Hopf algebra reading seminar

András Kornai

June 21 2023 2PM CET

Organization

- 2pm zoom
  https://us02web.zoom.us/j/84045659802?pwd=L3grbWtqREE4OE
- 11pm zoom
  https://us02web.zoom.us/j/89203668566?pwd=M1dRL2ozOWxBT
- 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
- Add github repo?

Goals

1. Our focus will be with the application(s) to natural language
2. HAs are a family of methods not a single unified theory – we need to understand how to build HAs for typical data structures used by linguists: (planar) trees, wild trees, finite functions, DAGs, codescriptive stuff,...
3. We want not just a shared understanding, but we want to (computer) verify it
4. Looking for volunteers to explain minimalism, HAs, ML, and TP from an expert perspective to non-experts. Sign up on the appropriate slack channel
5. If you have goals that don't look like a good fit with the above, explain them now

Background: symbolic computation with NNs

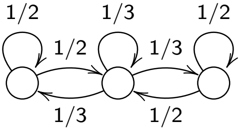
- We start with three papers: McCulloch and Pitts, 1943; Little, 1974 and Smolensky, 1990
- McC-P: We have a bunch of *neurons* (IRL 10^{12}) which we assume to be binary (on or off). Individually they change state on a millisecond, collectively on femtosecond scale. Each brain state is characterized by a *thought vector* $\Psi(t)$ of dim 10^{12}
- I/O neurons are clamped to some state (we will ignore these for the most part)
- Language acquisition takes place on a multi-year scale, with a child learning maybe (peak) 40 words/day. If state vectors are updated according to some $2^{10^{12}}$ by $2^{10^{12}}$ transition matrix P as $\Psi(t+1) = \Psi(t)P$, we see P change adiabatically (1 change in P for 10^{17} changes in Ψ). On a second/minute scale we may as well assume P *is constant*

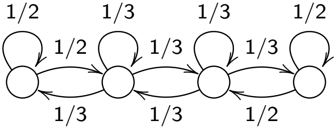
LITTLE:1974

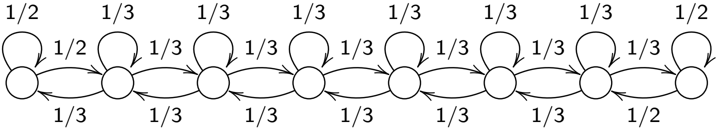
- P is stochastic, we have a Markov chain. (Once we know enough, we may invite Amy Pang to discuss her stuff.)
- By Perron-Frobenius, P has a unique largest eigenvalue $\lambda_0 = 1$, the corresponding eigenvector gives the long-term state distribution ϕ_0
- P is big. But with probability 1 it is diagonalizable, so we can select an eigenvector basis $\{\phi_r\}$ arranged in decreasing order of the eigenvalues λ_r so that for any vector α we have $\alpha = \sum_r a_r\phi_r$
- In this basis, $\langle\Psi(t+1)|P|\Psi(t)\rangle = \Sigma_r\lambda_r\phi_r(\Psi(t+1))\Psi_r(t)$

Machine Learning

Linear algebra for automata







Survey – beginner

- What is the P transition matrix for the top automaton?
- Can you compute the eigenvector basis?
- How about the second automaton?
- How about the third (8 state) automaton?

Survey – intermediate

- Can you generalize to 2^n states?
- What software are you using to compute the eigenvalues? How far can you push it?
- Can you provide a formula for the second largest eigenvalue?
- Can you prove it?
- Can you locate a computer-verified proof of Perron-Frobenius?

SURVEY – ADVANCED

- Can you generalize from linear connectivity (as on the examples) to planar, spatial, n-dim? How do the transition matrices P look like?
- What can you show about the gap between 1 and the 2nd largest eigenvalue?
- Can you coax a theorem prover into proving your formula?
- How about proving upper and lower bounds?
- Physicists do a lot of stuff the mathematicians find objectionable, but the results often stand up. How can TPs/verifiers support such arguments?

- Little, W. A. (1974). “The existence of persistent states in the brain”. In: *Mathematical Biosciences* 19, pp. 101–120.
- McCulloch, W.S. and W. Pitts (1943). “A logical calculus of the ideas immanent in nervous activity”. In: *Bulletin of mathematical biophysics* 5, pp. 115–133.
- Smolensky, Paul (1990). “Tensor product variable binding and the representation of symbolic structures in connectionist systems”. In: *Artificial intelligence* 46.1, pp. 159–216.