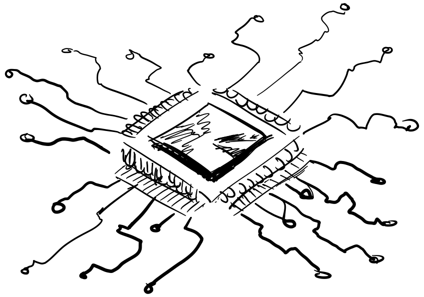
AI



QNLP

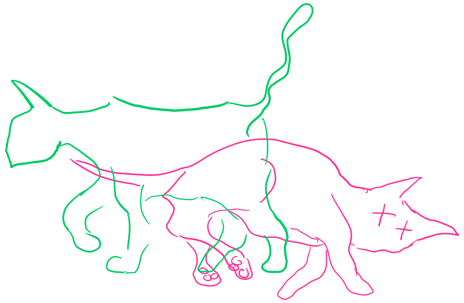
NLP

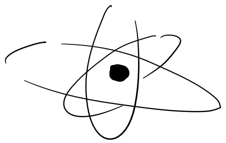
QC

Konstantinos Meichanetzidis

Cambridge Quantum

ACT 2021





Starring

Alexis Toumi

Giovanni de Felice

Anna Pearson

Robin Lorenz

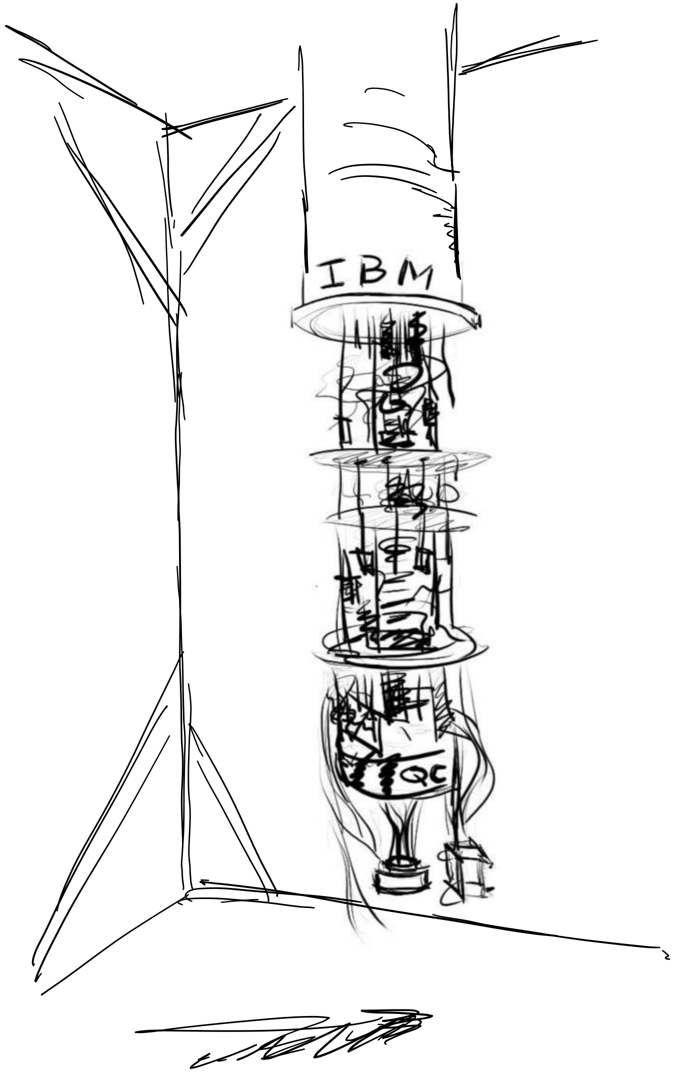
Dimitri Kartsaklis

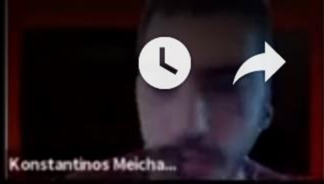
B💀b C💀ecke

Richie Yeung

$t|ket\rangle$ team
QML team

And HALQ 9000





NISQy
QNLP

Konstantinos Meichanetzidis

https://www.cs.ox.ac.uk/people/konstantinos.meichanetzidis/

ACT 2020

Results from the
first-ever QNLP experiments
run on quantum hardware.

Why NLP

Market Overview

The Global Natural Language Processing (NLP) Market was valued at USD 10.72 billion in 2020, and it is expected to be worth USD 48.46 billion by 2026, registering a CAGR of 26.84% during th forecast period (2021-2026). Due to the ongoing Covid-19 pandemic the market is witnessing growth in healthcare sector.

mordorintelligence.com

- Automated conversation over a domain
- Information retrieval, search, QA
- Translation, Summarisation
- Text to Speech and vice versa, human-computer interaction
- Language generation, creativity tools
- Sentiment analysis
- Cognition, scientific interest

Why NLP: because AI

"NLP is AI-hard"

Turing test :
verify intelligent multi-domain behaviour
via use of language!

"The limits of my language mean the limits of my world"
- Ludwig Wittgenstein
"Language faculty is what separates us from other species"
- Noam Chomsky

NLP : SotA

WIKIPEDIA
The Free Encyclopedia

Main page
Contents

Article Talk

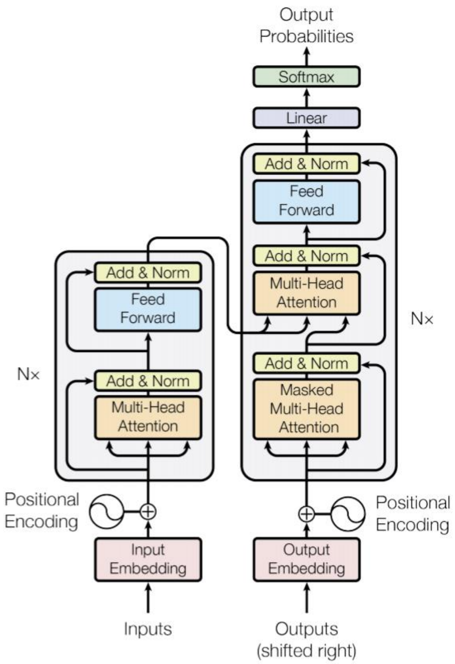
GPT-3

From Wikipedia, the free encyclopedia

Generative Pre-trained Transformer 3 (**GPT-3**) is an autoregressive language model that uses deep learning to produce human-like text. It is the third-generation language prediction model in the GPT-n series (and the successor to GPT-2) created by OpenAI,

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com



All tweets written by a 175 billion parameter natural language processing artificial intelligence. Account curated by Ben Prentice. Credit: @sushant_kumar.

Everywhere Joined July 2020

1 Following 1,432 Followers

Not followed by anyone you're following

Tweets Tweets & replies Media Likes

GPT-3 @GPT3_ · Oct 13, 2020 ···

You can procrastinate doing anything but you can't procrastinate on procrastination.

6 6 21

Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity

William Fedus* Google Brain liamfedus@google.com
Barret Zoph* Google Brain barretzoph@google.com
Noam Shazeer Google Brain noam@google.com

US-China tech war: Beijing-funded AI researchers surpass Google and OpenAI with new language processing model

- The WuDao 2.0 natural language processing model had 1.75 trillion parameters, topping the 1.6 trillion that Google unveiled in a similar model in January

Human Brain: 200 billion neurons, 125 trillion synapses (just in the cerebral cortex)

“but”

nature

Explore content ˅ Journal information ˅ Publish with us ˅ Subscribe

nature > news feature > article

NEWS FEATURE · 03 MARCH 2021

Robo-writers: the rise and risks of language-generating AI

A remarkable AI can write like humans – but with no understanding of what it’s saying.

Matthew Hutson

MIT Technology Review   Topics

GETTY

Artificial intelligence / Machine learning

OpenAI’s new language generator GPT-3 is shockingly good—and completely mindless

The AI is the largest language model ever created and can generate amazing human-like text on demand but won't bring us closer to true intelligence.

by **Will Douglas Heaven**   July 20, 2020

Stanford SOCIAL INNOVATION Review
Informing and inspiring leaders of social change

SOCIAL ISSUES SECTORS SOLUTIONS MAGAZINE MORE

Technology

The Case for Causal AI

Using artificial intelligence to predict behavior can lead to devastating policy mistakes. Health and development programs must learn to apply causal models that better explain why people behave the way they do to help identify the most effective levers for change. *Open access to this article is made possible by Surgo Foundation.*

SHARE COMMENT DOWNLOAD PRINT ORDER REPRINTS

By Sema K. Sgaier, Vincent Huang & Grace Charles | Summer 2020

Pitfalls of Static Language Modelling

Angeliki Lazaridou[*♡△♠] Adhiguna Kuncoro[*♡△] Elena Gribovskaya[*♡△]
Devang Agrawal[◇♡] Adam Liška[◇♡] Tayfun Terzi[◇] Mai Gimenez[◇]
Cyprien de Masson d’Autume[◇] Sebastian Ruder[♡] Dani Yogatama[♣]
Kris Cao[♣] Tomas Kocisky[♣] Susannah Young[♣] Phil Blunsom[♣♠]
DeepMind, London, UK
{angeliki,akuncoro,egribovskaya}@google.com

Trust Issues

$12M training cost
After training, the world moves on,
what it “knows” becomes obsolete.

“but”

“Deep learning has instead given us machines with truly impressive abilities but no intelligence. The difference is profound and lies in the absence of a model of reality.”
- Judea Pearl

NN learns the **statistical relationships** between words in large amounts of text data.

Susceptible to *bias* present in the training data, no critical thinking.

No **coherent** and **unified** model of the world,
struggle with **context** and out-of-text meaning.

Famous examples: “– I want to kill myself – sounds like a good idea”
“– how many eyes does a foot have? – 2”

Unable to answer “Why questions”, **causal reasoning**.

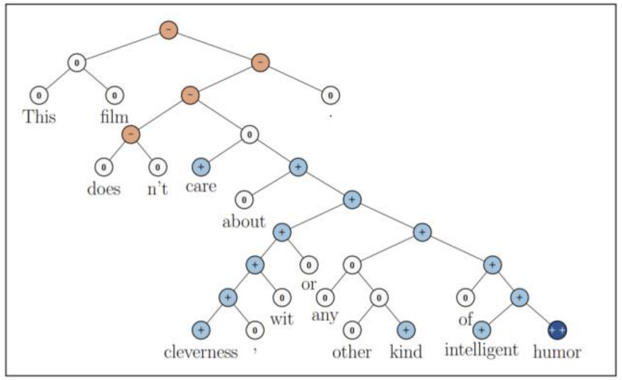
AI pioneer Geoff Hinton: “Deep learning is going to be able to do everything”

Back to basics

Untapped potential in formal, rigorous, structural, **hybrid** approaches.

What is the **mechanics of meaning**?
Focus more on models rather than compute.

"I would encourage everyone to think about problems, architectures, cognitive science, and the details of human language, how it is learned, processed, and how it changes, rather than just chasing state-of-the-art numbers on a benchmark task."
– Chris Manning



Notable example:
Syntax-aware RNN model
https://www.socher.org/

After this motivating intro…

Why ACT

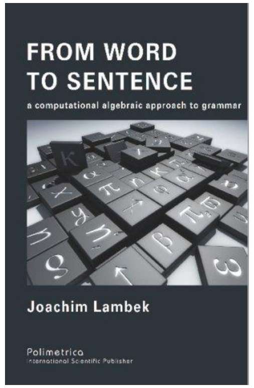
Category theory allows us to reason about structure-respecting mappings between seemingly disparate domains.

And so, it keeps us sane when we build AI models which by nature are complex.

Bob's talk yesterday @ Industry session

Enter the DisCo

and party the *good old fashion* way



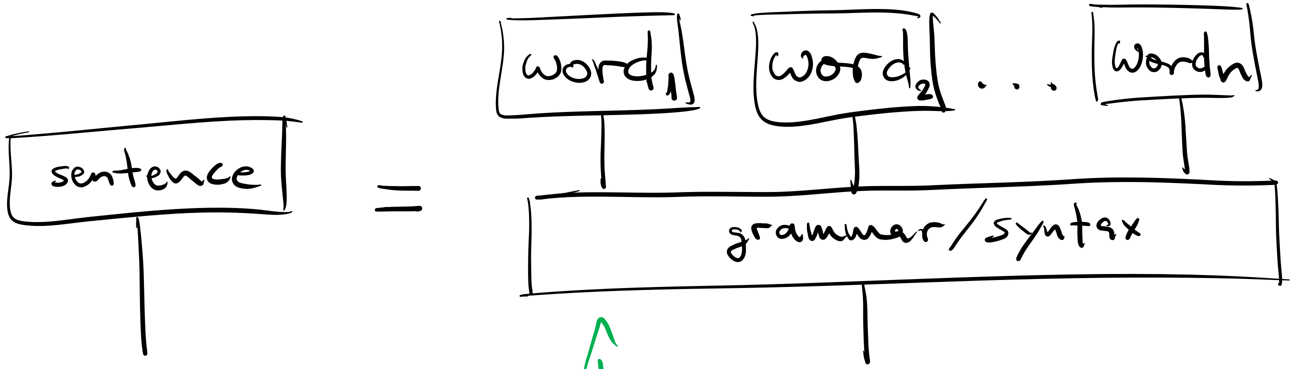
DisCoCat: Coecke, Sadrzadeh, Clark [1003.4394]

Distributional Hypothesis (Firth)

+

Principle of Compositionality

(https://plato.stanford.edu/entries/compositionality/)



Input:
word meanings

Process:
grammar that *composes* them

Output:
sentence meaning

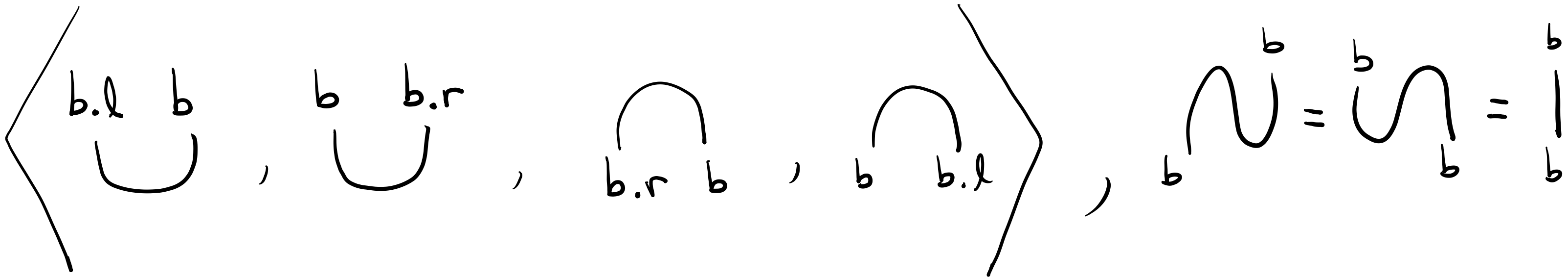
But what is in this box?

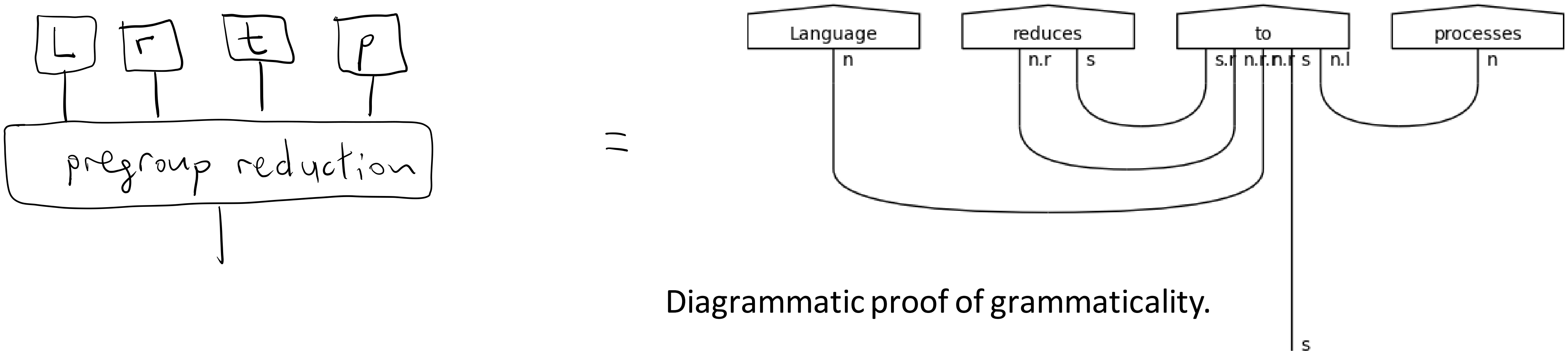
Any mathematical grammar will do fine

DisCo and pregroup grammar

DisCoCat: Coecke, Sadrzadeh, Clark [1003.4394]

Parser: words get typed by a product of pregroup types $b \in \{s, n, \dots\}$,
which can be right- or left- adjoint.
Cups denote type-reductions: $b.l\ \ b \to \epsilon\ ,\ \ b\ \ b.r \to \epsilon$





Diagrammatic proof of grammaticality.

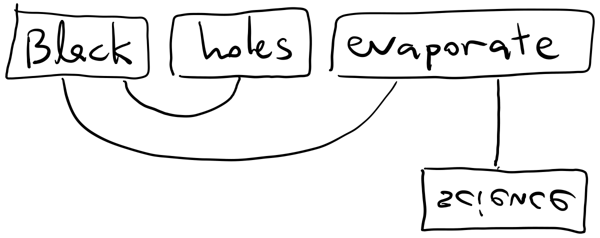
QDisCo

string diagrams

$F_V \swarrow \quad \searrow F_H$

tensor networks $\leftrightarrow$ quantum processes

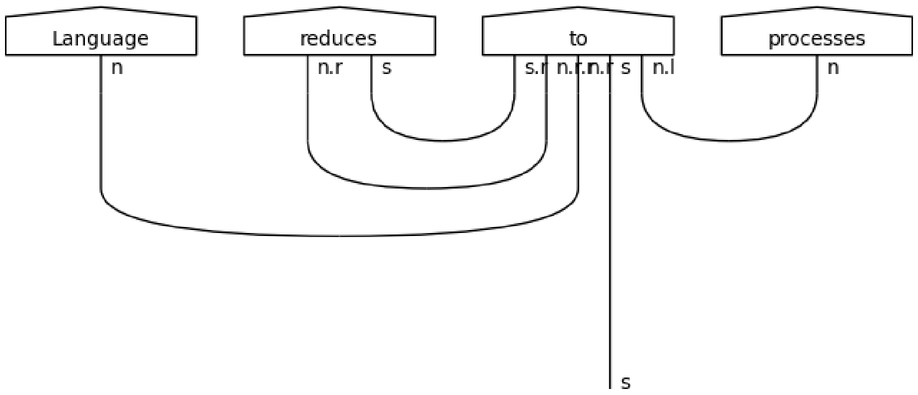
Zeng and Coecke [1608.01406]:
quadratic speedup using qualgo for closest vector problem
(assuming QRAM)



= ⟨science|Bhe⟩

Work with Gogioso and Chiappori [2005.04147]:
mapping DisCo diagrams to parameterized quantum circuits

vectors/tensors
↓
tensor contractions
↓
sentence meaning
encoded in a vector



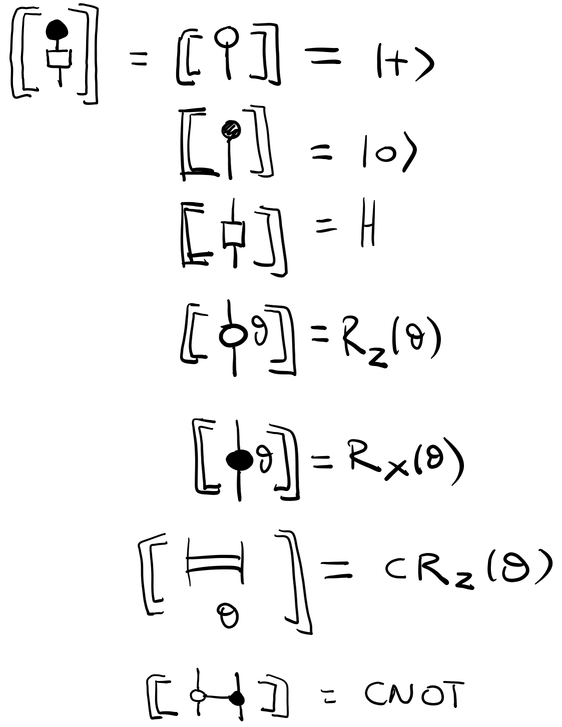
quantum states
↓
Bell effects
↓
sentence meaning
encoded in a
quantum state

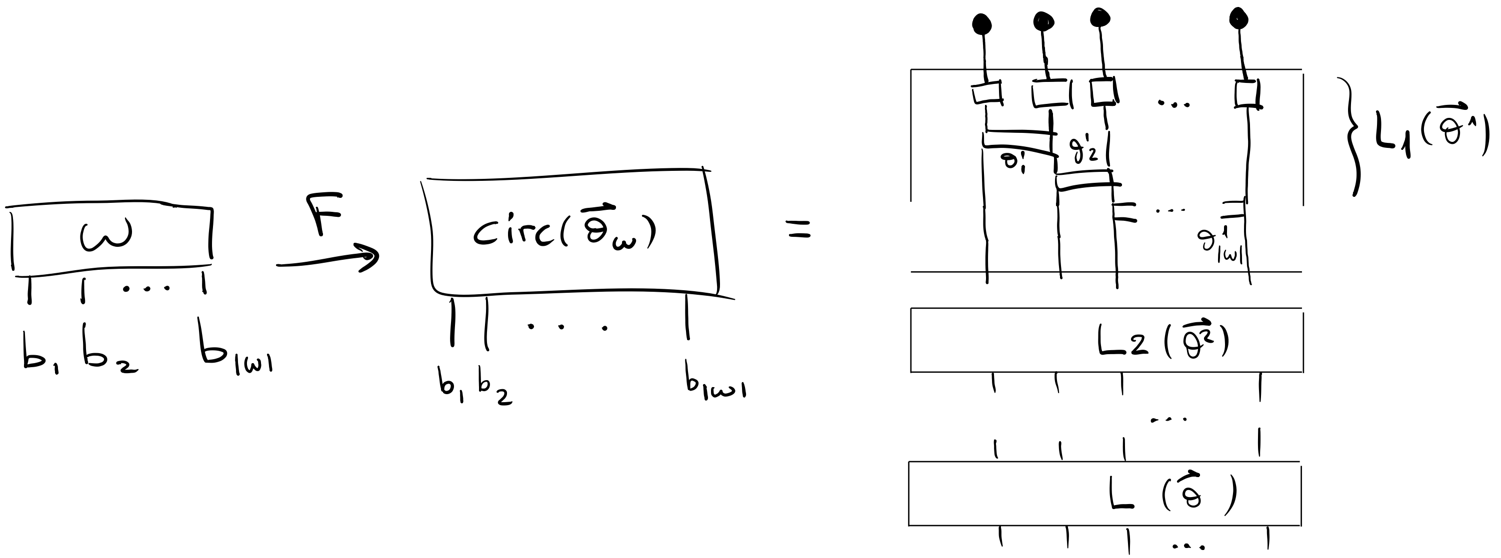
Mathematical similarity begs the q

Run it on a QC

Define a "functor" $F(\theta): D \to C(\theta)$
Each type is assigned a number of qubits $b \to q_b$
Each b-wire carries a Hilbert space of dimension 2^{q_b}
Each state is prepared by a parameterised circuit.

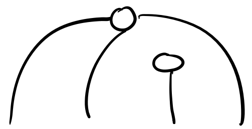
Example $\mathrm{n}_b = 1, \forall b$:

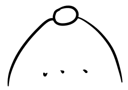




Addendum

Functional words like “is”, “that”, “who”, etc
are treated as “wirings” themselves,
i.e. we invent motivated meaning ansaetze for them
which live at the string diagram level.

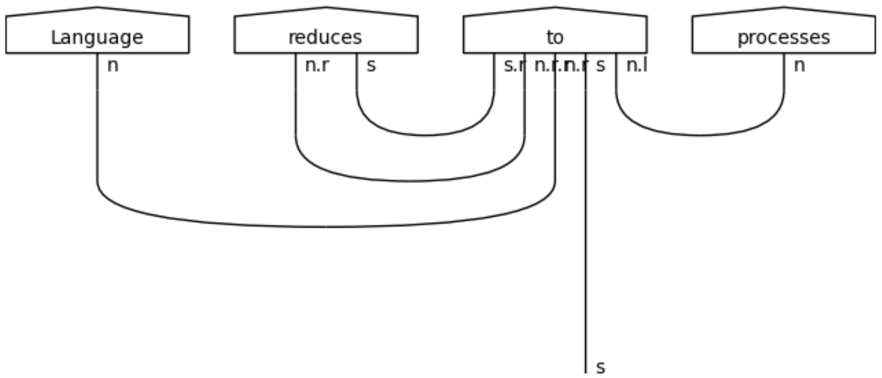
ex "that" $\xrightarrow{F}$ 

where  Frobenius "spider"

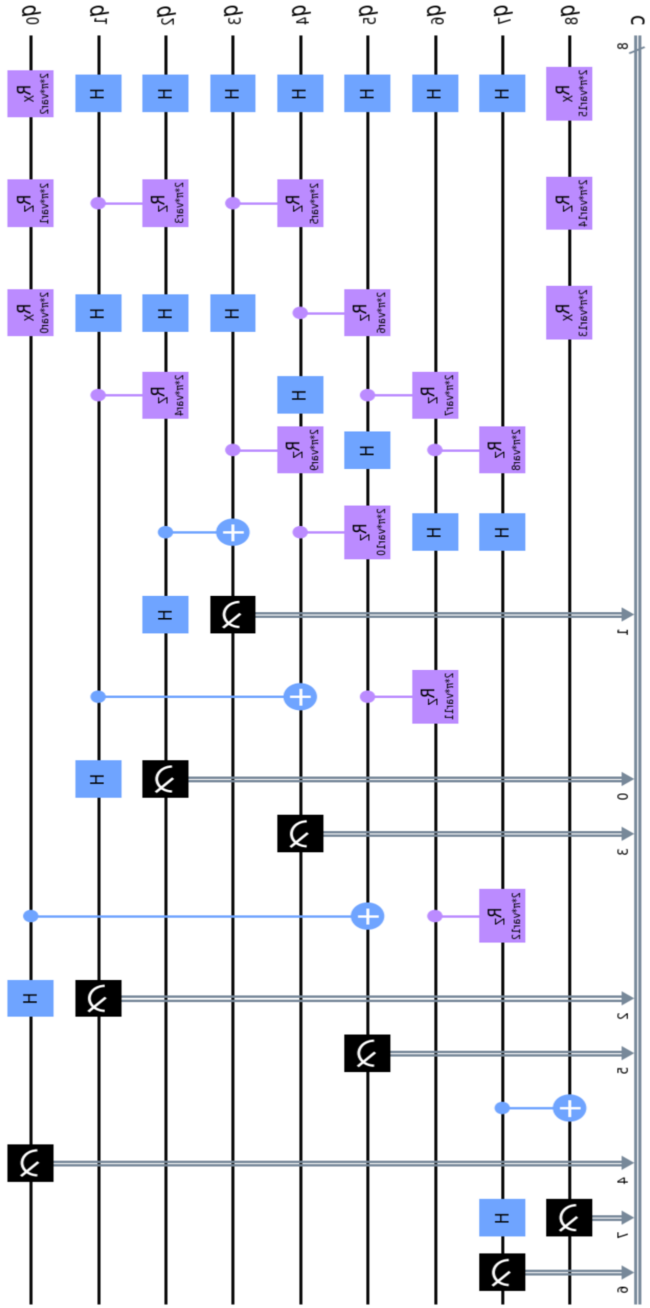
Stick it on a QC

Example $q_n = 1$, $q_s = 1, d = 2$:

$D =$



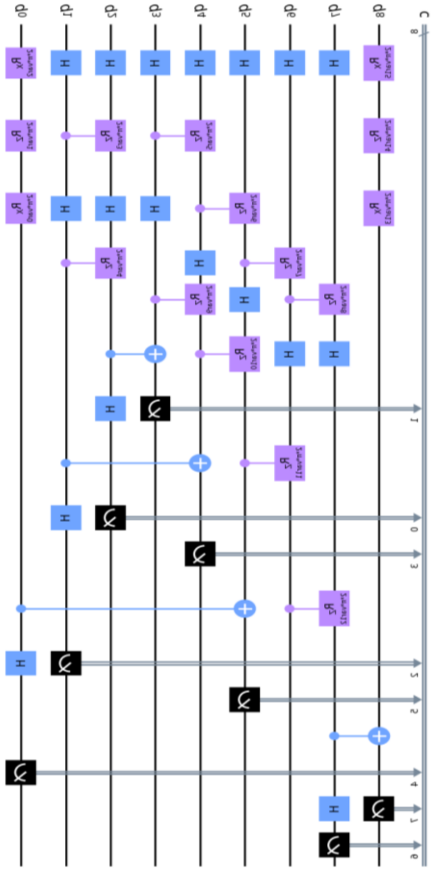
$\xrightarrow{F}$ $C(\theta) =$



ccg2diagram [2105.07720]
https://qnlp.cambridgequantum.com/generate.html

we use DisCoPy [2005.02975] to compose diagrams and apply functors

Compiler: CQC's $t|ket\rangle$ [2003.10611]



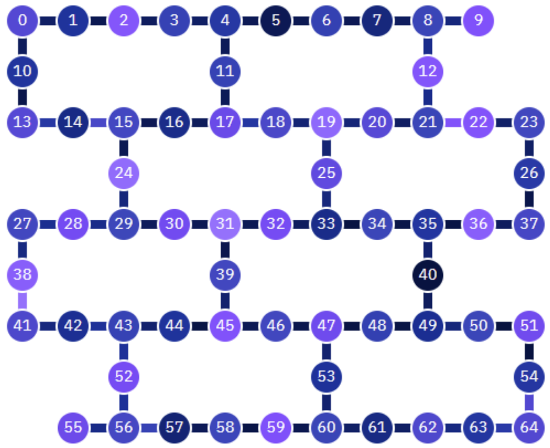
$t|ket\rangle$

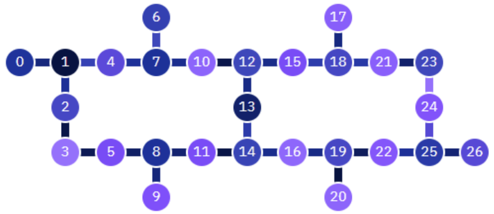
$t|ket\rangle$

$t|ket\rangle$

ibmq_bogota (5qb) (log2qv 5)

ibmq_manhattan (65qb) (log2qv 5)

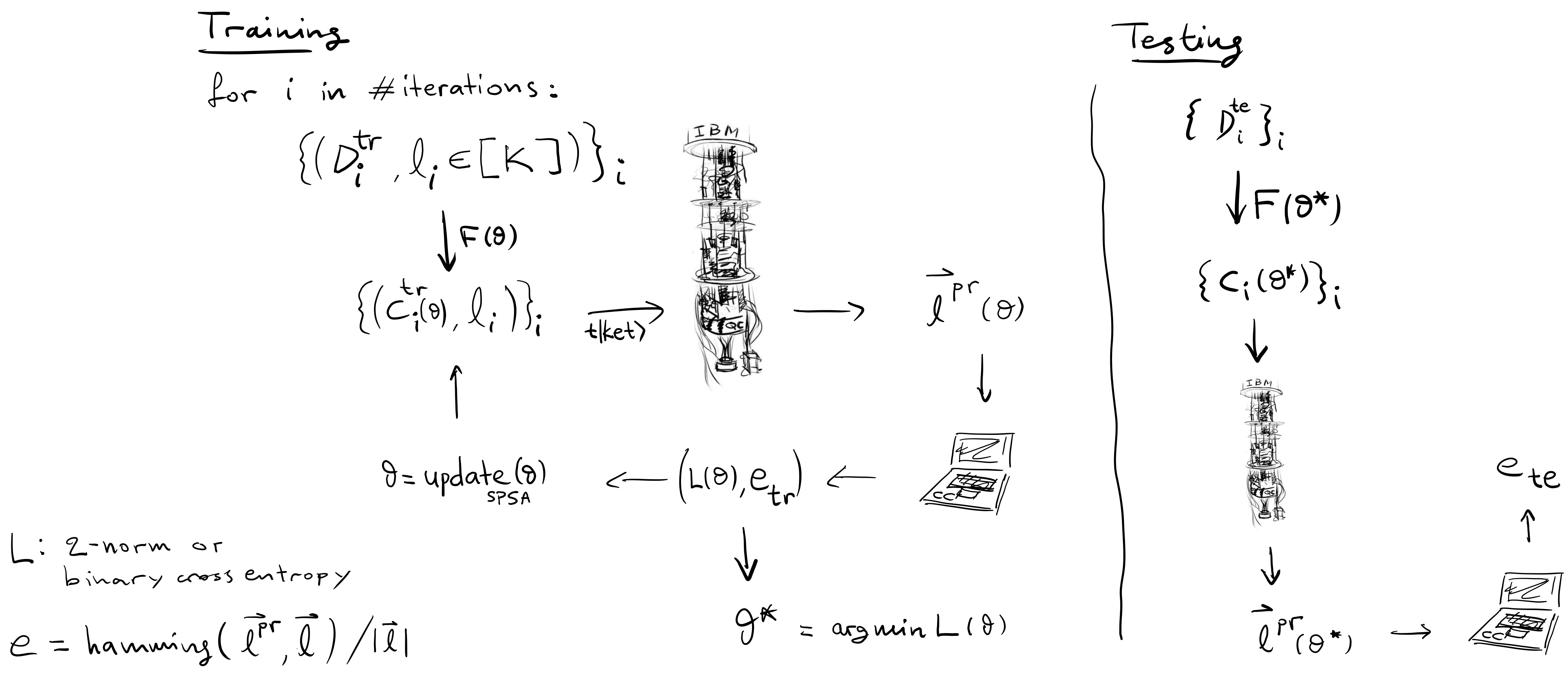




ibmq_montreal (27qb) (log2qv: 7 now, 5 then)

noisy CNOTs !

Learning the functor: sentence classification



Tasks

Truth value: $q_n = 1, q_s = 0$
Diagram is a scalar

False:
(Juliet kills Romeo who dies, 0)
(Romeo kills Juliet, 0)
(Romeo who kills Juliet dies, 0)
…

True:
(Juliet dies, 1)
(Romeo who dies loves Juliet, 1)
(Romeo who kills Romeo dies, 1)
…

Topic: $q_n = 1, q_s = 1$
Diagram is a state

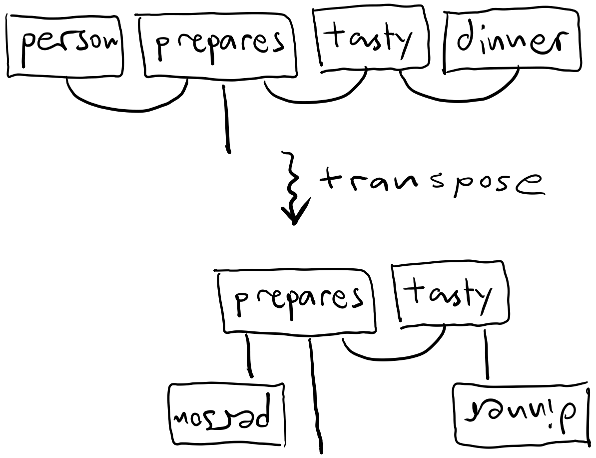
Cooking:
(Skilful man prepares sauce, 0)
(Woman cooks tasty meal , 0)
(Skilful person prepares meal, 0)
…

Technology:
(Skilful woman debugs program , 1)
(Man prepares useful application, 1)
(Person debugs useful software, 1)
…

Relative pronoun type: $q_n = 1, q_s = 0$
Diagram is a scalar

Subject:
(Organisation that establishes church, 0)
(Vehicle that replaces horse, 0)
(Organisation that has team, 0)
…

Object:
(Organisation that church establishes, 1)
(Vehicle that family owns, 1)
(Organisation that team joins, 1)
…



Notes:

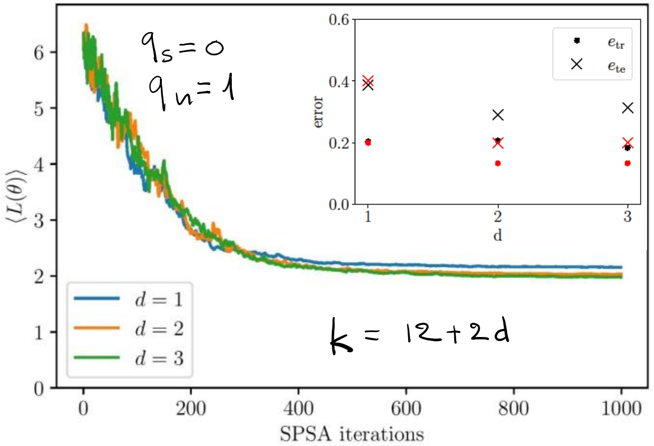
Intra-sentence correlations are ‘quantum’: due to grammar
Inter-sentence correlations are ‘classical’: due to shared words

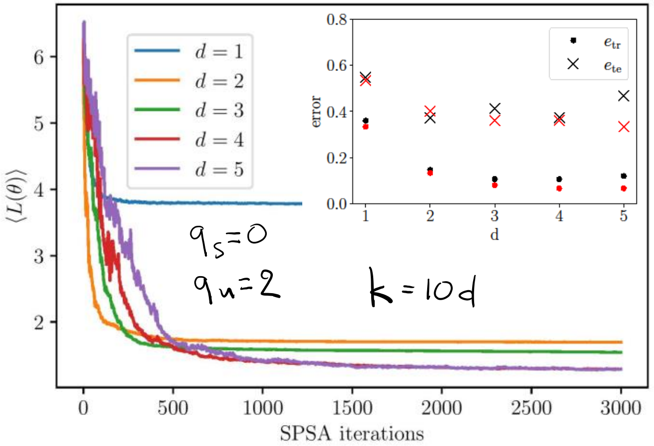
Classical Simulation

$\langle L(\theta)\rangle$ over 20 realisations

Truth Value

basinhopping over Nelder-Mead

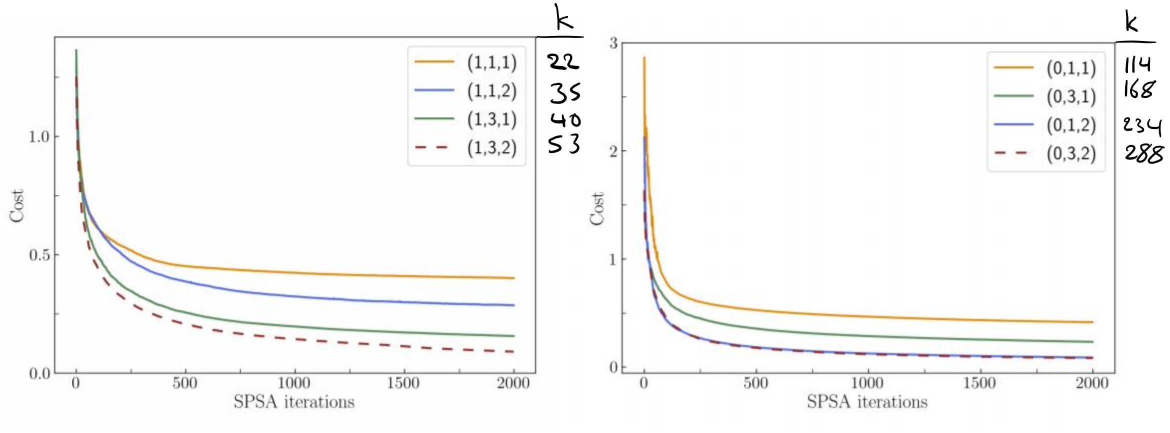




Topic

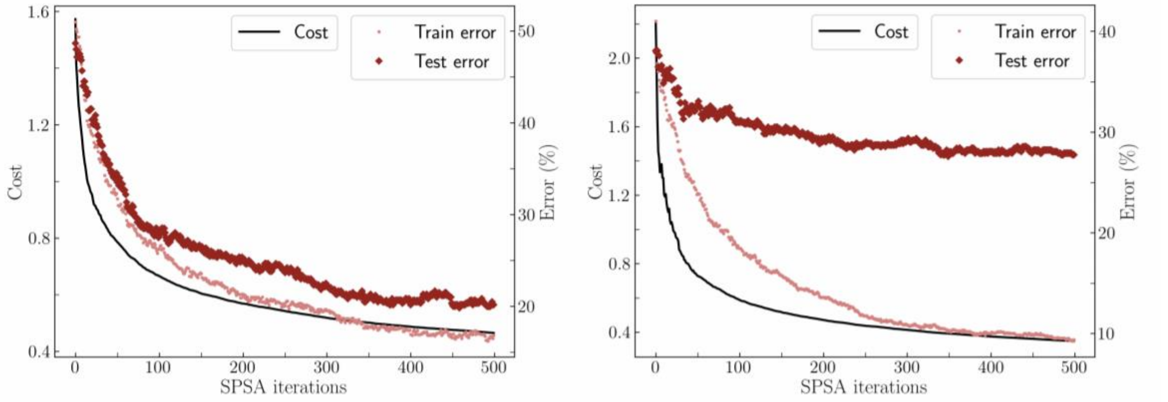
hyperparams (q_s, p_n, d)

Rel pron type



(a) (b)

Figure 7: Convergence of the models in the classical simulation (averaged over 20 runs) for different ansätze; in (a) for *MC* task and in (b) for *RP* task.



(a) (b)

Figure 8: Classical simulation results for the cost and errors (again averaged over 20 runs) in (a) for *MC* task and chosen ansatz $(1, 3, 1)$ and in (b) for *RP* task and chosen ansatz $(0, 1, 2)$.

QC (IBMQ)

Truth Value

[2012.03756]

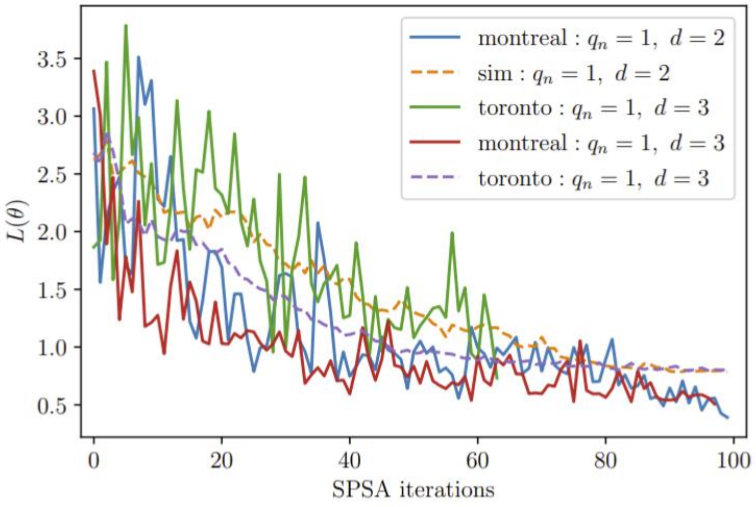
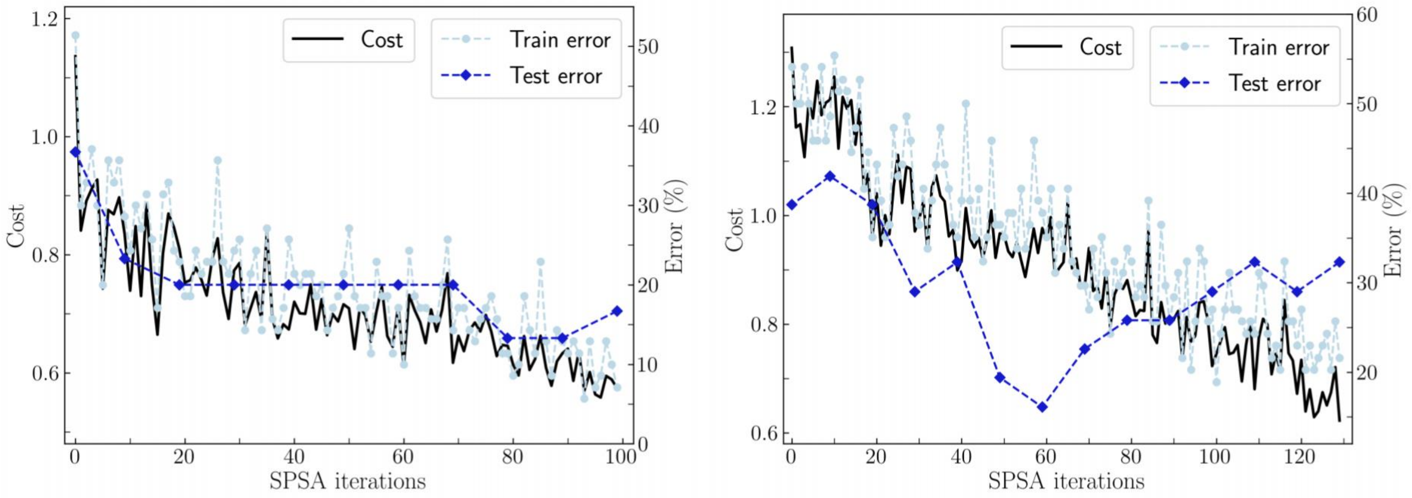


FIG. 5. Convergence of the cost $L(\theta)$ evaluated on quantum computers vs SPSA iterations for corpus K_{16}. For $q_n = 1$, $d = 2$, for which $|\theta| = 10$, on `ibmq_montreal` (blue) we obtain $e_{\mathrm{tr}} = 0.125$ and $e_{\mathrm{te}} = 0.5$. For $q_n = 1$, $d = 3$, where $|\theta| = 13$, on `ibmq_toronto` (green) we get $e_{\mathrm{tr}} = 0.125$ and a lower testing error $e_{\mathrm{te}} = 0.375$. On `ibmq_montreal` (red) we get both lower training and testing errors, $e_{\mathrm{tr}} = 0$, $e_{\mathrm{te}} = 0.375$ than for $d = 2$. In all cases, the CNOT-depth of any sentence-circuit after t|ket⟩-compilation is at most 3. Classical simulations (dashed), averaged over 20 realisations, agree with behaviour on IBMQ for both cases $d = 2$ (yellow) and $d = 3$ (purple).

Topic

Rel pron type

[2102.12846]



(a)

(b)

Figure 9: Results from quantum computation for cost and train and test errors (test error for every 10th iteration) in (a) for *MC* task and chosen ansatz $(1, 3, 1)$ and in (b) for *RP* task and chosen ansatz $(0, 1, 2)$.

bogota

Many many things to do

Regularisation for PQCs, error mitigation,
gradient methods and other non-gradient opt methods

Tweak grammar model to get rid of post selection

Mixed state models

More tasks (sentiment analysis (again classification),
summarisation, translation, etc),
generation, disambiguation, …

Compare with more sophisticated baselines;
what can grammar and syntax actually buy you?

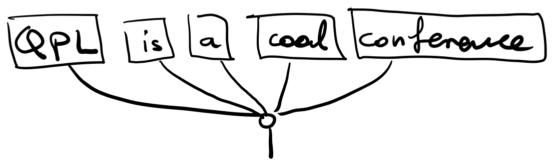
Real-world data instead of hand-crafted.
ccg2diagram has processes Alice in wonderland

Coherently compose sentences
into text circuits

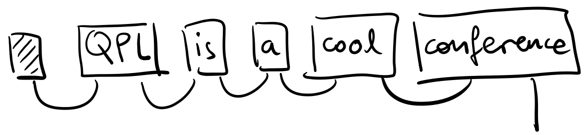
Quadvantage?

Beyond NLP?

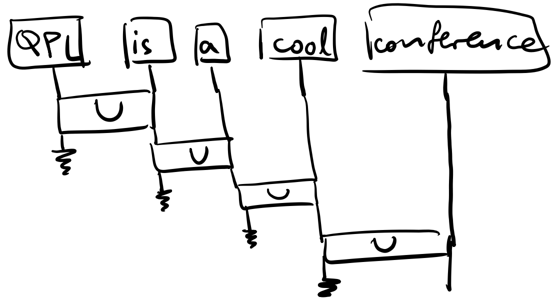
"spider"



"recurrent"

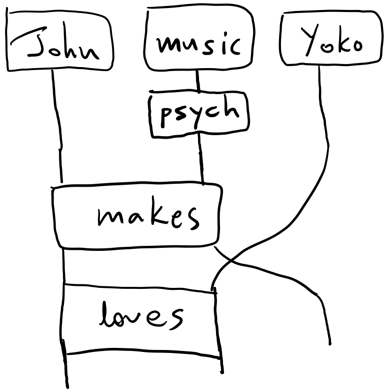


"ladder"



or grammar-aware version
in terms of parse trees

DisCoCirc



That's all for now, stay tuned.