Hopf algebra reading seminar

András Kornai

September 18 2023 10PM CET

ORGANIZATION

- Will (likely) alternate between 6:30PM and 10PM Monday zooms
- 6:30pm zoom
  https://us02web.zoom.us/j/85255485089?pwd=bFk5VWhva1RTQ
- 10pm zoom
  https://us02web.zoom.us/j/84731595225?pwd=ZVdzZTFmaXJZO
- Slack https://join.slack.com/t/slack-qyx1689/shared_invite/zt-1xppi4d00-WnJhAvg_ThoSBOw9xH7yIw
- Course webpage
  https://nessie.ilab.sztaki.hu/~kornai/2023/Hopf
  Also reachable as kornai.com → 2023 → Hopf
- Attendance sheet
  https://docs.google.com/spreadsheets/d/17cK-cl3_xdbo73_kHWCIAvwgkd-G6qz44J4D6tyFfAc/edit?usp=sharing Please fill in also if you are watching the recordings (without being present on zoom)

Plan for today

1. More on the $\mathcal{L}_{1.5}$ class of languages
2. Research questions
3. Plans for fall semester

Mildly context sensitive languages (MCS)

- Background: lecture 7a (Slides/hopf7a.pdf and Recordings/hopf7a.mp4): strings, stringsets, languages, language families

- $\mathcal{L}_{1.5}$ defined by desirable properties rather than specific generation/acceptance mechanism: polynomial parsability; constant growth or even semilinearity; limited cross-serial dependence *all of these are disputed*

- For polynomial parsability recall that $\mathcal{L}_5$ (finite) is $O(1)$ (independent of string length); $\mathcal{L}_3$ is real time; $\mathcal{L}_2$ is less than cubic ($n^{2.81}$, see Valiant, 1975; $\mathcal{L}_1$ is PSPACE-complete

- For lack of semilinearity, see Michaelis and Kracht, 1997 on Old Georgian; for lack of constant growth, see Radzinski, 1991.

- For the whole cross-serial business, see Huybregts, 1976; Shieber, 1985, but a good enough formal example is the language $\{ww | w \in \{a, b\}^*\}$

So why we like MCS anyway?

- Because they show the convergence of historically important grammar formalisms
- Combinatory Categorial Grammas (CCG) Steedman, 1987 grew out of categorial grammar Ajdukiewicz, 1935; Lambek, 1958 and have very close ties to Montague Grammar Montague, 1970; Montague, 1973
- Head grammars Pollard, 1984, grew out of the *head wrap* operation Bach, 1981 and became highly influential in the later Head-Driven Phrase-Structure Grammar (HPSG) Pollard and Sag, 1994
- Linear indexed grammars Gazdar, 1988 are a subclass of the classic indexed grammars Aho, 1968 and were recognized as a good generalization of the influential Generalized Phrase Structure Grammar (GPSG) Gazdar et al., 1985
- Tree Adjoining Grammars (TAG) Joshi, Levy, and Takahashi, 1975 grew out of string grammars Harris, 1962

"Convergence" is even deeper

- All these theories (CCG, HPSG, GPSG, TAG) grew out of roots preceding the "mainstream" Chomsky, 1957; Chomsky, 1965; Chomsky, 1981; Chomsky, 1995 etc. or, often, from various critiques of the mainstream

- Initially, the key technical device was a *transformation* (arbitrary tree to tree mapping) which became 'Move α' by GB and early Minimalism, which reformulated it as Merge

- This was formalized (to everyone's satisfaction except Chomsky's) by Collins and Stabler, 2016 which enabled a rigorous proof that Minimalist Grammar is yet another MCS formalism

- This was demonstrated using another extension of CFGs, Multiple CFGS (MCFGs) Seki et al., 1991, for a good survey see Clark, 1994.

Old versus new minimalism

- Avery Andrews discussed the difference between old and new in lecture 7p (see Slides/hopf7p.pdf, Recordings/hopf7p.mp4)
- Marcolli et al Marcolli, Chomsky, and Berwick, 2023; Marcolli, Berwick, and Chomsky, 2023 pin the difference on dealing with planar trees (old) or any kind of tree (new)
- This results in a different Hopf Algebra flavor, and eliminates the old MP distinction between "internal" and "external" merge
- Other advantages are also claimed, in particular a more compact theory Berwick, 2015
- But the overall formal power of the system remains unclear

Plans for the Fall

- Remember our motto (stolen from the Hungarian beer company): "The world is as much as you pour into it"
- Students need to formulate goals, find papers they want to present in class, or solve more complex research problems
- Suggestion: we start with MBC rather than MCB
- They essentially use bialgebras, what is the full story about the antipode?
- What is the (weak and strong) generative capacity?
- Continue with applying what we learn to understand LLMs

SCHEDULE UP TO MID-NOVEMBER

Sept 25	6:30pm	Adam Nemecek: Machine Learning with HAs
Oct 2	10pm	Richard Luo: Old Georgian and semilinearity
Oct 9	6:30pm	Isabelle Senturia: Learning CCGs
Oct 16	10pm	Facundo Calabró: Multiple CFGs
Oct 23		No meeting (Hungarian National Holiday)
Oct 30	6:30pm	Stepan Shabalin: TBA
Nov 6	10pm	Blanka Kövér: Bialgebras without an antipode
Nov 13	6:30pm	Michael Bukatin: Mathematical foundations of transformer analysis

📄 Aho, Alfred V. (1968). "Indexed Grammars – An Extension of Context-Free Grammars". In: *Journal of the ACM* 15.4, pp. 647–671.

📄 Ajdukiewicz, Kazimierz (1935). "Die Syntaktische Konnexität". In: *Studia Philosophica* 1, pp. 1–27.

📄 Bach, Emmon (1981). "Discontinuous constituents in generalized categorial grammars". In: *Proceedings of the NELS*. Vol. II, pp. 1–12.

📄 Berwick, Robert C. (2015). "Mind the gap". In: *50 Years Later: Reflections on Chomskyâs Aspects*. Vol. MIT Working Papers in Linguistics, pp. 1–13.

📄 Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.

📄 — (1965). *Aspects of the Theory of Syntax*. MIT Press.

📄 — (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

📄 — (1995). *The Minimalist Program*. MIT Press.

📄 Clark, Robin (1994). *Kolmogorov Complexity and the Information Content of Parameters*. IRCS Report 94-17. Institute for Research in Cognitive Science, University of Pennsylvania.

📄 Collins, Chris and Edward Stabler (2016). "A formalization of minimalist syntax". In: *Syntax* 19.1. DOI: 10.1111/synt.12117.

📄 Gazdar, Gerald (1988). "Applicability of Indexed Grammars to Natural Languages". In: *Natural Language Parsing and Linguistic Theories*. Ed. by U. Reyle and C. Rohrer. Reidel, pp. 69–94.

📄 Gazdar, Gerald et al. (1985). *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

📄 Harris, Zellig S. (1962). *String analysis of sentence structure*. Vol. 1. Papers on Formal Linguistics. Mouton.

📄 Huybregts, Rini (1976). "Overlapping dependencies in Dutch". In: *Utrecht Working Papers in Linguistics* 1, pp. 224–265.

📄 Joshi, Aravind K., L.S. Levy, and M. Takahashi (1975). "Tree adjunct grammars". In: *Journal of Computer and System Sciences* 10.1.

📄 Lambek, Joachim (1958). "The mathematics of sentence structure". In: *American Mathematical Monthly* 65, pp. 154–170.

📄 Marcolli, Matilde, Robert Berwick, and Noam Chomsky (2023). *Old and New Minimalism: a Hopf algebra comparison*. URL: https://lingbuzz.net/lingbuzz/007373.

📄 Marcolli, Matilde, Noam Chomsky, and Robert Berwick (2023). *Mathematical Structure of Syntactic Merge*. arXiv: 2305.18278 [cs.CL].

📄 Michaelis, Jens and Marcus Kracht (1997). "Semilinearity as a Syntactic Invariant". In: vol. 1328. Lecture Notes in Computer Science. Springer, 329â–345.

📄 Montague, Richard (1970). "Universal Grammar". In: *Theoria* 36, pp. 373–398.

📄 — (1973). "The proper treatment of quantification in ordinary English". In: *Formal Philosophy*. Ed. by R. Thomason. Yale University Press, pp. 247–270.

📄 Pollard, Carl (1984). *Generalized Phrase Structure Grammars, Head Grammars, and Natural Language*. Stanford University: PhD thesis.

📄 Pollard, Carl and Ivan Sag (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press.

📄 Radzinski, Daniel (1991). "Chinese number-names, Tree Aadjoining Languages, and Mild Context-Sensitivity". In: *Computational Linguistics* 17.3, pp. 277–299.

📄 Seki, Hiroyuki et al. (1991). "On multiple context-free grammars". In: *Theoretical Computer Science* 88, pp. 191–229.

📄 Shieber, Stuart M. (1985). "Evidence against the context-freeness of natural language". In: *Linguistics and Philosophy* 8, pp. 333–343.

📄 Steedman, Mark (1987). "Combinatory grammars and parasitic gaps". In: *Natural Language and Linguistic Theory* 5, pp. 403–439.

📄 Valiant, L.G. (1975). "General context-free recognition in less than cubic time". In: *Journal of Computer and System Sciences*

10.2, pp. 308–314. ISSN: 0022-0000. URL:
http://repository.cmu.edu/compsci/1752.