

# Using Higher-Order Logic Programming for Semantic Interpretation of Coordinate Constructs

Seth Kulick

University of Pennsylvania  
Computer and Information Science  
200 South 33rd Street  
Philadelphia, PA 19104-6389 USA  
skulick@linc.cis.upenn.edu

## Abstract

Many theories of semantic interpretation use  $\lambda$ -term manipulation to compositionally compute the meaning of a sentence. These theories are usually implemented in a language such as Prolog that can simulate  $\lambda$ -term operations with first-order unification. However, for some interesting cases, such as a Combinatory Categorical Grammar account of coordination constructs, this can only be done by obscuring the underlying linguistic theory with the “tricks” needed for implementation. This paper shows how the use of abstract syntax permitted by higher-order logic programming allows an elegant implementation of the semantics of Combinatory Categorical Grammar, including its handling of coordination constructs.

## 1 Introduction

Many theories of semantic interpretation use  $\lambda$ -term manipulation to compositionally compute the meaning of a sentence. These theories are usually implemented in a language such as Prolog that can simulate  $\lambda$ -term operations with first-order unification. However, there are cases in which this can only be done by obscuring the underlying linguistic theory with the “tricks” needed for implementation. For example, Combinatory Categorical Grammar (CCG) (Steedman, 1990) is a theory of syntax and semantic interpretation that has the attractive characteristic of handling many coordination constructs that other theories cannot. While many aspects of CCG semantics can be reasonably simulated in first-order unification, the simulation breaks down on some of the most interesting cases that CCG can theoretically handle. The problem in general, and for CCG in particular, is that the implementation language does not have sufficient expressive power to allow a more direct encoding. The solution given in this paper is to show how advances in logic programming allow the implementation of semantic theories in a

very direct and natural way, using CCG as a case study.

We begin by briefly illustrating why first-order unification is inadequate for some coordination constructs, and then review two proposed solutions. The sentence in (1a) usually has the logical form (LF) in (1b).

(1a) John and Bill run.

(1b) (and (run John) (run Bill))

CCG is one of several theories in which (1b) gets derived by raising *John* to be the LF  $\lambda P.(P \text{ john})$ , where  $P$  is a predicate that takes a NP as an argument to return a sentence. Likewise, *Bill* gets the LF  $\lambda P.(P \text{ bill})$ , and coordination results in the following LF for *John and Bill*:

(2)  $\lambda P.(\text{and } (P \text{ john}) (P \text{ bill}))$

When (2) is applied to the predicate, (1b) will result after  $\beta$ -reduction. However, under first-order unification, this needs to be simulated by having the variable  $x$  in  $\lambda x.run(x)$  unify both with Bill and John, and this is not possible. See (Jowsey, 1990) and (Moore, 1989) for a thorough discussion.

(Moore, 1989) suggests that the way to overcome this problem is to use explicit  $\lambda$ -terms and encode  $\beta$ -reduction to perform the needed reduction. For example, the logical form in (3) would be produced, where  $\mathbf{X}\backslash run(\mathbf{X})$  is the representation of  $\lambda x.run(x)$ .

(3) and(apply( $\mathbf{X}\backslash run(\mathbf{X})$ , john),

apply( $\mathbf{X}\backslash run(\mathbf{X})$ , bill))

This would then be reduced by the clauses for `apply` to result in (1b). For this small example, writing such an `apply` predicate is not difficult. However, as the semantic terms become more complex, it is no trivial matter to write  $\beta$ -reduction that will correctly handle variable capture. Also, if at some point it was desired to determine if the semantic forms of two different sentences were the same, a predicate would be needed to compare two lambda forms for  $\alpha$ -equivalence, which again is not a simple task. Essentially, the logic variable  $\mathbf{X}$  is meant to be interpreted as a bound variable, which requires an additional layer of programming.

(Park, 1992) proposes a solution within first-order unification that can handle not only sentence (1a), but also more complex examples with determiners. The method used is to introduce spurious bindings that subsequently get removed. For example, the semantics of (4a) would be (4b), which would then get simplified to (4c).

- (4a) A farmer and every senator talk
- (4b) `exists(X1, farmer(X1)
&(exists(X2, (X2=X1)&talk(X2))))
&forall(X3, senator(X3)
=>(exists(X2, (X2=X3)&talk(X2))))`
- (4c) `exists(X1, farmer(X1)&talk(X1))
&forall(X3, senator(X3)=>talk(X3))`

While this pushes first-order unification beyond what it had been previously shown capable of, there are two disadvantages to this technique: (1) For every possible category that can be conjoined, a separate lexical entry for *and* is required, and (2) As the conjoinable categories become more complex, the *and* entries become correspondingly more complex and greatly obscure the theoretical background of the grammar formalism.

The fundamental problem in both cases is that the concept of free and bound occurrences of variables is not supported by Prolog, but instead needs to be implemented by additional programming. While theoretically possible, it becomes quite problematic to actually implement. The solution given in this paper is to use a higher-order logic programming language,  $\lambda$ Prolog, that already implements these concepts, called “abstract syntax” in (Miller, 1991) and “higher-order abstract syntax” in (Pfenning and Elliot, 1988). This allows a natural and elegant implementation of the grammatical theory, with only one lexical entry for *and*. This paper is meant to be viewed as furthering the exploration of the utility of higher-order logic programming for computational linguistics - see, for example, (Miller & Nadathur, 1986), (Pareschi, 1989), and (Pereira, 1990).

## 2 CCG

CCG is a grammatical formalism in which there is a one-to-one correspondence between the rules of composition<sup>1</sup> at the level of syntax and logical form. Each word is (perhaps ambiguously) assigned a category and LF, and when the syntactical operations assign a new category to a constituent, the corresponding semantic operations produce a new LF for that constituent as well. The CCG rules shown in Figure 1 are implemented in the system described

<sup>1</sup>In the general sense, not specifically the CCG rule for function composition.

Function Application (>):  
 $X/Y : F \quad Y : y \Rightarrow X : Fy$

Function Application (<):  
 $Y : y \quad X \backslash Y : F \Rightarrow X : Fy$

Function Composition (> B):  
 $X/Y : F \quad Y/Z : G \Rightarrow X/Z : \lambda x.F(Gx)$

Function Composition (< B):  
 $Y \backslash Z : G \quad X \backslash Y : F \Rightarrow X \backslash Z : \lambda x.F(Gx)$

Type Raising (> T):  
 $np : x \Rightarrow s/(s \backslash np) : \lambda F.Fx$

Type Raising (< T):  
 $np : x \Rightarrow s \backslash (s/np) : \lambda F.Fx$

Figure 1: CCG rules

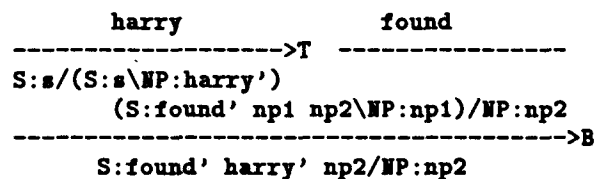


Figure 2: CCG derivation of *harry found* simulated by first-order unification

in this paper.<sup>2 3</sup> Each of the three operations have both a forward and backward variant.

As an illustration of how the semantic rules can be simulated in first-order unification, consider the derivation of the constituent *harry found*, where *harry* has the category *np* with LF *harry'* and *found* is a transitive verb of category  $(s \backslash np)/np$  with LF

(5)  $\lambda object. \lambda subject. (found' subject object)$

In the CCG formalism, the derivation is as follows: *harry* gets raised with the  $>T$  rule, and then forward composed by the  $>B$  rule with *found*, and the result is a category of type *s/np* with LF  $\lambda x.(found' harry' x)$ . In section 3 it will be seen how the use of abstract syntax allows this to be expressed directly. In first-order unification, it is simulated as shown in Figure 2.<sup>4</sup>

The final CCG rule to be considered is the coordination rule that specifies that only like categories can coordinate:

<sup>2</sup>The type-raising rules shown are actually a simplification of what has been implemented. In order to handle determiners, a system similar to NP-complement categories as discussed in (Dowty, 1988) is used. Although a worthwhile further demonstration of the use of abstract syntax, it has been left out of this paper for space reasons.

<sup>3</sup>The  $\backslash$  for a backward-looking category should not be confused with the  $\backslash$  for  $\lambda$ -abstraction.

<sup>4</sup>example adapted from (Steedman, 1990, p. 220).

(6)  $\mathbf{X} \text{ conj } \mathbf{X} \Rightarrow \mathbf{X}$

This is actually a schema for a family of rules, collectively called “generalized coordination”, since the semantic rule is different for each case.<sup>5</sup> For example, if  $\mathbf{X}$  is a unary function, then the semantic rule is (7a), and if the functions have two arguments, then the rule is (7b).<sup>6</sup>

(7a)  $\Phi FGH = \lambda x.F(Gx)(Hx)$

(7b)  $\Phi^2 FGH = \lambda x.\lambda y.F(Gxy)(Hxy)$

For example, when processing (1a), rule (7a) would be used with:

- $F = \lambda x.\lambda y.(\text{and}' x y)$
- $G = \lambda P.(P \text{ john}')$
- $H = \lambda P.(P \text{ bill}')$

with the result

$\phi FGH = \lambda x.(\text{and}' (x \text{ john}') (x \text{ bill}'))$

which is  $\alpha$ -equivalent to (2).

### 3 $\lambda$ PROLOG and Abstract Syntax

$\lambda$ Prolog is a logic programming language based on *higher-order hereditary Harrop formulae* (Miller et al., 1991). It differs from Prolog in that first-order terms and unification are replaced with simply-typed  $\lambda$ -terms and higher-order unification<sup>7</sup>, respectively. It also permits universal quantification and implication in the goals of clauses. The crucial aspect for this paper is that together these features permits the usage of abstract syntax to express the logical forms terms computed by CCG. The built-in  $\lambda$ -term manipulation is used as a “meta-language” in which the “object-language” of CCG logical forms is expressed, and variables in the object-language are mapped to variables in the meta-language.

The  $\lambda$ Prolog code fragment shown in Figure 3 declares how the CCG logical forms are represented. Each CCG LF is represented as an untyped  $\lambda$ -term, namely type  $\mathbf{tm}$ . **abs** represents object-level abstraction  $\lambda x.M$  by the meta-level expression (**abs**  $\mathbf{N}$ ),

<sup>5</sup>It is not established if this schema should actually produce an unbounded family of rules. See (Weir, 1988) and (Weir and Joshi, 1988) for a discussion of the implications for automata-theoretic power of generalized coordination and composition, and (Gazdar, 1988) for linguistic arguments that languages like Dutch may require this power, and (Steedman, 1990) for some further discussion of the issue. In this paper we use the generalized rule to illustrate the elegance of the representation, but it is an easy change to implement a bounded coordination rule.

<sup>6</sup>The  $\Phi$  notation is used because of the combinatory logic background of CCG. See (Steedman, 1990) for details.

<sup>7</sup>defined as the unification of simply typed  $\lambda$ -terms, modulo  $\beta\eta$  conversion.

<b>kind</b>	<b>tm</b>	<b>type.</b>
<b>type</b>	<b>abs</b>	$(\mathbf{tm} \rightarrow \mathbf{tm}) \rightarrow \mathbf{tm}.$
<b>type</b>	<b>app</b>	$\mathbf{tm} \rightarrow \mathbf{tm} \rightarrow \mathbf{tm}.$
<b>type</b>	<b>forall</b>	$(\mathbf{tm} \rightarrow \mathbf{tm}) \rightarrow \mathbf{tm}.$
<b>type</b>	<b>exists</b>	$(\mathbf{tm} \rightarrow \mathbf{tm}) \rightarrow \mathbf{tm}.$
<b>type</b>	<b>&gt;&gt;</b>	$\mathbf{tm} \rightarrow \mathbf{tm} \rightarrow \mathbf{tm}.$
<b>type</b>	<b>&amp;&amp;</b>	$\mathbf{tm} \rightarrow \mathbf{tm} \rightarrow \mathbf{tm}.$

Figure 3: Declarations for  $\lambda$ Prolog representation of CCG logical forms

where  $\mathbf{N}$  is a meta-level function of type  $\mathbf{tm} \rightarrow \mathbf{tm}$ . A meta-level  $\lambda$ -abstraction  $\lambda y.P$  is written  $\mathbf{y} \backslash \mathbf{P}$ .<sup>8</sup>

Thus, if **walked'** has type  $\mathbf{tm} \rightarrow \mathbf{tm}$ , then  $\mathbf{y} \backslash (\text{walked}' \mathbf{y})$  is a  $\lambda$ Prolog (meta-level) function with type  $\mathbf{tm} \rightarrow \mathbf{tm}$ , and (**abs**  $\mathbf{y} \backslash (\text{walked}' \mathbf{y})$ ) is the object-level representation, with type  $\mathbf{tm}$ . The LF for *found* shown in (5) would be represented as (**abs** **obj** \ ((**abs** **sub** \ (**found'** **sub** **obj**))). **app** encodes application, and so in the derivation of *harry found*, the type-raised *harry* has the  $\lambda$ Prolog value (**abs** **p** \ ((**app** **p** **harry'**)).<sup>9</sup>

The second part of Figure 3 shows declares how quantifiers are represented, which are required since the sentences to be processed may have determiners. **forall** and **exists** are encoded similarly to abstraction, in that they take a functional argument and so object-level binding of variables by quantifiers is handled by meta-level  $\lambda$ -abstraction. **>>** and **&&** are simple constructors for implication and conjunction, to be used with **forall** and **exists** respectively, in the typical manner (Pereira and Shieber, 1987). For example, the sentence *every man found a bone* has as a possible LF (8a), with the  $\lambda$ Prolog representation (8b)<sup>10</sup>:

<sup>8</sup>This is the same syntax for  $\lambda$ -abstraction as in (3). (Moore, 1989) in fact borrows the notation for  $\lambda$ -abstraction from  $\lambda$ Prolog. The difference, of course, is that here the abstraction is a meta-level, built-in construct, while in (3) the interpretation is dependent on an extra layer of programming. Bound variables in  $\lambda$ Prolog can be either upper or lower case, since they are not logic variables, and will be written in lower case in this paper.

<sup>9</sup>It is possible to represent the logical forms at the object-level without using **abs** and **app**, so that *harry* could be simply  $\mathbf{p} \backslash (\mathbf{p} \text{ harry}')$ . The original implementation of this system was in fact done in this manner. Space prohibits a full explanation, but essentially the fact that  $\lambda$ Prolog is a typed language leads to a good deal of formal clutter if this method is used.

<sup>10</sup>The LF for the determiner has the form of a Montaguevian generalized quantifier, giving rise to one fully scoped logical form for the sentence. It should be stressed that this particular kind of LF is assumed here purely for the sake of illustration, to make the point that composition at the level of derivation and LF are one-to-one. Section 4 contains an example for which such a



```

type    coord
      cat -> tm -> tm -> tm -> o.

coord (fs A B) (abs R) (abs S) (abs T) :-
  pi x \ (coord B (R x) (S x) (T x)).

coord (bs A B) (abs R) (abs S) (abs T) :-
  pi x \ (coord B (R x) (S x) (T x)).

coord B R S (and' R S) :- atomic-type B.

```

Figure 6: Implementation of coordination

to be constructors for forward and backward slash. For example, the CCG category for a transitive verb ( $s \backslash np / np$ ) would be represented as  $(fs \ np \ (bs \ np \ s))$ . Also, the predicate `atomic-type` is declared to be true for the four atomic categories. This will be used in the implementation of coordination as a test for termination of the recursion.

The implementation of coordination crucially uses the capability of  $\lambda$ Prolog for universal quantification in the goal of a clause.  $\pi$  is the meta-level operator for  $\forall$ , and  $\forall x.M$  is written as  $\pi \ x \ M$ . The operational semantics for  $\lambda$ Prolog state that  $\pi \ x \ G$  is provable if and only if  $[c/x]G$  is provable, where  $c$  is a new variable of the same type as  $x$  that does not otherwise occur in the current signature. In other words,  $c$  is a scoped constant and the current signature gets expanded with  $c$  for the proof of  $[c/x]G$ . Since  $c$  is meant to be treated as a generic placeholder for any arbitrary  $x$  of the proper type,  $c$  must not appear in any terms instantiated for logic variables during the proof of  $[c/x]G$ . The significance of this restriction will be illustrated shortly.

The code for coordination is shown in Figure 6. The four arguments to `coord` are a category and three terms that are the object-level LF representations of constituents of that category. The last argument will result from the coordination of the second and third arguments. Consider again the earlier problematic example (1a) of coordination. Recall that after *john* is type-raised, its LF will be  $(abs \ p \ (app \ p \ john'))$  and similarly for *bill*. They will both have the category  $(fs \ (bs \ np \ s))$ . Thus, to obtain the LF for *John and Bill*, the following query would be made:

```

?- coord (fs (bs np s) s)
      (abs p \ (app p john'))
      (abs p \ (app p bill'))
      M.

```

This will match with the first clause for `coord`, with

- $A$  instantiated to  $(bs \ np \ s)$
- $B$  to  $s$
- $R$  to  $(p \ (app \ p \ john'))$
- $S$  to  $(p \ (app \ p \ bill'))$

- and  $T$  a logic variable waiting instantiation.

Then, after the meta-level  $\beta$ -reduction using the new scoped constant  $c$ , the following goal is called:

```

?- coord s (app c john') (app c bill') M.

```

where  $M = (T \ c)$ . Since  $s$  is an atomic type, the third `coord` clause matches with

- $B$  instantiated to  $s$
- $R$  to  $(app \ c \ john')$
- $S$  to  $(app \ c \ bill')$
- $M$  to  $(and' \ (app \ c \ john') \ (app \ c \ bill'))$

Since  $M = (T \ c)$ , higher-order unification is used by  $\lambda$ Prolog to instantiate  $T$  by extracting  $c$  from  $M$  with the result

```

T = x \ (and' (app x john') (app x bill'))

```

and so  $M$  from the original query is

```

(abs x \ (and' (app x john') (app x bill')))

```

Note that since  $c$  is a scoped constant arising from the proof of an universal quantification, the instantiation

```

T = x \ (and' (app c john') (app x bill'))

```

is prohibited, along with the other extractions that do not remove  $c$  from the body of the abstraction.

This use of universal quantification to extract out  $c$  from a term containing  $c$  in this case gives the same result as a direct implementation of the rule for coordination of unary functions (7a) would. However, this same process of recursive descent via scoped constants will work for any member of the `conj` rule family. For example, the following query

```

?- coord
      (fs np (bs np s))
      (abs obj \ (abs sub \ (like' sub obj)))
      (abs obj \ (abs sub \ (hate' sub obj)))
      M.
M = (abs x \
      (abs x1 \
        (and' (like' x1 x)
              (hate' x1 x)))).

```

corresponds to rule (7b). Note also that the use of the same bound variable names `obj` and `sub` causes no difficulty since the use of scoped-constants, meta-level  $\beta$ -reduction, and higher-order unification is used to access and manipulate the inner terms. Also, whereas (Park, 1992) requires careful consideration of handling of determiners with coordination, here such sentences are handled just like any others. For example, the sentence *Mary gave every dog a bone and some policeman a flower* results in the LF 12.

<sup>12</sup>This is a case in which the particular LF assumed here fails to yield another available scoping. See footnote 10.

```
(and'
 (exists x\((bone' x) &&
  (forall x1\((dog' x1)
   >> (gave' mary' x x1))))))
 (exists x\((flower' x) &&
  (exists x1\((policeman' x1)
   && (gave' mary' x x1))))))
```

Thus, “generalized coordination”, instead of being a family of separate rules, can be expressed as a single rule on recursive descent through logical forms. (Steedman, 1990) also discusses “generalized composition”, and it may well be that a similar implementation is possible for that family of rules as well.

## 5 Conclusion

We have shown how higher-order logic programming can be used to elegantly implement the semantic theory of CCG, including the previously difficult case of its handling of coordination constructs. The techniques used here should allow similar advantages for a variety of such theories.

An argument can be made that the approach taken here relies on a formalism that entails implementation issues that are more difficult than for the other solutions and inherently not as efficient. However, the implementation issues, although more complex, are also well-understood and it can be expected that future work will bring further improvements. For example, it is a straightforward matter to transform the  $\lambda$ Prolog code into a logic called  $L_\lambda$  (Miller, 1990) which requires only a restricted form of unification that is decidable in linear time and space. Also, the declarative nature of  $\lambda$ Prolog programs opens up the possibility for applications of program transformations such as partial evaluation.

## 6 Acknowledgments

This work is supported by ARO grant DAAL03-89-0031, DARPA grant N00014-90-J-1863, and ARO grant DAAH04-94-G-0426. I would like to thank Aravind Joshi, Dale Miller, Jong Park, and Mark Steedman for valuable discussions and comments on earlier drafts.

## References

David Dowty. 1988. Type raising, functional composition, and non-constituent conjunction. In Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler, editors, *Categorial Grammars and Natural Language Structures*. Reidel, Dordrecht, pages 153-198.

Gerald Gazdar. 1988. Applicability of indexed grammars to natural languages. In U. Reyle and C. Rohrer, editors, *Natural language parsing and linguistic theories*. Reidel, Dordrecht, pages 69-94.

Jerry R. Hobbs and Stuart M. Shieber. 1987. An algorithm for generating quantifier scopings. *Computational Linguistics*, 13:47-63.

Einar Jowsey. 1990. *Constraining Montague Grammar for Computational Applications*. PhD thesis, University of Edinburgh.

Dale Miller. 1990. A logic programming language with lambda abstraction, function variables and simple unification. In P. Schroeder-Heister, editor, *Extensions of Logic Programming*, Lecture Notes in Artificial Intelligence, Springer-Verlag, 1990.

Dale Miller. 1991. Abstract syntax and logic programming. In *Proceedings of the Second Russian Conference on Logic Programming*, September 1991.

Dale Miller and Gopalan Nadathur. 1986. Some uses of higher-order logic in computational linguistics. In *24th Annual Meeting of the Association for Computational Linguistics*, pages 247-255.

Dale Miller, Gopalan Nadathur, Frank Pfenning, Andre Scedrov. 1991. Uniform proofs as a foundation for logic programming. In *Annals of Pure and Applied Logic*, 51:125-157.

Robert C. Moore. 1989. Unification-based semantic interpretation. In *27th Annual Meeting of the Association for Computational Linguistics*, pages 33-41.

Remo Pareschi. 1989. *Type-Driven Natural Language Analysis*. PhD thesis, University of Edinburgh.

Jong C. Park. 1992. A unification-based semantic interpretation for coordinate constructs. In *30th Annual Meeting of the Association for Computational Linguistics*, pages 209-215.

Jong C. Park. 1995. Quantifier scope and constituency. In *33rd Annual Meeting of the Association for Computational Linguistics* (this volume).

Barbara Partee and Mats Rooth. 1983. Generalized conjunction and type ambiguity. In Rainer Bauerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, Use, and Interpretation of Language*. W. de Gruyter, Berlin, pages 361-383.

Fernando C.N. Pereira. 1990. Semantic interpretation as higher-order deduction. In Jan van Eijck, editor, *Logics in AI: European Workshop JELIA '90*, Lecture Notes in Artificial Intelligence number 478, pages 78-96. Springer-Verlag, Berlin, Germany.

Fernando C.N. Pereira and Stuart M. Shieber. 1987. *Prolog and Natural-Language Analysis*. Number 10 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, California,

1985. Distributed by the University of Chicago Press.
- Frank Pfenning and Conal Elliot. 1988. Higher-order abstract syntax. In *Proceedings of the ACM-SIGPLAN Conference on Programming Language Design and Implementation*, 1988.
- Mark J. Steedman. 1990. Gapping as constituent coordination. In *Linguistics and Philosophy* 13, pages 207–263
- David Weir. 1988. *Characterizing Mildly Context-sensitive Grammar Formalism*. CIS-88-74, PhD thesis, University of Pennsylvania.
- David Weir and Aravind Joshi. 1988. Combinatory categorial grammars: generative power and relation to linear CF rewriting systems. In *26th Annual Meeting of the Association for Computational Linguistics*, pages 278–285.