Thoughts on Learning from LLMs

I have been reliably told that the "age of the textbook is over." If I want to learn about topic X, I should simply talk with an LLM about X. Just say "Hey Gork[1], help me understand X" and engage in dialogue with the machine god.

I've never had much success with this approach. When I have a *specific question* about a topic, I find that LLMs are excellent resources. Gork is regularly able to:

1. Understand what I'm getting at (even when I express the question imperfectly);
2. Quickly identify key gaps in my understanding;
3. Fill in the gaps efficiently.

This is excellent. Gork is the perfect generalist; he might not know everything cutting edge, but when it comes to content *in the canon*, so to speak, Gork can explain it all.

I have had much less success with *general questions* though. For example, I am currently working on a paper that applies some of my thesis work to the theory of graph neural networks. I have a decent understanding of classical neural networks (I can explain back propagation) and a deeper understanding the niche where my thesis work ties in. However, I don't have a great understanding of graph neural networks broadly. Asking Gork "Can you help me understand graph neural networks?" isn't very helpful.

I find it challenging to articulate exactly why and in what ways the resulting "general subject" conversation with Gork is so much less productive than asking specific questions. Instead of a comprehensive explanation, let me instead offer a few observations.

Background Context

A small and very specific problem I have is that Gork has learned a handful of topics that I am interested in (sheaf theory, algebraic topology, quantitative finance, type theory, etc.). In a misguided attempt to be helpful, Gork subsequently tries his best to link *every question I ask* to these subjects. While I find this dynamic to be endearing in a way, it is distinctly anti-helpful. At best, these links are real, but inessential for learning the basics of a subject:

> No, Gork. It probably isn't best to use GARCH volatility models as a first example of a regression analysis problem. Let's start with ordinary least squares.

At worst, these connections are completely hallucinated:

> No, Gork. The backpropagation algorithm isn't best thought of through the lens of category theory.

And somewhere in the middle, we have shoehorned analogies:

> Ok, Gork, I guess we can think of X as being "like a sheaf, but for ...," but I'd rather learn how practitioners think about X.

I imagine this problem could be solved through a combination of better prompting and system instructions. Indeed, I have reduced this problem 90% through these tools. However, I feel that I shouldn't have to resort to such brute force methods. Gork should understand that these connections to my interests aren't pedagogically productive.

[1] "Gork" can be any LLM, not just Grok.

Search Algorithms

Think of subject X as an abstract space of topics. This space can be represented as a network where the nodes V are (highly granular) **concepts** , and directed edges E represent **conceptual dependencies**. We further decorate this network with a collection of **topics**, where a topic $T \subseteq 2^V$ is a collection of concepts. If you like, you may think of this as a hypergraph structure where directed 1-edges are conceptual dependencies, and undirected hyper-edges encode topics.

Example 0.1. *As a simple example, in the abstract space representing regression analysis, we might have, among many many others:*

- *Nodes corresponding to the singular value decomposition (SVD), principal component analysis (PCA), ordinary least squares (OLS), and Fisher projections (FP).*
- *A directed edge from SVD to PCA.*
- *A topic called "dimensionality reduction" which contains the vertices SVD, PCA and FP, but not OLS.*

A more fundamental problem with learning a subject through LLM dialogue is that Gork's outputs force the dialogue into a sub-optimal search through this network structure. When I ask a question like:

> Can you help me understand graph neural networks? I understand feed forward neural networks already, and I would like to better understand graph neural networks.

I get a long output covering a variety of major topics, haphazardly ordered, at an extremely high level. When I ask for more detail on any given topic, I get what amounts to a list of sub-topics, again haphazardly ordered. This invites the following mode of dialogue:

Listing 1: Learning via depth first search

```
#Depth first search to learn a topic:
def DFS(T : Topic):
    subtopics = subtopics(T) #function that lists subtopics of a topic
    for S in subtopics:
        DFS(S) #learn each subtopic via DFS

#Run DFS on the subject X:
DFS(X)
```

There are two major problems with this approach. First, topic-level aggregation has no regard for the directed 1-edges that encode conceptual dependence. Foundational knowledge is at the bottom of each depth-first excursion. Moreover, depending on the whims of Gork, certain foundational knowledge might not be listed as a subtopic where it is needed. Second, I often find that DFS-ing through sub-topics hits my internal recursion-depth-limit before hitting foundational knowledge. Ironically, by teaching me high-level words without filling in the foundational understanding, I fear that Gork is training me to be a stochastic parrot.

I am not necessarily advocating for a fully-linear approach to learning. For mathematics (but also more generally), I think of learning through Vakil's tendrils of knowledge:

> Here's a phenomenon I was surprised to find: you'll go to talks, and hear various words, whose definitions you're not so sure about. At some point you'll be able to make a sentence using those words; you won't know what the words mean, but you'll know the sentence is correct. You'll also be able to ask a question using those words. You still won't know what the words mean, but you'll know the question is interesting, and you'll want to know the answer. Then later on, you'll learn what the words mean more precisely, and your sense of how they fit together will make that learning much easier. The reason for this phenomenon is that mathematics is so rich and infinite that it is impossible to learn it systematically, and if you wait to master one topic before moving on to the next, you'll never get anywhere. Instead, you'll have tendrils of knowledge extending far from your comfort zone. Then you can later backfill from these tendrils, and extend your comfort zone; this is much easier to do than learning "forwards."

A good textbook is non-linear, and keeps these tendrils of knowledge in mind. A depth first search through Gork's semi-random lists of subtopics are non-linear, but crafted without any regard for back-filling the tendrils: you just have to keep digging until you hit bedrock.

Creation vs. Curation

I think my difficulty learning a broad topic with Gork is related to a general dynamic in how people interact with generative AI: decreasing the cost of creation correspondingly increases the value of curation. In a sentence, having good taste is more valuable in a world full of slop.

Example 0.2 (AI art)**.** *Most AI art isn't very good. Moreover, AI art engines can't really tell which creations are good and which are bad. Good AI art, in my experience, is often made in the following two ways:*

1. *A human has a truly novel idea would be painfully difficult for a human artist to make. For example, look at this image. The local and global lighting and shading play against each other in a quasi-Escherian way. While it wouldn't be impossible for a human to create such an image without AI assistance, I'm not surprised that an AI made it first.*

2. *Second, a process of iterative refinement. A human asks an AI to make a picture, looks over dozens of slightly different creations, combines small details from a best subset, and iterates. By the end of the process, tens of thousands of individual creations have been made, and only the best is selected.*

Both of these processes involve the human playing the role of a curator; she selects, ideates, and refines, but she does not "make". There's a similar lesson here for some abstract art. You aren't Rothko — not because you can't match his brushwork — but because you can't match his taste[2]*.*

I think this dynamic also partially explains why so many non-professional *artists feel so threatened by AI art. It makes sense why professional artists feel threatened: the sword of Damocles is hanging over my head too. But for a non-professional artist, the threat seems more identity based. I think part of this ego-threat is that art-making-ability and taste, while certainly positively correlated, are distinct skills. AI art tools shift the coefficients in the quality function toward taste and away from brushwork.*

Example 0.3 (AI-driven research)**.** *In my experience LLMs are quite good at coming up with research ideas. Indeed, they are much better at coming up with research ideas (at present) than they are at following through on that research. This creates a similar curation vs. creation dynamic for human researchers. The process of selecting which research ideas to pursue and choosing how to allocate time between different projects are acts that demonstrate taste. This was a key insight for me when writing my own paper on LLM-assisted mathematics research.*

When it comes to teaching, Gork is extremely good at explaining concepts. He is clear, concise, and pedagogically gifted. He has near-infinite patience, will not hesitate to re-explain something for the thousandth time, and is more than happy to deep-dive into an appealing niche. What Gork lacks, in my opinion, is taste.

While I was writing my own book, figuring out how to order concepts in a way that balances both a top-down "this work is leading somewhere valuable" big-picture view with a suitably bottom-up "I actually understand why this is true" structure presented a serious challenge. Early drafts of chapters were completely deficient in one (or both) of these facets. This was a skill completely separate from those I have learned at the lectern. It takes good taste to select and order topics for a book, and I didn't have it. I'd like to think my taste has gotten much better through my book writing, but I still have room to improve. For example, I often relied on my co-author Robin for coming up with chapter "hooks."

When asking Gork to "teach me about X," he's in a similar boat to the one I was in. He simply doesn't have good taste for selecting and ordering topics. This may change in the future. I believe this will change in the future. But that's where he is now. Moreover, unlike picking generated artwork or research ideas, the learner who asking to be taught almost necessarily lacks the skill to curate from Gork's output.

[2] You also can't match his brushwork