Scope in an incremental context

Asad Sayeed, Ph.D.
Department of Philosophy, Linguistics, and Theory of Science, University of Gothenburg
Box 200, 40530 Gothenburg, Sweden
asad.sayeed@gu.se
https://clasp.gu.se/about/people/asad-sayeed

Course type Advanced, Language and Computation

Abstract This is an interdisciplinary course intended to bring together students from psycholinguistic and computational backgrounds in order to explore the question of the interpretation of ambiguous scopes under conditions of incremental understanding. Scope ambiguities (quantifier, negation, etc.) are common in language, but human language users usually have very little trouble settling on an interpretation in context. However, the interaction of contextual information with incrementally-built formal structure is not well-understood for scope. This course will go through the evidence for theories of time course of scope ambiguity identification and resolution from a psycholinguistic perspective, the adaptation of well-researched formal approaches to the incremental context of scope, and the computational basis for integrating pragmatic knowledge into the resolution process.

1 Introduction

This advanced course provides students with the opportunity to learn about an issue of increasing importance at the interface between syntax, semantics, and pragmatics: the interaction of pragmatic knowledge with the formal representation of scope in actual processing. For example: in understanding sentences like "The caregiver comforted the children every night", at what point do language-users decide on the relative number caregivers and children, how helpful is formal representation in understanding this process, and where is background knowledge invoked?

In this course, we will focus attention on the resolution of scope ambiguity in "online" sentence processing. The current state of research into the incremental processing of scope-ambiguous sentences—currently, mostly focused at the quantifier level—leads to conflicting conclusions. One is that scope ambiguities are mostly resolved using world knowledge; another is that formal representations significantly bias and constrain how the resolution proceeds. The role of world-knowledge constrains the extent to which the parser uses reanalysis or underspecification strategies.

In exploring recent research into the relationship between this type of world knowledge and "algorithmic" processing, this ESSLLI course will expand student understanding of cross-linguistic and cross-cultural analyses of common intuitions about scope-relevant phenomena. This will enable students to think about this knowledge both in their theoretical work and with potential for application in socially-realistic human-machine interaction research.

This overall educational aim will be enhanced by pursuing three overall subtopics: a *psycholinguistic* subtopic, seeking to expand the evidence base for the factors underlying scope resolution in incremental sentence processing; a *theoretical* subtopic, seeking to explore representations for scope processing that are theoretically defensible and approximate psycholinguistic observations; and a *computational* subtopic, looking at resources and techniques from machine learning and natural language processing that could characterize scope ambiguity resolution in a manner that may lead to applications (especially in dialogue technology) and theoretical progress. The computational and psycholinguistic subtopics will have a special focus on the integration of world-knowledge into formal representation.

2 Principal content

2.1 Theoretical basis

Scope ambiguity, particularly quantifier scope ambiguity, has a long history of research from a "static", non-incremental perspective. This course will rely on the extensive review by Ruys and Winter (2011) of "classical" analysis and approaches to scope ambiguity resolution, including the different levels at which scope phenomena take place, syntax-based approaches such as quantifier raising, operations on logical derivations such as Cooper storage and quantifying-in, and so on.

We will then discuss foundational representational issues in scope ambiguity in the incremental context, looking at two conflicting overall frameworks: reanalysis (Fodor and Ferreira, 1998) and underspecification (Bos, 2004). Under reanalysis, language users employ a default formal semantic structure and revise it when the context invalidates that structure. With underspecification, the parser makes few or no presuppositions as to the relative size of the sets involved (in the case of quantifier scope ambiguities), until evidence from the sentence or context (including world-knowledge) allows the parser to draw a conclusion. These conflicting approaches represent a spectrum of possible scope resolution mechanisms (Karimi and Ferreira, 2016).

This course will also take a cross-linguistic perspective and cross-phenomenal perspective, considering the differences in scope ambiguities in e.g. non-Indo-European languages (Scontras et al., 2014), including as they relate to interactions between different types of operators, such as between negations and quantifiers (Lee, 2009).

2.2 Computational basis

We will consider computational questions such as the computational tractability of incremental representation (Kroch and Joshi, 1985; Johnson and Lappin, 1999). We will discuss the reasons why parsing formalisms rarely include space for QR-style covert operations, instead preferring to resolve scope ambiguities through mechanisms like parallelism (Boston et al., 2011). We will also focus on the tradeoffs created by parallelism or lack thereof. For example, formalisms that rely on a highly compositional semantics either omit covert operations (Morrill, 2012) or exploit parallelism; but since scopal items are common, this approach may still result in the proliferation of parallel parses, even controlling for world-knowledge. An example resolution would be the approach of Steedman (2011), which makes use of Skolem terms to encode scope ordering relations in the context of combinatory categorial grammar (CCG) to avoid explicit covert operations.

The course will also cover the matter of computational modeling, particularly through discussion of corpus development and recent applications of machine learning. Corpus development for detecting and resolving scope ambiguities is in a nascent phase, but there are already several efforts we will discuss during the course (AnderBois et al., 2012; Manshadi et al., 2011; Vincze et al., 2008). In addition, there have been very recent results in learning quantifier interpretations from multimodal data via machine learning that we will discuss in the course (Pezzelle et al., 2018)—which will lead to discussion during the course of computational approaches to modeling world-knowledge in the quantifier scope context.

2.3 Psycholinguistic basis

A psycholinguistic basis for the time course of scope disambiguation has a nascent literature with some recent results. This course will cover the most recent developments with an eye towards introducing students with minimal experimental background to the basic relevant concepts and techniques (e.g. self-paced reading, eye-tracking, ERPs). This is the aspect of the course that will most closely examine the interaction with pragmatics and world-knowledge.

One example of recent work in the psycholinguistics of incremental scope is Dwivedi (2013), who used the context-continuation paradigm of Kurtzman and MacDonald (1993) in an investigation of the degree of underspecification in grammar. The context is a sentence containing two quantificational expressions with ambiguous scope; e.g., "Every child climbed a tree". The continuation is a sentence following the context that confirms one set-cardinality interpretation; e.g., it begins either with "The tree" or "The trees". Dwivedi used this paradigm in self-paced reading studies to demonstrate that heuristic processing relying on world-knowledge tends to domininate the process of interpretation, considering that most English-users would find it improbable that every child was climbing one tree, and that the formal process is somewhat underspecified. More recent ERP work (Dwivedi and Gibson, 2017) also suggests a heuristic-first approach. Complementary work (Dotlačil and Brasoveanu, 2015), however, suggests that more intricate higher-order specification of scopes may be possible for more complicated contexts. This course will cover extensively the potential conflicts between these results and interpretations.

2.4 Integrating the streams

Towards the end, this course will bring together different theoretical, cognitive, and computational threads in discussion with the class in order to come up with a integrated picture of the role of event knowledge (Metusalem et al., 2012; Amsel et al., 2015) in scope ambiguity resolution. We will discuss how to update existing nascent attempts at reconciling the different "moving parts" in scope ambiguity resolution (Sayeed, 2016; Koller et al., 2003), with an eye towards understanding what makes it possible for humans to perform this task with overall ease, and how to make it possible for an artifical model to do so in an analogous manner.

References

Amsel, B. D., DeLong, K. A., and Kutas, M. (2015). Close, but no garlic: Perceptuomotor and event knowledge activation during language comprehension. *Journal of memory and language*, 82:118–132.

AnderBois, S., Brasoveanu, A., and Henderson, R. (2012). The pragmatics of quantifier scope: A corpus study. In *Proceedings of Sinn und Bedeutung*, volume 16, pages 15–28.

Bos, J. (2004). Computational semantics in discourse: Underspecification, resolution, and inference. *Journal of Logic, Language and Information*, 13(2):139–157.

Boston, M. F., Hale, J. T., Vasishth, S., and Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, 26(3):301–349.

Dotlačil, J. and Brasoveanu, A. (2015). The manner and time course of updating quantifier scope representations in discourse. *Language, Cognition and Neuroscience*, 30(3):305–323.

Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PloS one*, 8(11):e81461.

Dwivedi, V. D. and Gibson, R. M. (2017). An erp investigation of quantifier scope ambiguous sentences: Evidence for number in events. *Journal of Neurolinguistics*, 42:63 – 82.

Fodor, J. and Ferreira, F., editors (1998). *Reanalysis in sentence processing*. Springer.

Johnson, D. E. and Lappin, S. (1999). *Local constraints vs. economy*. University of Chicago Press, Chicago, USA.

Karimi, H. and Ferreira, F. (2016). Good-enough linguistic representations and online cognitive equilibrium in language processing. *Quarterly Journal of Experimental Psychology*, 69(5):1013–1040.

Koller, A., Niehren, J., and Thater, S. (2003). Bridging the gap between underspecification formalisms: Hole semantics as dominance constraints. In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1*, pages 195–202. Association for Computational Linguistics.

Kroch, A. S. and Joshi, A. K. (1985). The linguistic relevance of tree adjoining grammar. Technical report, University of Pennsylvania.

Kurtzman, H. S. and MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, 48(3):243–279.

Lee, S. (2009). *Interpreting scope ambiguity in first and second language processing: Universal quantifiers and negation*. PhD thesis, University of Hawaii at Manoa.

Manshadi, M., Allen, J., and Swift, M. (2011). A corpus of scope-disambiguated english text. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 141–146. Association for Computational Linguistics.

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., and Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of memory and language*, 66(4):545–567.

Morrill, G. (2012). *Type logical grammar: Categorial logic of signs*. Springer Science & Business Media.

Pezzelle, S., Sorodoc, I.-T., and Bernardi, R. (2018). Comparatives, quantifiers, proportions: a multi-task model for the learning of quantities from vision. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. Association for Computational Linguistics.

Ruys, E. and Winter, Y. (2011). Quantifier scope in formal linguistics. In *Handbook of philosophical logic*, pages 159–225. Springer.

Sayeed, A. (2016). Representing the effort in resolving ambiguous scope. In *Sinn und Bedeutung 20*, Tübingen, Germany.

Scontras, G., Polinsky, M., Tsai, E. C.-Y., and Mai, K. (2014). Chinese scope: an experimental investigation. In *Proceedings of Sinn und Bedeutung 18*, Vitoria-Gasteiz, Spain.

Steedman, M. (2011). *Taking Scope*. MIT Press, Boston, USA.

Vincze, V., Szarvas, G., Farkas, R., Móra, G., and Csirik, J. (2008). The bioscope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC bioinformatics*, 9(11).

Tentative outline

- Day 1: Scope basics
  - Scopal operators in natural language
  - Scope at the syntax/semantics interface
  - Overview of "non-incremental" approaches to scope (quantifier raising, etc.)
  - Motivation: the problem of incrementality, scientific and application value
- Day 2: Formal and theoretical considerations
  - Incremental parsing from a syntactic perspective
  - Incrementality-compatible semantic representation
  - Typology of scope-bearing operators
  - Comparison of scope phenomena across languages
- Day 3: Psycholinguistic aspects of scope interpretation
  - Brief introduction to methods of investigation in sentence processing
  - The context-continuation experimental paradigm
  - Heuristic-first accounts of scope processing
  - Evidence for higher-order specification
  - **Class activity:** Data collection of multilingual scope judgements
- Day 4: Computational aspects of scope interpretation
  - Corpus-based approaches, corpus resources
  - Brief overview of basic modern machine learning concepts.
  - Machine-learning and multimodal approaches scope interpretation
  - **Class activity:** Analysis (qualitative or quantitative) of collected scope judgements
- Day 5: Integrating event- and world-knowledge
  - Brief introduction to generalized event knowledge
  - Possible methods of invoking event knowledge in scope interpretation
  - Revisiting incrementality-compatible semantic formalism
  - **Class activity:** Brief presentation of collected data and analysis

Prerequisites

This is an advanced but interdisplinary course intended to provoke thought and discussion between computational linguists and psycholinguists, among others, so the prerequisites cannot be very strict. Encouraged would be previous basic training in formal semantics. Exposure to either psycholinguistic experimental methodologies, parsing technologies, or experimental pragmatics would be helpful.

Funding

The organizer is presently located in Sweden (Gothenburg). Some travel support may be available from the Swedish Research Council-funded Centre for Linguistic Theory and Studies in Probability (CLASP).