Covert without overt: QR for movementless parsing frameworks

Asad Sayeed and Vera Demberg, MMCI Cluster of Excellence, Saarland University

After a two-decade period of relative absence, rich linguistic representation is returning to engineering applications, particularly incremental parsing and spoken dialogue systems. However, for reasons of structural ambiguity avoidance and representational convenience, movementless formalisms dominate the landscape, maintaining the on-going disconnection with potentially useful generalizations from syntactic and semantic theory. We suggest that one important area in which these generalizations can be restored to engineering-oriented formalisms is in the representation of scope ambiguity.

We propose a new way of representing quantifier scope ambiguities and their resolutions that is derived entirely from neo-Davidsonian semantic expressions. We use a limited form of movement and simple restrictions thereon in order to implement covert movement without imposing aspects of movement-based formalisms on movementless incremental parsers. Implementing only covert movement therefore allows the parsing algorithm to proceed unchanged.

The problem of scope ambiguity in the output representation has recently become a matter of debate (e.g. Joshi et al., 2008) in the parsing literature. Many of the current solutions proposed for Tree-Adjoining Grammars (TAGs), combinatory categorial grammars (CCGs), and so on have the property that the ambiguous structures are represented in the lexicon along with the scope items as logical forms. Ruys and Winter (2010) compare these logical approaches to quantifier raising (QR; May, 1985) approaches from theoretical syntax and find that the logical approaches do not capture some of the island generalizations that QR approaches do. Attempts to handle scope ambiguity through logical operations such as type-raising (Champollion, 2011) also require the bottom-up revision of semantic structures, which is incrementality-unfriendly.

A controversial example comes from Romero and Kallmeyer (2005), for which we give a neo-Davidsonian representation:

(1) a. Two politicians spied on someone from every city.

b. $2x_1\mathsf{Politician}(x_1)\&\exists x_2\mathsf{Person}(x_2)\&\forall x_3\mathsf{City}(x_3)\&\mathsf{From}(x_2,x_3)\rightarrow$
$\exists e\mathsf{Spyer}(e,x_1)\&\mathsf{SpiedUpon}(e,x_2)\&\mathsf{Spy}(e)$

Neo-Davidsonian semantics (Parsons, 1990) uses existentially-quantified event variables (in this example, e) to connect verb predicates with their arguments, and to assign theta roles to arguments. Conjunctions are used to produce a flat semantic representation, which simplifies the inferences required to update the semantic expression during incremental parsing and allows adjuncts to be integrated into the expression without already knowing their hosts.

In this example, Romero and Kallmeyer claim that valid scope orders include $2>\exists>\forall$, $2>\forall>\exists$, $\forall>\exists>2$, and $\exists>\forall,2$, where $\forall$ is in the restrictor of $\exists$. However, $\forall>2>\exists$ is excluded. Within the TAG framework, Joshi et al. accommodate these readings using multi-component trees with ambiguous attachments in the syntax. During an incremental parse, commitments may have been made that would thus have to be undone in an *ad hoc* manner in order to accommodate multiple readings.

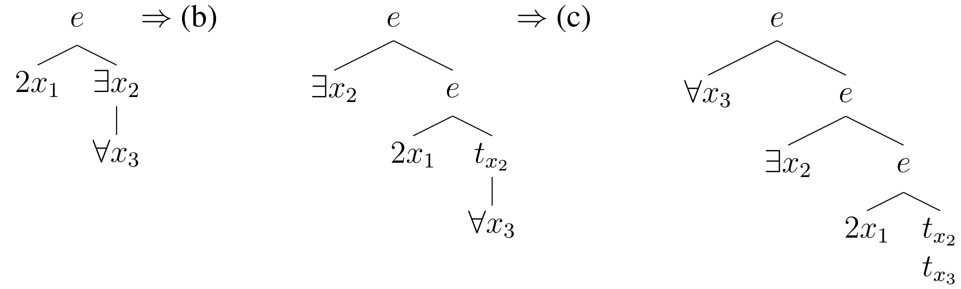
We reconcile these readings within a covert-only movement-based framework using a type of structure we call a variable scope tree (VST). Other approaches to scope ambiguity that involve the construction of a graph structure alongside the parse tree include Koller et al. (2003), which is fully lexicalized and non-incremental. We construct VSTs from neo-Davidsonian expressions using the following algorithm:

- Find the root event e corresponding to the outermost/matrix clause.
- Find every predicate in the expression mentioning e.
- For every variable v other than e mentioned in these predicates,
  - Make v a child of e in the VST. Label v's node with v's quantifier.

– Follow this procedure recursively for v. Ignore variables already met.

This constructs a spanning tree over variables in the semantic expression (2a). This can be applied at every incremental step or the tree can be grown dynamically.

(2) (a)



We can permit $2 > \forall > \exists$ by declaring that sisters can scope over each other. (As this is happening in covert syntax, asymmetry is not required for e.g. a linearization algorithm. Such asymmetries can be introduced in future work as needed without losing the substance of this proposal.) We can then permit $\exists > 2, \forall$ and $\forall > \exists > 2$ using a type of movement over this very limited structure (2b,c). We then propose the following definition of VST-movement to restrict the possible movements of quantified variables:

(3) **VST-move**: a variable v and its quantifier are permitted to move to e iff e is the immediate parent of v, or is the ancestor of v only via traces (t). A new copy of e is created as a parent to the original e and v.

We further stipulate that no event node may have only traces as children, to prevent infinite movement. This excludes $\forall > 2 > \exists$, because $\forall x_3$ cannot move as long as $\exists x_2$ is its parent. Event variables thus become analogous to CP in accounts with overt movement, and restrictions on covert movement are now defined in terms of proximity to the event. This gives us a principled way to understand the difference between these sentences from Kallmeyer and Romero.

(4) a. A student said you met every professor. (inverse scope forbidden)

b. A student wants (you) to meet every professor. (inverse scope permitted)

In the first example, inverse scope is forbidden by the presence of two fully-fledged events. *Every professor* would be represented in the VST as the child of a lower event. However, in the control verb condition, the lower verb is not a fully-fledged event with a separate conjunct. This insight is reflected in structures used by Asudeh and Toivonen (2012) who nest the lower verb's event inside the representation of the control verb. Consequently, the movement of the lower event would not be blocked.

For *wh-in-situ* languages such as Chinese, quantifiers are not permitted to take inverse scope, but *wh*-items are. We can treat this as a parametric variation by building the VST based on question-bound variables rather than quantifier-bound ones. In conclusion, our VST analysis allows a principled unification of covert movement analyses for movementless formalisms that do not readily accommodate May-style generalizations.

A. Asudeh and I. Toivonen. (2012) Copy raising and perception.

L. Champollion. (2011) Quantification and negation in event semantics.

A. Joshi, L. Kallmeyer, and M. Romero. (2007) Flexible composition in LTAG: quantifier scope and inverse linking.

A. Koller, J. Niehren, and S. Thater. (2003) Bridging the gap between underspecification formalisms: hole semantics as dominance constraints.

R. May. (1985) Logical form: its structure and derivation.

T. Parsons. (1990) Events in the semantics of English.

M. Romero and L. Kallmeyer. (2005) Scope and situation binding in LTAG using semantic unification.

E. Ruys and Y. Winter. (2010) Scope ambiguities in formal syntax and semantics.