Energy-based Modeling Approach Enables Reasoning entirely from Unsupervised Learning

note by David Sauerwein, 7/20/25

A new energy-based modeling approach enables reasoning entirely from unsupervised learning. This is an exciting push to break free from major constraints of today's reasoning models, with their narrow scope and reliance on external rewards, toward more data-efficient and generalizable models.

Human thinking is classified into System 1 (intuitive, fast) and System 2 (slow, deliberate reasoning). Current transformers excel at System 1 but struggle with System 2. Recent advances using reinforcement learning or test time computation are impressive but are still restricted to domains with easily verifiable rewards (math, programming).

To create systems that truly think independently, we need approaches that ideally rely entirely on unsupervised learning for System 2 thinking.

Particularly, they should address these three facets of human thinking that current LLMs lack:

1. Dynamic compute allocation: Adjusting computational effort to problem complexity. For example, humans contemplate career transitions much longer than lunch decisions.

2. Modeling uncertainty: Humans weigh uncertainty before committing to decisions. Quantifying this uncertainty is central to complex reasoning.

3. Verification of predictions: Verification is central to 1. and 2. Above. Moreover, verifying solutions is typically also easier than generating them. So, learning a verifier could be more data efficient and robust. However, current LLMs don’t naturally integrate verifiers, and creating them for domains that are hard to quantify (e.g. rate of a conversation) remains challenging.

Researchers have now proposed a new paradigm to address these challenges (link in comments).

They propose viewing thinking as optimization with learned verifiers that evaluate input-output compatibility. More precisely, they train energy-based transformers (EBTs) to learn energy landscapes where lower energy indicates higher compatibility.

Thinking then starts from random predictions and refines through energy minimization until convergence. Since the optimization duration depends on problem complexity, this enables dynamic compute allocation (facet 1). The energy values quantify uncertainty (facet 2) and serve as verifiers (facet 3).

Training energy-based models is notoriously hard to scale, but the researchers show how transformer properties (scalability, robustness, parallelizability) transfer to EBTs.

The results show EBTs achieve up to 35% higher scaling rates (across e.g. data, parameters) and 29% improved reasoning performance versus vanilla transformers. They superior scaling can probably be traced back to the fact that the EBT has also learned to verify, not only predict.

Of course, many questions remain before declaring this a new architectural breakthrough. EBTs are more complex to train, and scaling beyond 800M parameters is unclear.

But this is truly exciting work. I'm keen to see how people push this approach forward.

A diagram of energy landscape

Description automatically generated

# References

[1] [Original Linkedin post, David Sauerwein, 07/19/2025](https://www.linkedin.com/posts/davidsauerwein_ai-genai-agi-activity-7352576343349280768-77N3?utm_source=share&utm_medium=member_desktop&rcm=ACoAAAFZfUoBgPoGUucdnvtwuzPv79P8VHj6uvk)

[2] [Energy-Based Transformers are Scalable Learners and Thinkers, Alexi Gladstone et al, UVA, UIUC, Amazon, 2025](https://github.com/dimitarpg13/large_language_models/blob/main/articles/energy-based_methods/Energy-Based_Transformers_are_Scalable_Learners_and_Thinkers_Gladstone_2025.pdf)