Hopf algebra reading seminar

András Kornai

July 5 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

Plan for today

1. Zsolt Zombori on the analog/digital question
2. Gerald Penn on Smolensky (1990)
3. (Time permitting) AK: tensors in multilinear algebra

Analog vs Digital

Zsolt Zombori

July 2023

Representation

- Whatever can be represented by a Turing machine can be continuously approximated by a (recurrent) neural network
- Whatever can be represented by a neural network can be discretely approximated by a Turing machine

Gradient Descent Does not Work on Digital

- Current methods for optimizing neural networks use gradient descent
- Requires differentiable computation from parameters to optimisation target
- In a digital network, we don't have a good notion of "approaching the solution"

Gradient Descent is almost linear

- Time complexity of a single update step is somewhere between linear and quadratic
- No theoretical knowledge about sample complexity, but in practice it is linear
- One can stop training when over the budget and get a decent solution
- Massively parallelizable
- This is why people are optimizing networks with billions of paramters!

Discrete Optimisation is usually exponential

- For fitting to a set of samples, a naive solution simply checks all combinations
- Optimised methods cannot do much better

Gerald Penn on Smolensky 1990

See slides at Hopf/Slides/hopf3g.pdf

Graded algebras

- Family of algebras $A_i (i \in \mathbb{N})$ over some field or ring.
- Must have some *degree structure* so that formal products $v_i \in A_i \cdot v_j \in A_j$ are in A_{i+j}
- Hamel basis versus Schauder basis
- Further examples
- Is this the grading used in Hopf algebras? $A_i \otimes A_j$ is found in $A_{i \cdot j}$ not A_{i+j}: $\dim A_i \otimes A_j = \dim A_i \cdot \dim A_j$ (the shape parameter is additive: 2-tensor $\otimes$ 3-tensor $=$ 5-tensor)