UNIVERSAL COMPLEXITY CTR.

X@Solichorum

Perpetvvm@proton.me

August 25 2023

Feature Space Exploration, AI, Generative Models, Manipulation of Latent Variables, Systematic Approach to Machine Learning, Comparative Neurodynamics, Exploration of Latent Spaces, Machine Abstraction Techniques, Cognitive Drafting, Conceptual Engineering in AI, Linguistics in Early AI Development, AI Image Generation, Advanced Feature Representation, Universal Complexity Theory

How to Grab a Tidbitule

A Novel Approach to Manipulating Latent Variables in Generative Models

This method offers a systematic and efficient way to differentiate between otherwise likewise features in a consistent, provable, and repeatable manner.

The above quote could have added the phrase "expressions of" and used terms like "model parameters," "feature representations," or "learned features." These terms do not reflect the specificity of this method: we suggest calling one of these a "tidbitule." In our example, the tidbitule is one ball. This will be one of such formulations in a forthcoming glossary, reflecting the authors' goals in conceptual engineering around AI and comparative neurodynamics under the aegis of Universal Complexity Theory.
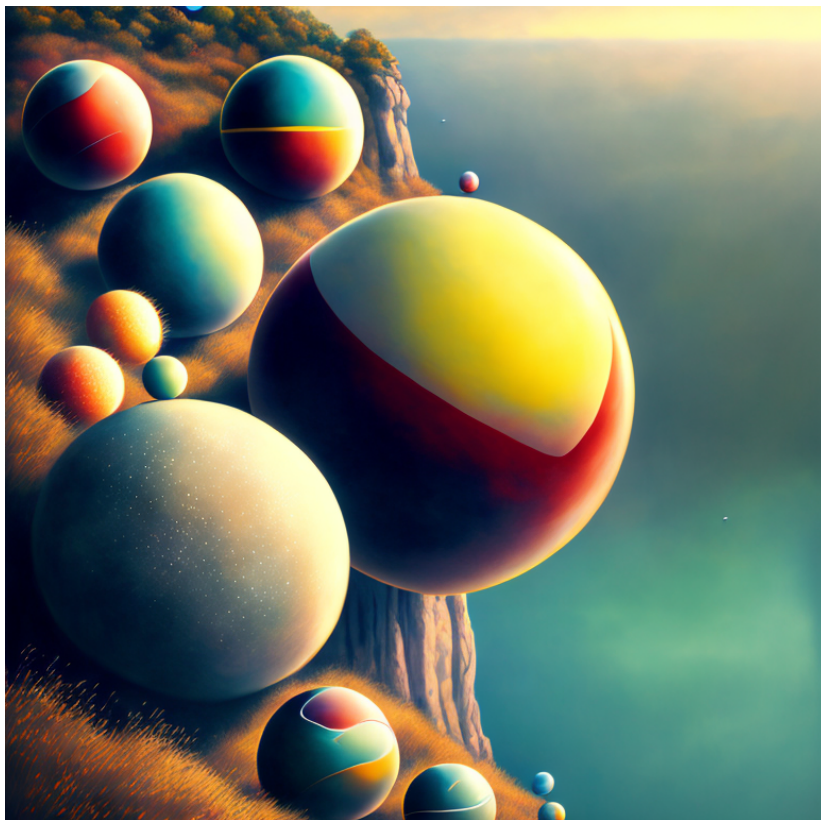

Fig II: 12balls.jpg

Preface

This paper presents a method for manipulating the latent variables in generative models, inspired by the problem of identifying a different ball among identical ones using a balance scale (Grossman, 1945). The method offers a systematic and efficient way to control the outputs of generative models, with potential applications in various areas of AI and machine learning. Such include adjusting the prompts to emphasize a target, manipulating the weights in the checkpoint model, merging models according to block weights, extracting learned representations, additional training of models, and version control for comparing updates. The method could be particularly useful for platforms like RunwayML, which could use it to guide its v3 model. Further implications are discussed after the fold.

Method:

The method begins by dividing the latent space into three groups. For example, one could divide the space based on the values of the latent variables, or based on the effects of moving along different directions in the space.

The first step of the method involves a comparison between two of the groups, analogous to weighing two groups of balls on the balance scale. In the context of the latent space, this comparison could involve measuring the effect on the output of the model when moving along the directions defined by the two groups. This could be done, for example, by generating images with the model and comparing the resulting images.

Depending on the outcome of this comparison, one of three scenarios will occur:

The two groups have the same 'weight', i.e., moving along the directions defined by the two groups has the same effect on the output of the model. In this case, the 'different ball', i.e., the direction in the latent space that we are interested in, is in the third group.

One group has a heavier 'weight' than the other, i.e., moving along the directions defined by one group has a greater effect on the output of the model than moving along the directions defined by the other group. In this case, the 'different ball' is either in the heavier group (if it is 'heavier' than the other balls) or in the lighter group (if it is 'lighter' than the other balls).

One group has a lighter 'weight' than the other, i.e., moving along the directions defined by one group has a lesser effect on the output of the model than moving along the directions defined by the other group. In this case, the 'different ball' is either in the lighter group (if it is 'lighter' than the other balls) or in the heavier group (if it is 'heavier' than the other balls).

In the second and third steps of the method, the process is repeated with the group that was identified as containing the 'different ball', until the tidbitule is identified.

This method promises to offer a systematic and efficient way to find specific expressions of features or effects in the latent space. It collapses the dimensional problem space into a single, guaranteed positive result.

To the best of our knowledge, this is the first presentation of the method in the literature.

Implications

Fig III: hands.jpg
https://www.bing.com/images/create/mr-hands/64922a4d28184999a953dc0a08cced93? Image of a human figure apalmé. Generated at bing.com/create on 20 Jun 2023.

Implications

Manipulating variables or "twiddling knobs" is the heart of creativity (Hofstadter, 1979) and this method offers one way to granularize this principle in machine systems. Translation to higher dimensions could reveal methods to manipulate quidditas (Aquinas, ~1270) in modular ways. Retroactive application in advanced systems could echo human processes that postfactually imbue experience with the light of meaning, then seen as willful sentience (Dennett, 1991). Echoing Dennett's term, we could call such retroapplication "machine cognitive drafting."

More immediately, let's continue to talk about images. See attached generated image of a person with too many fingers. With above method, we could grab a sixth finger and increase its expression. The reader may expect us to say this is an opportunity to remove a hallucination, but that would be acquiescent to descriptive frameworks that don't serve us at this level.

Remember the tidbitule: does a hand with six digits have five fingers and a hallucination? The distinction becomes absurd. In our view, it has six tidbitules. We'll take this opportunity to introduce terms that reveal the value-bind of the traditional vocabulary.

1. "Veritilla." A unit of output concordant with expectations. The "correct" five digits in a generated hand are veritillae.

2. "Fallaculus." Hallucination, sive "stuff we don't want." Sixth and seventh digits are fallaculi. (The classically-minded may object that "fallaculus" is inappropriately gender-creative; such concerns here were balanced toward euphony and intuitive pluralization by the many.)

Wielding these words, we have a clearer view of what's going on--we see it's a matter of judgment. But what if we have a hand with four digits? There's a place where we think a finger ought be. We'll name that too:

3. "Latent void." The emptiness where a latent space should be, ie a missing finger.

This term isn't relegated to the scale of the tidbitule. In our estimation, it could find more use describing higher missing pieces. For example, if we asked for a picture of a cat, and it gave us a realistic photo of a hairless breed, one could see that as an expression of a latent void, since our human latent spaces for cats include associations like "fur, fuzzy, soft."

Implications, abbreviated: It can remove a sixth finger, and you might be able to make abstractive reasoning with it.

Morgan Corrigan

The Universal Complexity Center

References:

KwegiboHB. (2023, June 8). On the Magic that is Weights and Biases. Something Awful Forums.

Grossman, H (1945). Scripta Mathematica XI.

Hofstadter, Doug (1979). Gödel, Escher, Bach: An Eternal Golden Braid. Basic Books.

Aquinas, Thos (~1270). Summa Theologica, I-III,4.

"Quiddity or essence signifies that whereby a thing is what it is; and this belongs to the definition of a thing, which signifies what a thing is."

(Quidditas vel essentia significat id quo res est quod est; et hoc pertinet ad rationem definitionis, quae significat quid est res.)

Dennett, Dan (1991). Consciousness Explained. Little, Brown and Co.