Transformers for Scattering Amplitudes

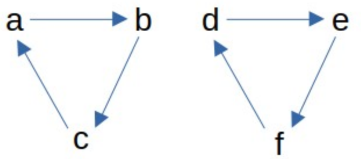
Garrett Merz, Kyle Cranmer, Francois Charton, Lance Dixon, Tianji Cai, Matthias Wilhelm, Niklas Nolte

Scattering Amplitudes and the Symbol

- In N=4 planar super-Yang Mills theory, Feynman integrals are often composed of functions called multiple polylogarithms
- Polylogarithms can be decomposed into products of elementary tensors: we can write the two-loop term of the three-gluon MHV form factor*:

$$\begin{aligned}\mathcal{S}[F_3^{(2)}] = 4\Big[& b\otimes d\otimes d\otimes d + c\otimes e\otimes e\otimes e + a\otimes f\otimes f\otimes f \\ & + b\otimes f\otimes f\otimes f + c\otimes d\otimes d\otimes d + a\otimes e\otimes e\otimes e\Big] \\ + 2\Big[& b\otimes b\otimes b\otimes d + c\otimes c\otimes c\otimes e + a\otimes a\otimes a\otimes f \\ & + b\otimes b\otimes b\otimes f + c\otimes c\otimes c\otimes d + a\otimes a\otimes a\otimes e\Big]\end{aligned}$$

- The symbol (composed of words of length w=2L over six letters & their coefficients) follows some very nice rules:

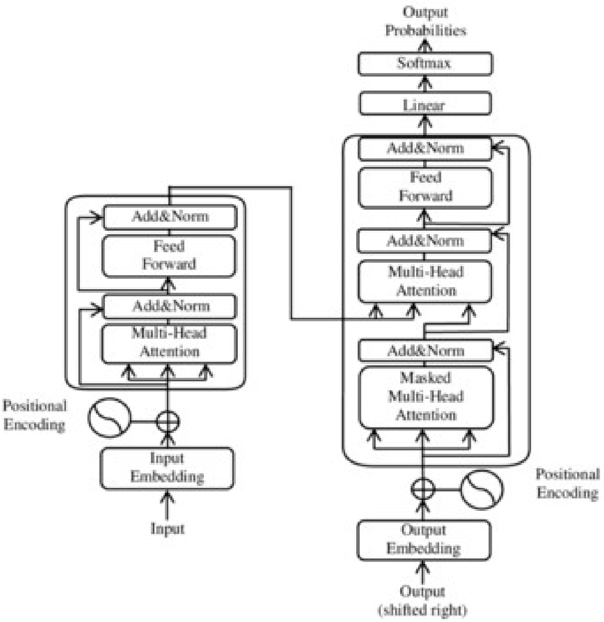


  - Invariance under dihedral symmetry
  - Forbidden letter adjacencies
  - Words must start with a,b,c and end with d,e,f
  - Linear relations: ex. fix the first 2L-4 letters of a word as anything {denote these as ****}: then for all L, we have relations like:
    - Coef(****ddbd)-Coef(****dbdd)=0
    - Coef(****cbdd)-Coef(****dbdd)=0

(*specifically, this is the two-loop term of the normalized BDS remainder function of the form factor)

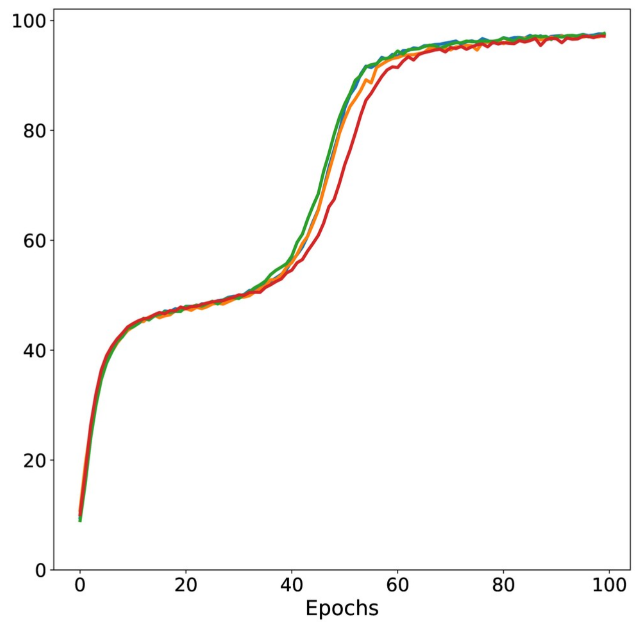
Can A Transformer Learn It?

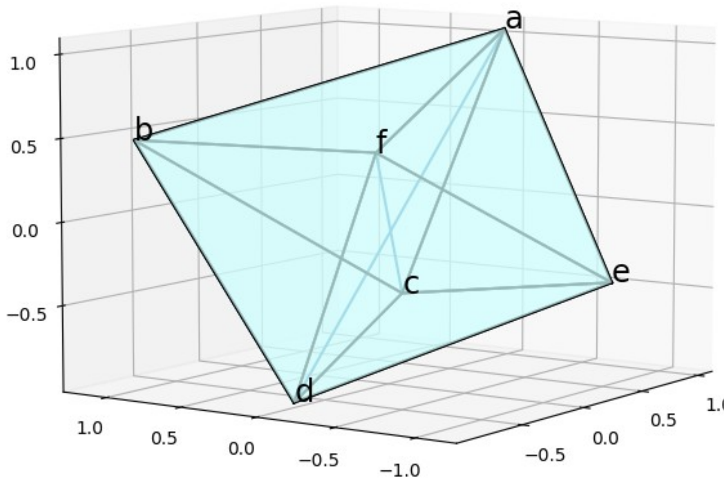
- Current methods of calculating the symbol become computationally intractable beyond eight loops
- Lots of use-cases for ML here! Symbol looks “kind of like language”, and because solution is “easy to check” using linear relations but “difficult to guess”, we can ask whether a language model can:
  - propose an ansatz for coefficients at unknown loops, or otherwise narrow down the dimensionality of the solution space
  - Find unknown recurrence relations that connect terms across loop orders (i.e., symbolic regression)
  - Expedite difficult stages of the existing symbol calculation procedure



Task 1: Coefficient From Word

- At a fixed loop (6 loops), given a “word”, can a model predict its coefficient?
  - This is essentially a “tensor completion” task!
- Train a 2-layer encoder-decoder Transformer (d=256, 8 heads) on some percentage of nonzero terms in the full symbol (as low as 20% still works!), validate on a holdout set (1M elements train; 100k test)
  - Tokenize coefs in base 1000 with sign as a separate token, cross-entropy loss
- Model performs very well (98% accuracy after 120 epochs)
- Unusual double-plateau structure: at the first plateau, the model gets coef magnitudes correct, but cannot resolve sign
  - This is present whether the sign is first or last token: likely not an artifact of autoregressive decoder!
- At 6 loops, PCA on input token embeddings shows dihedral symmetry!





Task 2: Strikeout

- Can we predict a coefficient at loop L from a set of related “parent” coefs at L-1?

$$\cancel{a}\cancel{a}cf = cf, \cancel{a}a\cancel{c}f = af, \cancel{a}ac\cancel{f} = ac$$

$$a\cancel{a}\cancel{c}f = af, a\cancel{a}c\cancel{f} = ac, aa\cancel{c}\cancel{f} = aa$$

  - {Input: list of loop L-1 coefficients in strikeout order} → {Output: loop L coefficient}
  - Remove duplicate values from dataset (since parents of dihedral images are dihedral images)
  - Training: 773500 training, 10k test examples.

- Giving a reduced set of parents (limiting strikeout distance), shuffling the parents, or giving only whether parents are {+,-,or 0} still allows reconstruction of the coefficients!

	Accuracy	Magnitude accuracy	Sign accuracy
Strike-two, all parents	98.1	98.4	99.6
Strike-two, $k = 5$	98.3	98.6	99.7
Strike-two, $k = 3$	98.4	98.7	99.7
Strike-two, $k = 2$	98.1	98.3	99.5
Strike-two, $k = 1$	94.3	95.2	98.5
Randomly shuffled parents, all parents	95.2	99.1	96.3
Randomly shuffled parents, $k = 2$	93.5	98.1	95.0
Sorted parents, $k = 5$	93.9	95.4	97.9
Parent signs only	93.3	93.5	99.0
Parent magnitudes only	81.8	98.4	83.2

Conclusion

- Transformers are able to learn some amount of the structure underlying planar N=4 SYM amplitude symbols
  - Though we're still trying to understand what properties are being learned by our models
- Hope to use these models (or models like them) to:
  - Uncover unknown recurrences in structure of scattering amplitudes
  - Predict symbol coefficients at as-yet-unseen loops
  - Shed further light on structure of planar N=4 SYM (antipodal duality?)
  - Expedite difficult theory calculations
- But this will almost definitely require a more sophisticated "next phase" problem formulation: new tasks, new data, new models, etc.

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

- Scattering Amplitudes and the Symbol Map

$$a = \frac{u}{vw}\,, \qquad b = \frac{v}{wu}\,, \qquad c = \frac{w}{uv}\,,$$
$$d = \frac{1-u}{u}\,, \qquad e = \frac{1-v}{v}\,, \qquad f = \frac{1-w}{w}$$

(*specifically, this is the two-loop term of the normalized BDS remainder function of the form factor)