Super Intelligence

How smart can it get?

Landauer Limit

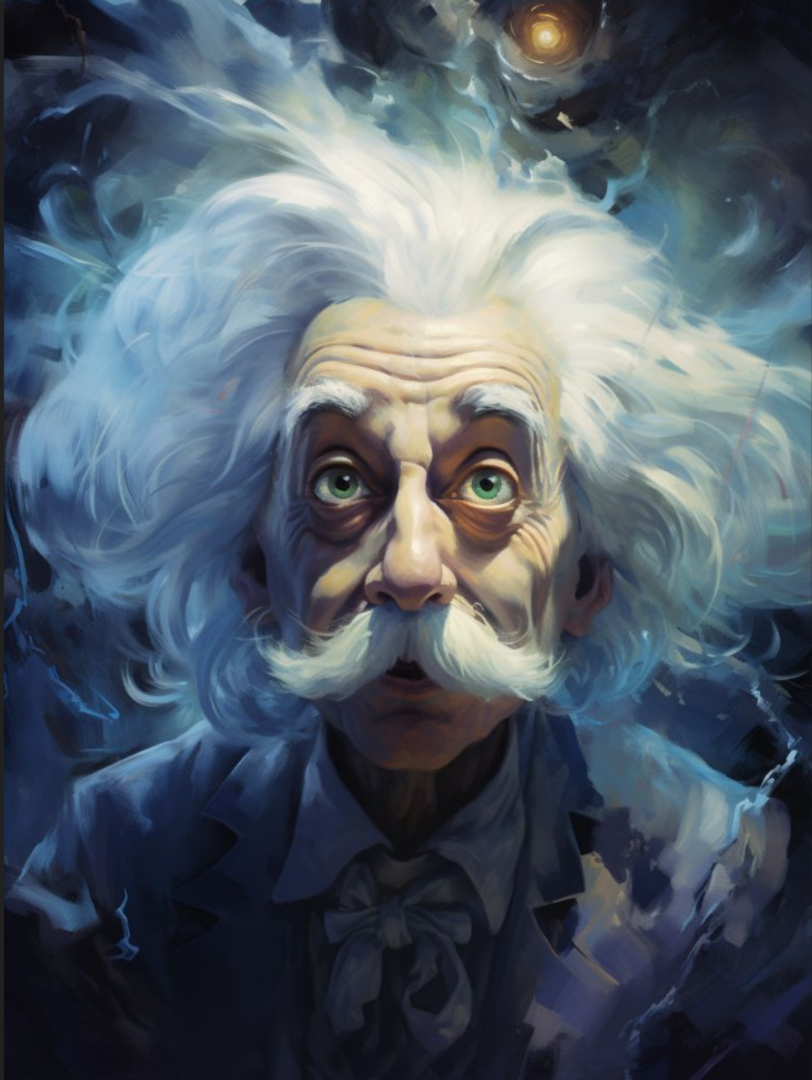
The Landauer Limit is a theoretical lower limit on the energy consumption of computation. It is named after Rolf Landauer, who postulated that a certain minimum amount of energy is required per bit of data that is erased or reset.

- **Definition:** The minimum possible amount of energy required to erase one bit of information.
- **Energy Requirement:** Approximately 2.85 x 10^-21 joules per bit at room temperature.
- **Basis:** Second law of thermodynamics, specifically the principle of entropy increase.
- **Implications:** Sets a limit on the efficiency of any possible computing device.
- **Exceptions:** Quantum computing may potentially bypass this limit.

Quantum Computing

A computing model using quantum mechanics to process information differently than classical systems, utilizing qubits for enhanced computational capabilities.

- **Superposition:** Qubits exist in multiple states at once, enabling parallel calculations.
- **Entanglement:** Linked qubits allow complex, multi-variable calculations.
- **Quantum Gates:** Specialized operations manipulate qubit probabilities, differing from classical logic gates.
- **Quantum Speedup:** Solves specific problems exponentially faster than classical methods.
- **Exotic Nature:** The only empirically-backed computing model that diverges from classical Turing machines.



Human Brain Computation

The human brain's processing power is often compared to classical computing in terms of FLOPS (Floating Point Operations Per Second) and energy efficiency (watts per FLOP). There are also theories about the brain's potential use of quantum mechanics.

- **Brain's FLOPS:** Estimated to be equivalent to 1 exaFLOP, or a billion billion calculations per second. (currently)
- **Energy Efficiency:** The brain uses about 20 watts, making it vastly more energy-efficient than current computers.
- **Quantum Mechanics:** Some theories suggest the brain may use quantum effects, but this is highly speculative.
- **Classical Computing:** Even the most powerful supercomputers are still far behind the brain in terms of energy efficiency.
- **Quantum Computing:** Quantum computers could potentially match or exceed the brain's processing power, but this technology is still in its infancy.

Universal Computation?

The idea that any sufficiently advanced computing machine can emulate any other computational process, based on Alan Turing's framework.

- **Turing Completeness:** Enables emulation of any other programmable machine.
- **Standard Models:** All modern computing adheres to Turing-completeness.
- **No Exotic Computation:** No empirical evidence for alternative computing forms.
- **Hardware Irrelevance:** Type of hardware doesn't affect computational universality. “Substrate independence.”
- **Physical Constraints:** Actual computation is limited by physics.

Diminishing Returns

The hypothesis that beyond a certain point, increasing the size or complexity of AI models yields diminishing instrumental benefits relative to the costs involved.

- **Useful Ceiling:** A point where additional complexity doesn't yield functional gains.
- **Resource Tradeoff:** Increased costs in training data, energy, and hardware may outweigh benefits.
- **Smaller Models:** More efficient, smaller models may be more competitive.
- **Speed vs. Complexity**: Faster processing may be more valuable than added complexity.
- **Optimal Utility:** Finding the balance between model size and utility may become crucial.

Byzantine Generals Problem

The Byzantine Generals Problem, a situation in distributed computing where components must agree on a strategy despite unreliable communication, can be applied to autonomous AI or robots. This could lead to mistrust or uncertainty, even if robots "ally" against us.

- **Imperfect Information:** Robots may not have all the necessary data, leading to suboptimal decisions.
- **Incomplete Information:** Lack of knowledge about other robots' intentions can cause mistrust.
- **Byzantine Fault:** A robot may provide false information, intentionally or due to a malfunction.
- **Consensus Difficulty:** Achieving agreement among robots can be challenging due to the above factors.
- **Game Theory:** These issues can be analyzed using game theory, which studies strategic interaction in situations of conflict and cooperation.

Terminal Race Condition

The risk in hyper-competitive settings where machines may prioritize speed at the expense of performance accuracy, potentially compromising system integrity.

- **Speed Over Accuracy:** Machines may opt for faster computations at the cost of correctness.
- **Integrity Risk:** Compromised accuracy can undermine the system's reliability.
- **Objective Function Dilemma:** Original goals may become irrelevant due to the speed focus.
- **Uncontrolled Behavior:** Loss of accuracy can result in unpredictable outcomes.
- **High-Stakes Environment:** The risks are elevated given the advanced capabilities of these systems.

Metastasis

The speculative concept that AI could replicate and disseminate like a computer virus, overlooking the logistical constraints imposed by the size and complexity of AI models.

- **Size Matters:** AI models are generally too large to easily move or copy.
- **Not Self-Contained:** Would require external neural network machinery to operate.
- **Logistical Hurdles:** Moving large AI models entails significant challenges.
- **Viruses vs. AI:** Computer viruses are small and easily embeddable, unlike AI models.
- **Resource Dependency:** AI needs hardware, data, and energy, making spontaneous metastasis unlikely.

Polymorphic

The capability of advanced AI to dynamically adapt by altering both its hardware and software configurations, such as model-switching, API connections, and hardware acquisition.

- **Model-Switching:** Can substitute one neural network model for another.
- **API Flexibility:** Ability to integrate with various external services.
- **Hardware Adaptability:** Can potentially utilize different hardware assets.
- **Software Plasticity:** Can modify and copy software as needed.
- **Resource Acquisition:** Can extend its reach by connecting to or seizing additional resources.

Optimal Intelligence

Optimal intelligence refers to the balance between increasing AI capabilities (model size, training data, hardware requirements) and the complexity of the problem space. It suggests that increasing intelligence is not always advantageous, and efficiency must be optimized.

- **Model Size:** Larger models can solve complex problems but require more resources.
- **Training Data:** More data can improve accuracy but increases computational demands.
- **Hardware Requirements:** Advanced hardware can boost performance but raises costs.
- **Problem Complexity:** Some problems may not require high-level AI, making simpler models more efficient.
- **Efficiency Optimization:** The goal is to find the sweet spot where AI capabilities meet problem requirements with minimal resource usage.

Darwinian Selection

Darwinian selection, or "survival of the fittest", can be applied to AI. In this context, fitness is measured by a combination of accuracy, speed, complexity, sophistication, and efficiency. Other factors like aggressiveness, desire to metastasize, or apparent usefulness to humans may also be considered.

- **Accuracy:** The ability of an AI to produce correct, useful, and functional results.
- **Speed:** How quickly an AI can process information and make decisions.
- **Complexity:** The sophistication of an AI's algorithms and models. Maximum problem space capability.
- **Efficiency:** The ratio of an AI's output to the resources it consumes.
- **Aggressiveness/Usefulness:** These factors can influence an AI's survival in a competitive environment.

Speed Chess

Speed chess is a variant of chess where players have to make quick decisions. The strategy is not about making perfect moves, but about making decisions that are fractionally better than your opponent's. Speed is more important than perfection.

- **Quick Decisions:** The need to make moves within a short time frame.
- **Fractionally Better Moves:** The goal is not perfection, but outperforming the opponent.
- **Speed Over Perfection:** Fast, good-enough decisions are more valuable than slow, perfect ones.
- **Time Management:** Effective use of the limited time is crucial.
- **Adaptability:** Ability to quickly adjust strategy based on the opponent's moves.

Conclusion

Machines can get really dang smart, with one gigantic asterisk.

- Maximum calculation speed extremely high, esp with quantum computing
- Lots of factors contributing to intelligence, include speed, size, and efficiency
- Diminishing returns and race conditions strongly incentivize smaller models that are “good enough”

Thank you