Computer Synthesis

Tom Cunningham

2022-04-11

- See also: LLM backwards problems
- See also 2023-05-11-generative-ai-economics.qmd

Summary

1. **Arificial neural nets (ANNs) work by finding latent hierarchical structures in data.** If you feed them images they will learn to represent both low-level features (brightness, contrast, texture, line direction) and more abstract features (objects, light direction, style). If you feed them text they will learn to represent both low-level features (vocabulary, sentences, paragraphs, chapters) and more abstract features (tone, syntactical rules, logical structure).

   It is difficult to know how much of the performance of neural nets reflects them learning deep abstract representation vs just doing shallow permutations of the training data, in part because the training sets and the parameter spaces are both so large (about a trillion of each). Without knowing the depth of representation we can still observe that the outputs are useful for a broad range of practical tasks.

2. **It seems plausible that neural nets will find latent structures that humans haven't discovered yet.** They are good at discovering hidden structure: they might find a visual way of discriminating between male and female chicks (something that humans cannot do),

they might find a new assocation between diet and cancer.

3. **Generative models produce what is *probable* not what is *good.*** Many generative models (both text and image) can be thought of as generating data that is *likely* conditional on a prompt. Thus they will produce an output that has good qualities (logical, harmonious, perspicacious) just to the degree that the training data has those good qualities, but not because they are trying to achieve those good qualities. The typical training data does have good qualities compared to random data.

   If flaws in the training dataset are random (orthogonal to each other) then generative models can do better than the average human. If flaws in the training set are consistently in the same direction then generated content will reproduce the same flaws.

   The situation is more complicated for models that have been fine-tuned with human feedback, e.g. chatGPT which has been which is trained with an additional layer of explicit rating for quality (RLHF).

4. **Unlike humans, computers can use their representations to synthesize new artefacts.** Human judgment has a notable asymmetry: we can easily recognize when an object satisfies a given property, but it's much harder to synthesize a new object that satisfies that property. E.g. we can trivially tell whether a joke is funny, or a poem rhymes, or a picture looks realistic, but we find it difficult to create a funny joke, a rhyming poem, or a realistic picture. This is sometimes attributed to the hierarchical and feed-foward nature of how the brain processes information.

   Artificial neural nets do not suffer from the same degree of asymmetry. They can be reversed to synthesize new objects that satisfy a given property. This means computers will be able to do qualitatively different things that humans cannot do, even with the same amount of knowledge.

5. **The economics: it seems plausible that OpenAI would become the dominant firm.** The costs in AI are mostly fixed costs and so we might expect an equilibrium with a single dominant firm whose pricing power is limited by the threat of entry from other firms.[1] I think there is an analogy to the economics of web search where Google is the dominant firm, but is threatened by other potential entrants, and for that reason can only capture a small share of the total surplus generated. In most industries the first-mover retains their position for a long time so it seems plausible that OpenAI will become the dominant firm in AI. There also seem to be substantial economies of scope across use-cases, thus it doesn't seem likely that there would be different firms specialized for different use-cases (the same is true for search: it has not fractured into different markets).

[1] John Sutton's model of "endogenous sunk costs" seems a good fit: as the market gets larger the sunk costs increase, and so concentration does not fall.

6. **The effect on entertainment: platforms produce their own synthetic content.** It seems likely that computers could synthesize adequate-quality entertainment: stories, porn, music, television, video clips, computer games. The typical adult spends perhaps 5 hours/day on passive entertainment and much of it is fairly generic, meaning it could be generated by a computer and it would be similarly entertaining, and the audience wouldn't be upset. The firms likely to synthesize the best content would be those with access to large datasets on both content and engagement, e.g. Facebook, YouTube, TikTok, Netflix. Some platforms which are two-sided might become mostly one-sided because once they synthesize their own content they no longer need to attract creators. Platforms which host communication between real-life friends (e.g. FB, IG, WA) would retain a competitive edge because that content cannot be so easily displaced by synthetic content.

   In the short-run we could expect computers to produce content that is similar to existing content, but in the medium-term they might learn what features make it maximally engaging, in a similar way to how headline-writers have learned how to maximize peoples' propensity to click. As a consequence it's hard to anticipate how entertainment might evolve.

Peoples' tastes depend heavily on their experience: people like what they're familiar with. As a consequence we might expect that in the short-run synthetic content will resemble what people already consume, but in the long-run it could get very weird (just through the evolution of tastes, setting aside any technological advances).

7. **The effect on art: computers could make art that is better than humans can.** I think there's a compelling argument that computers will be able to make art that is *better* than humans can in certain ways. The success of an artwork is often judged by its ability to satisfy multiple different criteria simultaneously. For example a good picture both (a) depicts a specific scene, and (b) is a harmonious arrangement of colours. A good poem both (a) tells a story, (b) rhymes, and (c) scans. A good piece of music has symmetry at many levels. It's easy for humans to recognize when an existing artwork successfully satisfies multiple criteria but hard to create a new artwork that satisfies those criteria. Neural nets allow computers to learn to represent the characteristics of artworks, and so it seems plausible they would become better than humans in being able to find permutations that satisfy them.[2]

8. **The effect on programming: increased efficiency.** GPTs are already being used a lot by programmers, I think their primary uses are (1) saving time looking up knowledge, e.g. the names of function calls; (2) filling in boilerplate details.

It seems paradoxical that a short prompt can be reliably translated into a long piece of code because it implies that the generated code contains redundant information. Why would programming languages have so much redundancy? However there is an asymmetry between writing and reading code: it makes sense to write code with the details implicit (it saves time) but read code with the details explicit (it helps with debugging and confirming the logic).[3]

[2] Computers have been better than humans at solving well-specified combinatorial problems for 70 years. The difference with newer generations of algorithms is that they can now perceive categories that humans learn unconsciously: harmony, rhythm, representation, semantics, etc.

[3] Relatedly Djikstra argued that advances in mathematics were accompanied by formalizations of the notation (On the foolishness of "natural language programming",)

It seems likely that NN-augmented code would have a bigger relative effect on the productivity of high-skilled programmers than on low-skilled programmers (AKA skill-biased technical change).

9. **The effect on adversarial problems: help the bad guys more than the good guys.** In adversarial situations like spam-detection computers will help both spammers and spam-detectors, but they seem likely to shift the balance in favor of the spammers because of the relative advantage computers have at synthesizing over recognizing content (compared to humans). Humans can intuitively discriminate between real and fake artefacts, but that ability will become less valuable when computers have learned the cues that humans use to make those discriminations.

10. **The effect on communication systems: greater reliance on *provenance* than on *content*.** When evaluating a message we can use signals either from the content or the provenance (i.e. where the message come from). Many communication systems already rely heavily on provenance, e.g. spam detection is mostly based on the sender, not what is sent. However as generative ML improves we should expect equilibrium to move even further in the direction of provenance and reputation and away from content.

11. **The effect on employment.** LLMs are able to produce outputs that are near-indistinguishable from humans in many domains, so it seems likely it would displace humans.

    However previous computer technologies had much smaller effects on output and employment than were initially predicted: (1) in the 1950s many people thought optimization algorithms (e.g. linear programming) would greatly increase productivity but I think the effect was small; (2) in the 1970s many people thought that experts would be replaced by simple algorithms (e.g. linear regression), but I think the effect was small; (3) in the 1990s many people thought computers and the internet would greatly boost productivity, but I think the effect was small.

Predictions

Fraction of entertainment-time is synthetic in 2024	25%
Fraction of porn consumed is synthetic in 2024	50%

Media.

Illustration: mostly synthetic. Generative models can already create illustrations to the quality of an above-average advertisement or children's book. So it seems like that 100s of thousands of people who are working in illustration will find themselves replaced by algorithms. I would guess that demand for illustration is pretty inelastic, so increases in productivity decrease employment. 100 years ago mechanical reproduction of artworks decreased employment of musicians, painters, carvers, etc.

Precedents:

(A) the ability to reproduce artworks (printing press, colour reproductions, stamped decorations, recorded music) led to a substantial loss of employment.

(B) electronic synthesis of musical instruments probably led to unemployment of musicians.

Entertainment: mostly synthetic. People spend perhaps 10% of their waking life watching entertainment: TV sitcoms and soaps, funny videos on TikTok, Instagram and Facebook. It seems likely that 50% of this could be replaced by entirely synthetic content in 3 years. In the short run I guess ML will produce content that's indistinguishable from regular content but after a while it could get very weird.

Ads: mostly synthetic. These models could be used to create brand new ads designed to maximize propensity to click on them, or to buy the product. I think these could be quite powerful because (1) Facebook and Google already have databases with hundreds of millions of ads and trillions of interactions to learn from; (2) ML could experiment very efficiently to figure out what maximizes conversion rates.

In the past few years advertisers have been learning what makes people click: "1 simple trick for belly fat", "see this pimple pop", "see child stars now." Advertisers have always been doing they volume and granularity allows them to do it quicker. I expect ML accelerate it again, helping identify the sensitive parts of our brains which make us curious.

Spam: no dramatic change. We can define "spam" as advertising trying to disguise itself as organic content.

Over the last few decades spam has increased as the price of communication has fallen (SMS, email, social media).

Platforms have been able to keep it under control by looking for patterns in the *quantity* of messages: (a) detecting the same message being sent a million times; or (b) detecting that a million messages are all coming from a common source. New ML models will make it harder to distinguish based on content because it's cheap to make variations, though it will remain difficult to counterfeit the source.

Informational content: not sure. News, informative articles & videos, product reviews, celebrity gossip, discussion boards.

Taste: shift towards aesthetic minimalism. The ability to cheaply generate new artworks seems likely to change tastes. I believe something similar happened in the late 19th and early 20th century. In the 19th century it became much cheaper to add ornamentat onto objects, e.g. printed patterns on clothes, machine-woven fabric, stamped tinwork, moulded plaster. In the early 20th century ornamentation gradually fell out of fashion and there followed a general taste for simple minimalist objects which has mostly survived until today.

Effects on Entertainment and Art

Historical Trends in the Economics of Creation

I think it's useful to compare the introduction of synthetic composition, AKA generative ML, to three prior technological changes which affected creation:

1. The introduction of synthetic materials.
2. The mechanical reproduction of artworks.
3. Photography.

(1) The introduction of synthetic materials. Early humans would make everything from wood, stone, and animal skins. We then gradually learned how to use metal, glass, oil, plastic, etc., and we can now create almost any combination of shape, colour, texture, taste, smell.

Despite the freedom we have, many everyday products are still faux-wood, faux-stone, and faux-animal skin. We seem to have cultural preferences that are shaped by historical circumstances and which last many generations. Put another way: consumption choices at a given point in time seem highly influenced by budget constraints from prior points in time.

(2) The invention of mechanical reproduction. We have gradually invented many means of reproducing individual activities:

- 5000BC: cast metal into molds
- 3000BC: writing (reproduction of speech)
- 600BC: printing
- 1000s: printing with movable type
- 1800s: colour printing
- 1840s: photography
- 1890s: recorded sound and music
- 1900s: moving pictures
- 1930s: colour film
- 2020s: virtual reality

The equilibrium impact of mechanical reproduction depends on the preference for variety.[4] If people have a strong taste for

[4] Suppose that goods come in various different varieties and you can consume a certain quantity of each variety. We can say that before mechanical reproduction you would produce exactly one unit of each variety, and after mechanical reproduction you could produce multiple units of an existing variety at a cost lower than producing a new variety.

variety, meaning people strongly prefer an idiosyncratic object over an object that is identical to others, then mechanical reproduction would have little impact. Taste for variety could come either from a strict preference people have for differentiation (different from what they're familiar with or different from their neighbours), or from each person having idiosyncratic needs (the degree to which tailored clothing is more comfortable than off-the-rack).

Overall I think the preference for variety has not been sufficiently strong to prevent mass production from dramatically reshaping production and consumption.

It used to be that every family or village would have its own musician, story-teller, tailor, painter, cook, carpenter, blacksmith. Now most of the creative work is done by few guys in New York or LA and packaged up for the rest of us. They compose a song, a story, draw a picture, design a jacket or a piece of furniture, or come up with a recipe, which is then reproduced and shipped to every house in the world. There are a few professions remaining where people do one-off creative work: hairdressers, chefs, tattoo artists.

Although demand for artisans has substantially shrunk there are still perhaps surprisingly many. Most cities still have professional orchestras and opera companies, there are still many professional artists.

(3) Photography. Since photography was invented drawing and painting have become much less focussed on realism. Instead popular drawing and painting tends to be stylized in one way or another.

Observations on Taste

We live in a corner of the room We value things that are familiar.

Lowering cost will lower value. Walter Benjamin: mass production caused the individual work to have less value. Similarly we should expect synthesis to render the *type* of work less valuable, not just the individual work.

People like things that are superficially real. We watch a lot of reality shows, true-crime, documentaries. They are constrained to be factual, but people don't really care if they're actually true. Professional wrestling, celebrity drama (Kardashians). All in a world of pretend-reality.

People still go to concerts, despite record players.

People still want new music, despite old music existing.

Scarcity makes a thing more valuable. Gold, diamonds, pearls, purple robes.

Difficulty makes a thing more valuable. Elaborate embroidery, elaborate carving. Ornament was valuable for thousands of years. Fell out of favor and minimalism became ascendant as ornament became cheap. (upward-sloping demand, Veblen good.)

New technologies imitate the constraints of prior technologies. Synthetic glass, rubber, margarine, music.

How Neural Nets Work

Neural nets decompose a set of inputs into a hierarchy of latent components. E.g. when given a set of images neural nets will discover they are best organized by object and pose and lighting. When given a set of text inputs they will be organized into language and syntax and a set of semantic features. Early neural nets would output a single scalar prediction but once a net has learned a latent representation we can also reverse the process and find inputs that match the output, e.g. generating images to go with text, etc..

Neural nets work because the world is hierarchical.

They don't work as well when the data does not have a latent hierarchical structure, e.g. neural nets aren't particularly impressive with tabular or time-series data.

They don't work as well when we have a small dataset: with a small dataset a "structural" model will tend to do better,