technical University of Denmark
International Language Studies and Computational Linguistics
Copenhagen Business School
Email: bas@imm.dtu.dk

Abstract—Categorial Grammar is a well established tool for describing natural language semantics [1]. In the current paper we discuss some of its drawbacks and how it could be extended to overcome them. We use the extended version for deriving ontological semantics from text. A proof-of-concept implementation is also presented.

I. INTRODUCTION

In this paper we extend the usual Categorial Grammar framework in order to achieve more flexibility. We also present how it can be used with an ontological component, which imposes well-formedness restrictions on sentences.

The objective is to integrate formal ontologies with semantic domains. Such an ontology-oriented semantics may be useful, e.g., in content-based text search. We introduce type ω as concepts in the ontology is in line with the adoption of the Davidsonian view. We also disregard meaning of determiners in the present fragment, as the resulting semantics is intended to be used for content-based text search. We don’t want to be specific about ω, but in the current example it could be replaced by α, which is traditionally used for representing sets. We will represent intersection of concepts using constant ∩ααα, e.g., child ∩ tall will be represented as ∩αααChildα ∩ Tallα.

A simple skeleton ontology can be represented using a set of factual clauses of the form:

\[ Sub_{\alpha\alpha\alpha}Child_{\alpha}Person_{\alpha} \]
\[ Sub_{\alpha\alpha\alpha}Person_{\alpha}Physical_{\alpha} \]
\[ Sub_{\alpha\alpha\alpha}Person_{\alpha}Animate_{\alpha} \]
\[ Sub_{\alpha\alpha\alpha}Running_{\alpha}Action_{\alpha} \]

Sub_{\alpha\alpha\alpha} is a direct descendant relation, representing the lines that are drawn in the Hasse diagrams. Isa_{\alpha\alpha\alpha} can be specified as the (reflexive) transitive closure of Sub_{\alpha\alpha\alpha} in the following way:

\[ ∀cα[Isa_{\alpha\alpha\alpha}cαcα] \]
\[ ∀cα∀aα[Isa_{\alpha\alpha\alpha}cαaα ⊃ Sub_{\alpha\alpha\alpha}cαpα ∧ Isa_{\alpha\alpha\alpha}pαaα] \]

We use constants of type ρ to represent roles (binary relations), e.g., Agt_{ρ}, Loc_{ρ}. We can use Peirce product [6] to create compound concepts. We use constant :ααα ρ, for that purpose, e.g., agt : child will be represented as :ααα ρ Child_{α}. Such a concept formation is well-known from Description Logics, where it would be represented as ∀agt.child.

For a sample sentence “The tall kid runs”, we would like to derive the following ontologic–algebraic meaning:

\[ running ∩ agt : (child ∩ tall) \]

This can be represented in type theory as a wff_{α}:

\[ ∩αααRunning_{α}[ααα ρ Agt_{ρ}[ααα Child_{α}Tall_{α}]] \]

We shall extend the specification of subsumption relation to accommodate for the intersection:

\[ ∀xα∀yα∀zα[[Isa_{ααα}∩αααxαyαzα] ⊃ [Isa_{ααα}xαyαzα]] \]
\[ ∀xα∀yα∀zα[[Isa_{ααα}∩αααxαyαzα] ⊃ [Isa_{ααα}yαzα]] \]

III. MODIFIED CATEGORIAL GRAMMAR

We introduce type ω. Constants of this type represent words in English, e.g., W_{kidω}, W_{tallω}, W_{runsω}. Notice that we use different names for constants denoting words and those representing concepts, e.g., W_{tallω} and Tallα. In this way they are not confused.

Let us define a new type, say, κ, which we will use for lexical entries. We also define three constants, which act as type constructors:
Lexical entries are represented using the predicate constant \( L_{\text{soc}} \). The lexicon consists of the set of factual clauses of the form:

\[
L_{\text{soc}} \Rightarrow W_{\text{kid}}[E_{\text{soc}} \text{Child}_a] \\
L_{\text{soc}} \Rightarrow W_{\text{tall}}[F_{\text{soc}} \text{Child}_a] \\
E_{\text{soc}} 
\]

The novelty in our approach is the use of predicates in meanings of words. In standard Categorial Grammar, lambda terms are used for that purpose. They are combined using \( \beta \)-reduction, with the only provision of categorial agreement. Consider the sentence “vitamin smiles.” In traditional CG, “vitamin” has category \( np \) and meaning \( \lambda x.\text{smile}(x) \). Since the categories fit together, the meanings get combined using \( \beta \)-reduction, and the sentence gets the meaning \( \text{smile(vitamin)} \). While the sentence is correct syntactically, it’s incorrect from an ontological point of view. Unfortunately, usual CG does not allow us to introduce any ontological restrictions on the semantics. Our approach can reject this sentence, as we can use the following lexical entry for “smiles”:

\[
L_{\text{soc}} \Rightarrow W_{\text{smiles}}[B_{\text{soc}}] \\
\forall c_a[[P_{\text{tall}}[E_{\text{soc}}]c_a]][E_{\text{soc}}] \supset [E_{\text{soc}}] \supset [\text{Child}_a] \\
\supset [\text{Physical}_a] \\
\]

The inability of \( \beta \)-reduction to fail has been already recognized as a problem by G. Ben–Avi and N. Francez in [7]. They have introduced a new formalism, which includes “\( \beta \)-reduction for ontologically–well typed \( \lambda \)-terms”, among 12 definitions that constitute “the Ontological Lambek Calculus”. Our proposal, however, has a few further advantages:

- It’s very formal – it’s formalized fully within \( C \).
- It’s very simple – it consists of a few formulas only.
- It’s very flexible – it allows adding arbitrary restrictions, e.g. one might like to use restrictions based on parthood relation rather than subsumption.

- It can be implemented in a straight-forward way, as presented in Section V.

For those reasons, rather than using \( \beta \)-reduction, we propose using a general proof system, which is a machinery having inherently the notion of failure at disposal. Now, desired semantic restriction can be handled by non-provability of a certain statement, e.g.:

\[
K \not\vdash \exists x_a[P_{\text{smiles}}[E_{\text{soc}}]Vitamin_a]x_a \\
\]

In the above \( K \) denotes a set of formulas consisting of lexical assertions and rules introduced throughout this paper. So a vitamin cannot smile, but a child certainly can:

\[
K \vdash \exists x_a[P_{\text{smiles}}[E_{\text{soc}}]\text{Child}_a]x_a \\
\]

Not only are we interested in the provability of such a goal, but we would also like to know what is the resulting semantics. In this case it’s: \( E_{\text{soc}}[\text{Child}_a] \supset [\text{Child}_a] \supset [\text{Aggr}_a] \)

In other words:

\[
K \vdash [P_{\text{smiles}}[E_{\text{soc}}]\text{Child}_a] \\
\supset [E_{\text{soc}}] \supset [\text{Child}_a] \supset [\text{Aggr}_a] \\
\]

IV. ELIMINATION RULES

Recall that the forward sequent rule of the natural–deduction Lambeck Calculus takes the form:

\[
a \Rightarrow \Phi_1 : A/B \quad b \Rightarrow \Phi_2 : B \\
\quad \vdash \Phi_2 : A \supset \Phi_1(B) \\
\quad \vdash \Phi_2 : B \supset A \\
\]

and the backward rule:

\[
a \Rightarrow \Phi_1 : B \quad b \Rightarrow \Phi_2 : B \supset A \\
\quad \vdash \Phi_2 : A \supset \Phi_1(B) \\
\quad \vdash \Phi_2 : B \supset A \\
\]

In the above \( a, b \) is the concatenation of \( a \) and \( b \). Notice that the only provision for the elimination to take place is the agreement of syntactic categories. This is because \( \Phi_1(\Phi_2) \) is always a well-formed \( \lambda \)-term, which can be \( \beta \)-reduced.

For the ease of reading, let us present the forward elimination rule of the generalized grammar in a similar, though quite informal way:

\[
a_0 \Rightarrow F_{\text{soc}} B_{\text{soc}} \supset b_0 \Rightarrow j_{\text{soc}} K \vdash [B_{\text{soc}}]x_0 B_{\text{soc}} \\
\quad \vdash a_0, b_0 \Rightarrow j_{\text{soc}} K \\
\]

Let us formalize our generalized elimination rules in \( C \). The derivation relation \( (\Rightarrow) \) can be represented using a constant \( R_{\text{soc}} \). The sequences of meanings will be represented using a well–known logic representation for lists. For the empty list we use the \( N_{\text{soc}} \) constant, and the list constructor is represented by \( L_{\text{soc}} \).

The list concatenation can be represented using the list appending well–known from logic programming, though here defined for non–empty lists only:

\[
\forall x_0 \forall y_0 \forall t_0 \forall z_0 [A_{\text{soc}}][L_{\text{soc}}x_0]\supset [L_{\text{soc}}y_0]t_0 \supset [L_{\text{soc}}y_0]z_0 \supset [A_{\text{soc}}][L_{\text{soc}}y_0]t_0 \supset [L_{\text{soc}}y_0]z_0 \\
\forall h_0 \forall y_0 \forall m_0 \forall t_0 \forall [A_{\text{soc}}][L_{\text{soc}}h_0]t_0 \supset [L_{\text{soc}}h_0]m_0 \supset [A_{\text{soc}}][L_{\text{soc}}h_0]m_0 \\
\]
The above use of $A_{\alpha \beta \gamma}$ asserts that the concatenation of $a$ and $b$ yields $g$, or $g$ is split into $a$ and $b$. Now the forward elimination rule can be formalized in $C$:

$$\forall a \forall b \forall g \forall j_k \forall k \alpha \beta \gamma [R_{\alpha \beta \gamma} g j_k k] \supset [A_{\alpha \beta \gamma} a b g j_k k] \land [R_{\alpha \beta \gamma} g j_k k] \land [p_{\alpha \beta \gamma} j_k k]$$

Similarly, the backward elimination rule:

$$\forall a \forall b \forall g \forall j_k \forall k \alpha \beta \gamma [R_{\alpha \beta \gamma} a b g j_k k] \supset [A_{\alpha \beta \gamma} a b g j_k k] \land [R_{\alpha \beta \gamma} a b g j_k k] \land [P_{\alpha \beta \gamma} j_k k]$$

We also need the following grammatical axiom, which expresses that if the list of meanings contains only one element, that element is the resulting meaning:

$$\forall k \alpha \beta \gamma [R_{\alpha \beta \gamma} L_{\beta \gamma} k_\alpha k_\beta k_\gamma]$$

The English text can also be represented formally in $C$. We introduce a new type $\zeta$ for that purpose. We will represent the text as a list of words. An empty list of words will be represented by a constant $\text{nil}_\zeta$ and a list constructor by a constant $T_{\zeta \xi}$.

V. IMPLEMENTATION OUTLINE

Using $C$ instead of some kind of untyped logic as the underlying formalism has important advantages regarding the implementation. It forces us to think of the type of every formula, subformula and symbol. Thanks to that, the resulting formalization of the grammar is well-suited for implementation in a strongly-typed programming language. Using such a language (e.g. Mercury instead of Prolog) helps avoid very hard-to-find bugs and allows the compiler to generate much faster code.

We are interested in translating the formal specification given so far to a Prolog-like logic programming language, so that we can execute specific queries. For that purpose, we have formalized the grammar in $C$ using only definite clauses. Furthermore, we perform the following steps:

- We write all the constants in lower-case
- We write all variables in upper-case
- We use $i$ for intersection constant and $p$ for the Peirce product constant
- We drop the subscripts indicating types in all formulas
- We replace ‘$\supset$’ and ‘$\land$’ with ‘$\to$’ and ‘$\cdot$’, respectively.
- We add a period at the end of each formula.
- We replace curried notation of argument application with a non-curried one.
- We simply remove the universal quantification over all variables, as it’s implicit.

Following all the given steps results in the program presented below. The result is a perfectly valid program in the Mercury logic programming language. Mercury is however strongly typed, so we need to add a few auxiliary definitions in order to compile it.

The type $\alpha$ can be modeled as follows:

```mercury
:-type alpha ---> child; tall; physical; action; person; running; smiling; animate; ... i(alpha,alpha); p(role,alpha).
```

We need to add the following type, mode and determinism specification for the subsumption predicate:

```mercury
:-pred sub(alpha::in,alpha::out) is nondet.
```

This tells Mercury that predicate `sub` takes an entity of type $\alpha$ as input, and computes an entity of type $\alpha$ as output. Furthermore, we specify that `sub` is a nondeterministic predicate, meaning that it can compute multiple outputs for one input. We need to provide similar specifications for all predicates in our program.

The factual subsumption database, translated directly from our previous definition in $C$:

```mercury
sub(child,person).
sub(person,physical).
sub(person/animate).
sub(running,action).
sub(smiling,action).
...
```

The type $\rho$ of roles can be defined in Mercury as:

```mercury
:-type role --->
tmp ; loc ; prp ; wrt ; chr ; cum ; bmo ; cby ; cau ; cmp ; pof ; agt ; pnt ; src ; rst ; dst ; via ; ...
```

Clauses defining the `isa` relation:

```mercury
isa(C,C).
isa(C,A):-sub(C,P),isa(P,A).
isa(i(X,Y),Z):-isa(X,Z).
isa(i(X,Y),Z):-isa(Y,Z).

isa(C,animate).
isa(C,physical).
```
The lexicon, translated to Mercury:

\[
\begin{align*}
\text{lex}(w_{\text{kid}}, e(\text{child})). \\
\text{lex}(w_{\text{tall}}, f(p_{\text{tall}})). \\
\text{lex}(w_{\text{runs}}, b(p_{\text{runs}})). \\
\text{lex}(w_{\text{smiles}}, b(p_{\text{smiles}})). \\
\text{lex}(w_{\text{the}}, f(p_{\text{the}})). \\
\text{lex}(w_{\text{a}}, f(p_{\text{the}})). \\
\end{align*}
\]

The definition of non-empty list appending:

\[
\begin{align*}
\text{a}(l(X,n), l(Y,T), l(X,l(Y,T))). \\
\text{a}(l(H,L), M, l(H,T)) :- \text{a}(L,M,T).
\end{align*}
\]

The rules formalized in \( C \) take the following form in Mercury:

\[
\begin{align*}
\text{r}(l(X,n), X). \\
\text{r}(G,E1) :- \\
\quad \text{a}(G1,G2,G), \\
\quad \text{r}(G1,f(P)), \\
\quad \text{r}(G2,E2), \\
\quad P(E2,E1). \\
\text{r}(G,E1) :- \\
\quad \text{a}(G1,G2,G), \\
\quad \text{r}(G1,E2), \\
\quad \text{r}(G2,b(P)), \\
\quad P(E2,E1).
\end{align*}
\]

The rules are the only place in the program, where we use higher-order predicates.

The following predicate assigns a meaning to English text:

\[
\begin{align*}
\text{m}(T,C) :- \\
\quad \text{map_lex}(T,L), \\
\quad \text{r}(L,e(C)).
\end{align*}
\]

Let us add an auxiliary predicate \textit{work} for testing purposes:

\[
\begin{align*}
\text{work}(C) :- \\
\quad \text{m}(t(w_{\text{the}}, t(w_{\text{tall}}, t(w_{\text{smiles}}, tnil)))}, C).
\end{align*}
\]

The \textit{map_lex} predicate is used for lexicon look-up of a list of words.

\[
\begin{align*}
\text{map_lex}(tnil,n). \\
\text{map_lex}(t(W,T), l(I,MT)) :- \\
\quad \text{lex}(W,I), \text{map_lex}(T,MT).
\end{align*}
\]

VI. CONCLUSION

We have presented an extension to Categorial Grammar, which allows constructing arbitrary semantics and enforcing arbitrary semantic restrictions in a very flexible manner. The functional composition and \( \beta \)-reduction are replaced with proof rule application. We have utilized it for constructing ontologico-algebraic meaning using ontological restrictions, dropping at the same time syntactic categories and syntactic restrictions. Finally, a proof-of-concept implementation has been given.

ACKNOWLEDGMENT

I am very grateful to my supervisor, prof. Jørgen Fischer Nilsson, for his most kind help.

REFERENCES


