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 include 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.

Throughout the paper we use classical Church’s type theory,
C, as presented in [2]. We use the convention that functional
types associate to the left, i.e. type γβα is the same as (γβ)α,
which is also sometimes denoted as α → (β → γ).

The work presented here takes place within the SIABO
project [3].

II. ONTOLOGICAL SEMANTICS

We wish to construct ontological semantics for a fragment of
English, following the approach in [4] and [5].

We can represent concepts appearing in a skeleton ontology
as constants of type α: Childα, Tallα, Runningα, . . . Those
concepts are to be thought of as sets, e.g. Childα is a set
of all imaginable children. Tallα is the set of all tall objects,
thus properties such as “tall” are introduced in the ontology
on the par with classes. Runningα is a set of all imaginable
actions of running, etc. Introducing nominalized verb forms
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 α, but in the current example
it could be replaced by ω, which is traditionally used for
representing sets. We will represent intersection of concepts
using constant ξαα, e.g. child ∩ tall will be represented as
ξαα Childα ∩ Tallα.

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

\[ \text{Sub}_\text{ααα} \text{Child}_α \text{Person}_α \]
\[ \text{Sub}_\text{ααα} \text{Person}_α \text{Physical}_α \]
\[ \text{Sub}_\text{ααα} \text{Person}_α \text{Animate}_α \]
\[ \text{Sub}_\text{ααα} \text{Running}_α \text{Action}_α \]

\[ \text{Sub}_\text{ααα} \] is a direct descendant relation, representing the lines
that are drawn in the Hasse diagrams. Isaαaa can be specified
as the (reflexive) transitive closure of Subαaa in the following way:

\[ \forall c_α [\text{Isa}_ααα c_α c_α] \]
\[ \forall c_α \forall a_α \forall p_α [\text{Isa}_ααα c_α a_α \supset \text{Sub}_ααα c_α p_α \land \text{Isa}_ααα p_α a_α] \]

We use constants of type ρ to represent roles (binary relations),
e.g. Agtρ, Locρ. We can use Peirce product [6] to create
compound concepts. We use constant :ααρ for that purpose,
e.g. agt : child will be represented as :ααρ Childα. 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:

\[ \text{running} \land \text{agt} : (\text{child} \cap \text{tall}) \]

This can be represented in type theory as a wffα:

\[ \cap_α \text{Running}_α [\text{ααρ} \text{Agt}_ρ [\cap_α \text{Child}_α \text{Tall}_α]] \]

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

\[ \forall x_α \forall y_α \forall z_α [[\text{Isa}_ααα [\cap_α α x_α y_α] z_α] \supset [\text{Isa}_ααα x_α z_α]] \]
\[ \forall x_α \forall y_α \forall z_α [[\text{Isa}_ααα [\cap_α α x_α y_α] z_α] \supset [\text{Isa}_ααα y_α z_α]] \]

III. MODIFIED CATEGORIAL GRAMMAR

We introduce type ω. Constants of this type represent words
in English, e.g. Wkidω, Wtallω, Wrunsω. Notice that we
use different names for constants denoting words and those
representing concepts, e.g. Wtallω and Tallω. 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:

\[ \text{κ} \]

\[ \text{κ} \]

\[ \text{κ} \]
Lexical entries are represented using the predicate constant \( L_{\text{exo}} \). The lexicon consists of the set of factual clauses of the form:

\[
L_{\text{exo}} \cdot W_{\text{kid}}[E_{\text{ex}} \cdot \text{Child}_a]
\]

The type constructor \( E_{\text{ex}} \) is used in the lexical entry in case the meaning of a word is some “fixed” ontological concept. The other constructors, \( F_e \) and \( B_k \) are reminiscent of the Categorial Grammar’s backward and forward slashes, respectively. A lexical entry for word “tall” might look as follows:

\[
L_{\text{exo}} \cdot W_{\text{tall}}[F_k(E_{\text{ex}}) \cdot P_{\text{tall}}]
\]

In the above, \( P_{\text{tall}} \) is a predicate, which might be defined in the following way:

\[
\forall c_a \cdot [P_{\text{tall}}[E_{\text{ex}} \cdot c_a] \cdot E_{\text{ex}} \cdot \text{Physical}_a] \approxeq [\text{Isa}_{\text{ex}} \cdot c_a \cdot \text{Physical}_a]
\]

The novelty in our approach is the use of predicates in meanings of words. In standard Categorial Grammar, lambda restrictions on the semantics. Our approach can reject this sen-

- 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.:

\[
K \not\vdash \exists x_a\cdot [P_{\text{smiles}}[E_{\text{ex}} \cdot \text{Vitamin}_x] \cdot x_a]
\]

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:

\[
K \vdash \exists x_a\cdot [P_{\text{smiles}}[E_{\text{ex}} \cdot \text{Child}_x] \cdot x_a]
\]

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:

\[
E_{\text{ex}} \cdot \text{Smiling}_a \cdot \text{Agt}_p \cdot \text{Child}_a
\]

In other words:

\[
K \vdash [P_{\text{smiles}}[E_{\text{ex}} \cdot \text{Child}_x] \cdot E_{\text{ex}} \cdot \text{Smiling}_a \cdot \text{Agt}_p \cdot \text{Child}_a]
\]

IV. ELIMINATION RULES

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

\[
a \Rightarrow \Phi_1 : A / B \quad b \Rightarrow \Phi_2 : B
\]

and the backward rule:

\[
a \Rightarrow \Phi_2 : B / A \quad b \Rightarrow \Phi_1 : A
\]

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:

\[
o_a \Rightarrow \Phi_k : p_{\text{ex}} \quad b_0 \Rightarrow \Phi_k : k\bar{c}
\]

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

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

\[
\forall x_a \cdot [\forall y_a \cdot [\forall t_\theta \cdot \text{Agt}_\theta : \text{Child}_x] \cdot \text{Agt}_\theta : \text{Child}_x] \cdot \text{Agt}_\theta : \text{Child}_x] \cdot \text{Agt}_\theta : \text{Child}_x]
\]

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on July 30,2010 at 09:55:46 UTC from IEEE Xplore. Restrictions apply.
The above use of $A_{abb}$ 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 $C$:

$$\forall a \forall b \forall g \forall j \forall k \forall \theta \forall p_{onk}; R_{onk}g = k \sigma [A_{abb}a b g b g]$$

Similarly, the backward elimination rule:

$$\forall a \forall b \forall g \forall j \forall k \forall \theta \forall p_{onk}; R_{onk}g = k \sigma [A_{abb}a b g b g]$$

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:

$$\forall k \sigma [R_{onk}L_\emptyset_k k_\emptyset_N_0]$$

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 $T_{nil}$ and a list constructor by a constant $T_{\varepsilon\varepsilon}$.

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 ‘$\forall$’ and ‘$\wedge$’ with ‘‘$\sigma$’’ and ‘‘$\tau$’’, 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;
prt; 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).
```

The predicates included in lexical entries take the following form in Mercury:

```mercury
p_tall(e(C), e(i(tall, C))):-
isa(C, physical).
p_runs(e(C), e(i(running, p(agt, C))):-
isa(C, animate).
p_smiles(e(C), e(i(smiling, p(agt, C))):-
isa(C, animate).
p_the(e(I), e(I)).
...
```
The lexicon, translated to Mercury:

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

\[
\text{...}
\]

The definition of non-empty list appending:

\[
a(l(X,n), l(Y,T), l(X,l(Y,T))).
\]

\[
a(l(H,L), M, l(H,T)) :- a(L,M,T).
\]

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

\[
r(l(X,n), X).
\]

\[
r(G,E1) :- \\
a(G1,G2,G), \\
r(G1,f(P)), \\
r(G2,E2), \\
P(E2,E1).
\]

\[
r(G,E1) :- \\
a(G1,G2,G), \\
r(G1,E2), \\
r(G2,b(P)), \\
P(E2,E1).
\]

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

The following predicate assigns a meaning to English text:

\[
m(T,C) :- \\
\text{map_lex}(T,L), \\
r(L,e(C)).
\]

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

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

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

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

Notice that the program does not require higher order unification, except for the simplest case where a variable is bound to a predefined predicate and such a predicate is called. Most logic programming languages provide a facility for such behaviour. We could change the syntax slightly and turn our code into a valid Prolog program by utilizing the \text{call} library predicate.

Notice that the program is a direct implementation of the theory, hence quite inefficient. Instead of the presented top-down approach, we could derive the semantics in the bottom-up manner in order to avoid unneeded search.

For the sample query:

\[
? \text{work}(C).
\]

We get the following (exactly one) result:

\[
C = i(\text{smiling}, p(\text{agt}, i(tall, child)))
\]

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