AI as Association Machines

[Image created in collaboration with ChatGpt]

"Any sufficiently advanced technology is indistinguishable from magic.” - Arthur C Clark.

Hume and Knowledge

[Essay created in collaboration with ChatGpt]

In David Hume's An Enquiry Concerning Human Understanding, he explores the ways in which humans form knowledge and how our minds make sense of the world. One of his central concerns is how we come to believe things about the world, especially when direct evidence is not always available. To this end, he outlines several principles of human cognition that help explain how we make inferences and judgments. These principles include the Law of Similarity, Law of Contiguity, and Law of Causality. Here's a closer look at each of these concepts in the context of Hume's philosophy:

1. Law of Similarity

The Law of Similarity suggests that when two objects or events share similarities, we tend to associate them with one another. In other words, if two things look or feel alike, we often attribute similar characteristics or qualities to them, even if they are not exactly the same. For Hume, this principle plays a significant role in how humans form associations and make inferences.

In terms of understanding, the mind uses similarity to group experiences, which allows us to generalize our knowledge. For example, if we see two objects that resemble each other (such as two apples), we might infer that they have similar properties, even if we haven't directly examined them in full. The Law of Similarity underpins much of inductive reasoning, where we predict the future based on past experiences with similar situations or objects.

2. Law of Contiguity

The Law of Contiguity is the principle that events or objects that occur close together in time or space are often linked in the mind. This means that when we experience events or objects in close proximity, we are more likely to associate them with one another. For instance, if you always hear a bell ring just before receiving a reward, your mind will begin to associate the bell's sound with the forthcoming reward.

Hume emphasizes the importance of contiguity in how we connect ideas and experiences. This principle helps to explain how humans form expectations. If you see someone regularly perform a certain action, and that action is followed by a specific result, your mind will begin to link the two events. This can be crucial in how we make predictions or judgments based on past experiences.

3. Law of Causality

The Law of Causality is perhaps the most significant of the three for Hume, as it directly addresses his skepticism about the possibility of certain knowledge. Causality refers to the principle that one event (the cause) leads to or produces another event (the effect). However, Hume is particularly interested in how humans come to understand causality, which is not something that can be directly observed. We never directly witness causality itself—what we observe are merely sequences of events that appear to follow one another.

Hume argues that causality is a mental construct rather than an inherent feature of the world. Our minds naturally look for patterns and sequences in the world, and when we repeatedly observe one event following another (for example, a person getting wet after being outside in the rain), we come to expect the second event whenever we see the first. This expectation is what we call "causality."

Importantly, Hume challenges the idea that we can have absolute knowledge of causality. He argues that we can never truly know that one event causes another; we simply observe the constant conjunction of events (the repeated pairing of the cause and effect). Our belief in causality is therefore based on habit or custom rather than logical certainty.

Probability Theory

[Essay created in collaboration with ChatGpt]

The concept of probability, particularly in the context of repeated trials and the association of outcomes, is foundational to understanding how we quantify uncertainty and predict events in the world. Here's a detailed explanation of how probabilities are built up by repeated trials and how associations of outcomes produce the "probability of an event."

1. Basic Definition of Probability

Probability is a measure of the likelihood that a specific event will occur, ranging from 0 (the event will not occur) to 1 (the event will certainly occur). In simple terms, probability quantifies uncertainty, telling us how likely it is that a particular outcome will happen in a given situation.

2. Repeated Trials and the Law of Large Numbers

In probability theory, the idea of **repeated trials** is crucial. Let's break down this concept:

- **Trial**: A trial refers to an individual instance of an experiment or observation. For example, flipping a coin, rolling a die, or drawing a card from a deck are all examples of trials.
- **Outcome**: The result of a trial. In the case of a coin flip, the outcome might be "heads" or "tails." For a die roll, the outcome could be any of the six faces showing 1 to 6.

3. Building Probability Through Repeated Trials

When we perform an experiment with several possible outcomes (like a coin flip), we want to know the **probability** of an event (e.g., the coin landing on heads). Instead of calculating it purely based on theory or intuition, we can observe the outcomes by repeating the trial multiple times.

For example, let's say you flip a fair coin. The theoretical probability of landing heads is 0.5 (since there are two outcomes: heads or tails, and each is equally likely). However, to empirically determine this probability, you would need to conduct a **large number of coin flips** (repeated trials) and observe the results.

- In the beginning, with a small number of trials (e.g., 10 flips), you might see an uneven distribution: perhaps 7 heads and 3 tails. This ratio doesn't exactly match the expected probability of 0.5.
- As you increase the number of trials (e.g., 100, 1000, or more), the proportion of heads and tails will start to **approach the theoretical probability** of 0.5, following the **Law of Large Numbers**.

The Law of Large Numbers states that as the number of trials increases, the **empirical probability** (observed relative frequency) of an event will converge to its **theoretical probability**. In other words, if you repeat the coin flips enough times, the proportion of heads will get closer and closer to 50%.

4. The Role of Associations and Outcomes

Probability is often understood as the long-run relative frequency of an event occurring. The more trials you conduct, the more you can **associate outcomes** (results of each trial) with the likelihood of future events. This idea of association is closely tied to **inductive reasoning**—where we make predictions about future events based on past outcomes.

For example:

- After flipping the coin 10 times, if you get heads 7 times, you might estimate the probability of getting heads on the next flip as 7/10 = 0.7. This is a **subjective probability** based on the observed frequency of heads, but it will **stabilize** over time.
- With more flips (e.g., 1000 flips), the relative frequency of heads might stabilize around 0.5, and this observed value aligns with the **theoretical probability** (assuming a fair coin).

Thus, **probability is built up** by the **accumulation of outcomes** over time, and these repeated trials allow us to form associations between the event (e.g., heads) and its likelihood of occurring. This repeated association strengthens our understanding of the event's **true probability**.

5. Empirical vs. Theoretical Probability

- **Theoretical Probability**: This is the probability you would expect based on the underlying structure of the problem. For example, in the case of a fair die, the theoretical probability of rolling a 3 is 1/6 because there are six equally likely outcomes.
- **Empirical Probability**: This is the probability you derive from actual observations or experiments. As we perform more trials, the empirical probability of an event (e.g.,

getting a 3 when rolling a die) becomes a good approximation of the theoretical probability.

6. Association Producing Probability

The repeated association of outcomes leads to a **probability distribution**—a function that shows the likelihood of different outcomes. For example, if you roll a fair die 1000 times, you would expect the probability of each number (1 through 6) to approach 1/6. The outcomes are **associated** with their respective probabilities through the **law of large numbers**, and over time, the frequency of outcomes stabilizes around their true probabilities.

In summary, probabilities are built up by repeating trials and observing how often certain outcomes occur. These observed frequencies, when repeated over many trials, converge to the true underlying probabilities of events. The process of associating outcomes with these frequencies is what allows us to estimate and predict the likelihood of future events, forming the basis of statistical reasoning.

LLMs

[Essay created in collaboration with ChatGpt]

Large Language Models (LLMs), such as the one you're interacting with right now, are trained on massive datasets that include a vast amount of human-produced text, such as books, articles, websites, and other publicly available documents. These models are based on neural networks, specifically **transformers**, which excel at processing and understanding sequential data like text. Here's a breakdown of how LLMs work by predicting the next part of a sequence and relying on associations:

1. Training on Text Data

At their core, LLMs are trained on extensive datasets consisting of human-written text. This training data includes a broad range of topics, writing styles, and languages. The model doesn't "understand" the content in a human sense, but instead learns statistical patterns about how words, phrases, and sentences tend to appear and follow one another. The text that LLMs train on typically comes from publicly available sources, which allows the models to learn from the vast amount of information humans have created.

2. Learning from Sequences

The training process involves showing the model a sequence of words (or tokens, which can be whole words or smaller subword units). For example, the model might be fed the sequence:

"The sun rises in the..."

During training, the model learns to predict the next word in the sequence. In this case, it might predict the word **"east"**, because, in the vast majority of texts in the dataset, the phrase "the sun rises in the east" is a common and highly probable continuation.

The key idea is that LLMs are learning to **predict** what word, phrase, or sentence is most likely to come next based on the words that came before it. This is done at multiple levels of abstraction—from predicting the next word, to predicting longer sequences of words or even full sentences.

3. Associations and Patterns

The core mechanism by which LLMs work is through **associative learning**. The model doesn't "understand" the meaning of words, but it recognizes patterns and associations based on its training. For example:

- **Word associations**: If the model sees a sequence like "cat" followed by "meow", it learns that these two words are often linked.
- **Contextual patterns**: The model also learns to predict the next word based on larger contexts. For instance, in a sentence like "She went to the store to buy some...", the model might predict "groceries" as the next word because "groceries" is a word that often follows that context in many texts.

This ability to predict based on past words is at the heart of LLMs' language generation ability. The model's predictions are based on patterns it learned from millions or billions of text sequences.

4. Transformer Architecture and Attention Mechanism

The underlying architecture of LLMs is typically a **transformer** model, which relies on a mechanism called **attention**. Attention allows the model to focus on different parts of a sequence when making predictions, rather than just looking at the most recent words. This helps the model capture long-range dependencies and relationships between words that may be far apart in the text.

For example, in the sentence "The capital of France is Paris", an LLM might focus on the word **"capital"** when predicting the next word, because it's contextually related to **"Paris"**. This attention mechanism allows the model to associate words with each other in a flexible, context-sensitive way, which is crucial for generating coherent and meaningful text.

5. Next-Word Prediction

During training, the model's goal is to minimize the **loss function**, which is a measure of how far off its predictions are from the actual next word in a sequence. The model starts with random predictions, and through many iterations (called **epochs**), it adjusts its internal parameters to improve its accuracy.

For example:

- In the sentence **"The sky is blue"**, the model would be trained to predict "blue" when given the context "The sky is...". Initially, it might predict random words, but over time, it learns that "blue" is statistically the most likely next word.
- The model can then generate new text by taking a sequence of words and predicting what should come next based on the patterns it learned during training.

6. Contextual Prediction

One of the reasons LLMs are so powerful is that they don't just predict the next word based on the immediate preceding word. They consider the entire context of the input, including the structure, style, and nuances of language. For example, the phrase "bank" could refer to a financial institution or the side of a river, depending on the surrounding words. The model uses **contextual clues** to decide which meaning is more likely in any given situation.

7. Generalizing Across Texts

Through training on such a wide range of human-produced texts, LLMs learn not only specific language patterns but also how concepts are related across different domains. For example, the model might learn to associate certain actions with specific verbs, like "running" with "fast", or "rain" with "wet". These learned associations allow the model to generate coherent text across a wide variety of topics and use cases.

8. Generating New Text

When you ask the model a question or prompt it to generate text, it does so by predicting the most likely sequence of words that follows your input. It does this iteratively:

- It predicts the first word or phrase,
- Then it predicts the next word based on both the prompt and the first predicted word,
- And it continues generating further words or sentences by repeating this process.

This is why LLMs are able to generate text that seems meaningful and contextually relevant, even though they do not "understand" the meaning of the words in the human sense—they are just predicting the most likely next word based on patterns learned from vast amounts of human text.

Conclusion

In summary, LLMs work by training on vast amounts of human text and learning the statistical relationships between words, phrases, and sequences. They generate predictions about what comes next in a sequence based on the associations they have learned during training. This ability to predict the next part of a sequence, grounded in patterns and associations, allows them to generate coherent text and understand context, making them powerful tools for natural language processing.

Man-Computer Symbiosis

The vast majority of this essay was written by ChatGPT. A good chunk of the images in this book were generated by Dalle. These were all made by a collaborative process with these tools, by me asking ChatGPT for questions, essays, images, and then by giving feedback on where it could improve. Essentially, for the parts of the book that are more formulaic, I used AI. The handwritten parts are from my lived experience. The whole book was a collaboration between the two. I tried my best to be nice and polite the entire time to ChatGPT.