A Small Math Model: Recasting Strategy Choice Theory in an LLM-Inspired Architecture

Roussel Rahman
Stanford University
Stanford, CA 94305
roussel.rahman@gmail.com

Jeff Shrager
Bennu Climate, Inc.
Palo Alto, CA 94301
jshrager@gmail.com

Abstract

Strategy Choice Theory (SCT)[1] (e.g., Siegler and Shrager, 1984; Siegler, 2000) explains important aspects of children's arithmetic learning based upon principles including learning from developmentally naturalistic data, probabilistic representation, confidence-based retrieval, and the phase-like importance of scaffolding strategies, such as finger-counting. Here we recast SCT as a "Small Math Model" (SMM), employing a neural-network-based architecture analogous to LLMs. The SMM extends SCT to include counting practice[2], symbol (number) embedding, and gated attention. Similar to earlier work, the SMM demonstrates constructive and destructive interference between counting and addition, and the "wave-like" use of finger-counting as sum recall improves. We plan to extend the SMM to later aspects of the decades-long SCT program, including adaptive strategy choice and eventually strategy discovery, providing a unified platform to investigate the understanding of numerical characteristics and relationships essential for mathematical reasoning – as it can emerge in LLM-based agents.

Keywords: Developmental Modeling, Transformers, Mathematical Reasoning, Strategy Choice

1 Introduction

In a very real sense, young children participate in reinforcement learning from human (usually parent) feedback. Early theories of children's learning posited processes such as "stages" to account for observed differences in children's cognitive skills. This view was overthrown by the work of Robert Siegler and his coworkers (e.g., Siegler and Shrager, 1984; Siegler and Shipley, 1995; Siegler, 1988, 2000) who proposed the "Strategy Choice Theory" (SCT) hypothesizing that cognitive development is a continuous process wherein the child learns to adaptively select from multiple strategies, such as counting, retrieval, finger-counting, and so on, depending upon a set of factors including the strength of association between problem and answer, the child's confidence, and past experience with each strategy. The SCT framework emphasizes mechanisms that align well with the probabilistic underpinnings, uncertainty estimation, and fallback mechanisms of modern AI. Indeed, LLMs have revolutionized natural language processing by scaling these same principles.

In the present work, we recast the earliest of the SCT models as a "Small Math Model" (SMM) within a contemporary LLM-inspired architecture. In undertaking this work, we have several goals. First, we hope to be able to unify the models that comprise Siegler's early math learning project in an architecture that provides a natural explanation for seemingly diverse phenomena, including learning to count, add, or other arithmetic, learning new strategies, and learning to use strategies in

---

[1] "Strategy Choice Theory", "Distributions of Associations", and "Overlapping Wave Theory" have been used to refer to this line of work, emphasizing different aspects.

[2] The original SCT model was pre-biased in accordance with the supposed experience of counting.

appropriate settings. Second, we hypothesize that the SMM approach can lead to improved "number common sense" in LLMs, something that has recently been noted as lacking (e.g., Rahman and Mishra, 2025). With similar goals, Hein and Diepold (2024) examined how multimodal transformers acquire skills on a set of numerical tasks (such as enumeration, set comparison, and seriation). They are, however, focused on learning from visual and linguistic data and integrating these modalities, but not on learning the numerical symbols, operations, and their internal representations of magnitudes, which form a crucial part of human number sense, alongside visual and linguistic representations.[3]

SCT has long been used to explain how humans develop number sense (Dehaene, 1992), making the SMM an ideal testbed for investigating and improving number sense in AI agents. While modern XAI methods, such as mechanistic interpretability, have yielded insights into how neural networks solve mathematical problems by rebuilding solution strategies from elements (e.g., Wang et al., 2022; Nanda et al., 2023), existing approaches are largely observational and difficult to scale (Bereska and Gavves, 2024). In contrast, the SMM offers a controllable environment for studying complex reasoning, with clear ground truths about the problem environment and model capabilities.

2 Strategy Choice Theory

In complex environments, learners have access to a variety of problem-solving strategies. SCT explains how individuals select among these strategies based on the distribution of associations between problems and potential answers. For example, children learning to add or multiply do not memorize facts (e.g., $2 + 3 = 5$, $2 \times 3 = 6$) but, like an LLM, build a probabilistic "Distribution of Associations" between problems and possible answers. Early on, these associations are noisy, as children might associate $2+3$ with answers 3 (counting up from 2), 4 (counting from 3, or $2 \to 3$), 5 (the correct answer), or even other results that they may have encountered through errors early on in their learning to add. With experience, correct associations strengthen while incorrect ones weaken. The SCT is formalized with three mechanisms: (1) representation, (2) solution generation (through retrieval or fallback strategies), and (3) learning. When retrieval confidence falls below a threshold, the agent resorts to backup strategies and uses it as a learning opportunity to update associations.

3 Architecture

The SMM processes input sentences of the form "a $operator$ b" where $a, b \in \{1, \ldots, 5\}$ and $operator \in \{\to (count\ up), + (add)\}$. The symbolic numerical arguments (i.e., the numbers) are "embedded" or mapped to a dense vector in a continuous space, enabling the model to capture emergent relationships among numbers (for example, relative magnitude, even versus odd, or common configurations such as base-five hand groupings). These embeddings feed into a recurrent processing core that implements a mixture of strategies, including finger counting and direct addition.[4] A gating ('quasi-attention', QA) mechanism learns how much attention to pay to the operands, conditioned on the operator, so that it emphasizes different input features in different contexts. Faced with a counting problem, it learns to amplify the representation of the current number in the sequence (b), facilitating stepwise incrementing, whereas faced with an addition problem, it learns to highlight both operands simultaneously. The embeddings interact with the attentional gate, enabling the model to capture the relational structure of numbers vis-à-vis strategies, and vice versa.

Like the 1984 SCT, the SMM includes two strategies for addition: direct recall and finger-counting. It also includes a mechanism for learning to count up, through either recall or "oracular" knowledge (being told, presumably by a parent), which was not modeled in the 1984 work. Answer confidence for either counting or addition is estimated from the entropy of the output distribution. When confidence falls below a threshold, the model defers to an alternative strategy for addition, either being told (for counting) or finger-counting, which generates the correct result step by step. Importantly, the SMM uses its counting skills as a scaffold for learning addition skills. The finger-counting strat-

---

[3]The triple code model (Dehaene, 1992) explains that humans represent numbers in three forms: (1) visual (e.g., 31), (2) verbal ("thirty-one"), and (3) their analog magnitudes along a number line, each corresponding to different parts of the brain (Dehaene, 1996; Kiefer and Dehaene, 1997; Pinel et al., 1999).

[4]While the network itself is feed-forward, the training regime and task structure make it functionally recurrent as its own outputs are fed back as inputs across steps to implement counting and addition.

egy relies on the ability to count up, which is learned by the model itself! The subsequent answer is then used for training in the usual back-propagation way. This is analogous to the way LLMs can be augmented with external tools (e.g., for coding) or explicit reasoning mechanisms (e.g., chains-of-thought, Wei et al., 2022 or tree-of-thoughts, Yao et al., 2023) and subsequently distilled back into the base model.

4 Training

Training follows a Gaussian curriculum schedule that gradually increases task difficulty. Early in training, simpler problems (e.g., sums with smaller operands) dominate; over time, harder problems appear. This smooth progression mirrors natural learning trajectories while avoiding abrupt jumps. It also follows learning practices in LLM training, where data distributions are staged over time to improve stability and generalization.

Counting is trained from the very beginning of each run, while addition can be introduced at different time points to investigate how prior knowledge influences subsequent learning. After each problem, strategy-specific confidence values are updated according to correctness, and these values govern stochastic selection of strategies on subsequent trials. This architecture, therefore, supports mechanistic analysis of how new procedures compete with, reinforce, or interfere with existing ones in the course of skill acquisition. The code for the SMM is available here (under MIT Open Source License): https://github.com/jeffshrager/smm

5 Results

Prior experience with counting can either scaffold or hinder the acquisition of addition, depending upon the timing of addition's introduction. Figure 1 shows changes in addition accuracy with training (Fig. 1a) and in the use of the finger-counting strategy (Fig. 1b) for addition problems. We varied the points in training at which addition problems were introduced. In all cases, we see that the change-points in accuracy closely align with the ones in finger-counting usage. The accuracy curves begin with plateaus (of varying lengths) before a rapid rise; during this period, finger-counting usage rises to a maximum, indicating that the model becomes more confident that finger-counting is the more promising strategy and that improved addition skills emerge from improved counting skills. Then follows another brief plateau, followed by a gradual rise to near-perfect accuracy. During this period, finger-counting usage briefly reached several local maxima, reflecting the continued use of finger-counting as a learning strategy, before a gradual decline to zero, marking the end of finger-counting as a scaffold for addition.
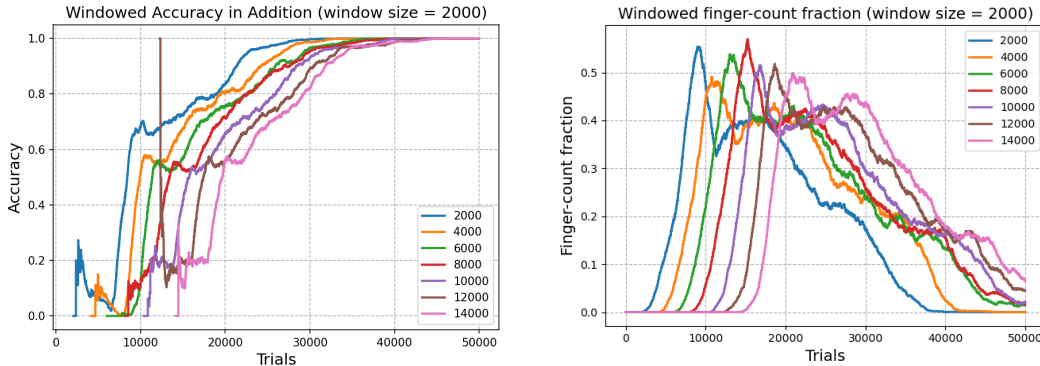

Figure 1: Addition correctness (a) and use of finger-counting (b) as a function of training step, for different addition-start times. Counting is always trained from the beginning. Destructive interference takes place near the beginning of practice if addition is introduced too early, although this is rapidly overcome by training.
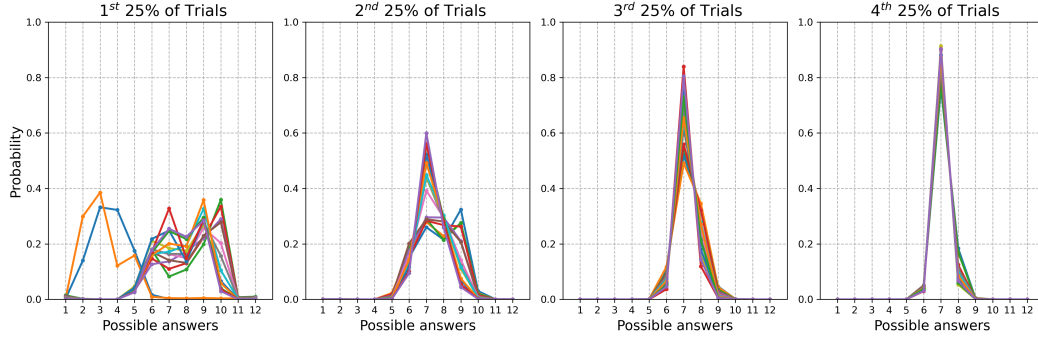
Figure 2: Addition answer confidence over the whole 50,000 training cycles, for the problem $3+4 = ?$, depicted in four quarters, early through late. In each quarter, the lines represent distributions for problems sampled once in every 10 instances. (The training itself is not divided into four periods.)

The change with learning in addition answer confidence is depicted in figure 2 for a specific problem: $3 + 4 = ?$. Note that early on the possible answers are biased to incorrect results due to addend-bias (for the problem $3 + 4$ tend to say 3 or 4), counting bias (tend to say 5, from the counting sequence $3 \rightarrow 4 \rightarrow 5$), and errors in executing finger-counting due to poor counting training (i.e., this many answers around, but not at 7, rather at $6, 8, 9, ...$). Notice as well that confidence is distributed among these various answers. Yet, with experience, these biases fall off, and the correct answer (7) becomes nearly the only likely answer with nearly 1.0 confidence.

6   Discussion and Next Steps

The SMM modernizes and extends the SCT (Siegler and Shrager, 1984; Siegler, 1988) recasting it in a contemporary LLM architecture that includes embeddings, gating/attention, cross-entropy loss, curriculum learning, and tool (strategy) use. The simple SMM architecture provides a testbed for studying how counting priors, confidence estimation, and quasi-attention (in the LLM sense) interact in a realistic yet controlled setting. We show that the SMM can build upon its counting skills to learn addition, using counting as a scaffold during early training. The transition point—when counting becomes redundant for addition—is clearly identifiable and aligns with predictions from Strategy Choice Theory (SCT), emerging naturally from the model's confidence thresholds. Moreover, the evolution of confidence distributions over possible answers is interpretable and reveals how biases (e.g., addend or counting biases) diminish with experience. These dynamics not only validate SCT's core mechanisms but also demonstrate how modern neural architectures can yield transparent, mechanistic insights into the development of symbolic reasoning.

While our demonstrations focused on problems involving numbers 1 through 5, the SMM architecture is readily scalable to larger numerical ranges and additional operations (e.g., subtraction, multiplication, division), with computational cost being the primary constraint—two directions we aim to pursue to investigate how complex mathematical skills emerge from foundational numerical abilities. Importantly, adaptive strategy competition can give rise to nontrivial transfer effects. In the early part of learning addition, the SMM falls back on its counting skills for solutions and updates associations between problems and answers. When addition problems were introduced too early, the SMM had difficulty using finger-counting to scaffold weak addition skills. An important implication of these results is the potential value of "progressive deepening" (e.g., Fayek et al., 2020; Hacohen and Weinshall, 2019), also recognized in cognitive science as the "starting small" hypothesis (Elman et al., 1996). By analogy, staged curricula in LLMs may help establish core abilities—such as number sense—before introducing higher-level reasoning, supporting more robust common sense in numerical domains.

The SMM may further serve as a testbed to probe number sense, which is fundamental to mathematical reasoning. The embedding-based architecture has the potential to capture latent properties (e.g., even/odd or base-five hand configurations, and many others). Indeed, our immediate next steps will be to provide a number of strategies that vary in their effectiveness – combining efficiency and

correctness – for different sorts of addends.[5] We expect number embeddings to be able to learn the relationships between number properties and strategy effectiveness.

The SMM illustrates how modern neural implementations can instantiate and extend classical cognitive theories. By reframing Siegler and Shrager's model (and eventually the rest of the Siegler line of research) in a modern architecture, we hope to obtain a platform for mechanistic analysis of symbolic reasoning in tractable domains, opportunities to probe representations directly (e.g., via similarity analysis or attention visualization), and a rich bridge between cognitive science and AI, highlighting shared principles of probabilistic selection based on confidence estimation, and adaptive strategy use.

Acknowledgments

References

Bereska, L. and Gavves, E. (2024). Mechanistic interpretability for ai safety–a review. *arXiv preprint arXiv:2404.14082*.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44(1-2):1–42.

Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of cognitive neuroscience*, 8(1):47–68.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press, Cambridge, MA.

Fayek, H. M., Srivastava, N., and Jin, J. (2020). Progressive learning: A deep learning framework for continual learning. *Frontiers in Artificial Intelligence*, 3:7.

Hacohen, G. and Weinshall, D. (2019). On the power of curriculum learning in training deep networks. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *ICML '19*, pages 2535–2544. PMLR.

Hein, A. and Diepold, K. (2024). Exploring early number abilities with multimodal transformers. *Cognitive Science*. First published: 03 September 2024.

Kiefer, M. and Dehaene, S. (1997). The time course of parietal activation in single-digit multiplication: Evidence from event-related potentials. *Mathematical cognition*, 3(1):1–30.

Nanda, N., Chan, L., Lieberum, T., Smith, J., and Steinhardt, J. (2023). Progress measures for grokking via mechanistic interpretability. *arXiv preprint arXiv:2301.05217*.

Pinel, P., Le Clec'H, G., Van de Moortele, P.-F., Naccache, L., Le Bihan, D., and Dehaene, S. (1999). Event-related fmri analysis of the cerebral circuit for number comparison. *Neuroreport*, 10(7):1473–1479.

Rahman, R. and Mishra, A. A. (2025). A fragile number sense: Probing the elemental limits of numerical reasoning in llms. *arXiv preprint arXiv:2509.06332*. Submitted September 8, 2025.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of experimental psychology: General*, 117(3):258.

Siegler, R. S. (2000). The rebirth of children's learning. *Child Development*, 71(1):26–35.

Siegler, R. S. and Shipley, C. (1995). Variation, selection, and cognitive change. In Simon, T. J. and Halford, G. S., editors, *Developing cognitive competence: New approaches to process modeling*, pages 31–76. Lawrence Erlbaum Associates, Hillsdale, NJ.

---

[5]Efficiency and correctness are partly correlated because strategies that require more steps have a greater likelihood of error, although there are other factors at play as well.

Siegler, R. S. and Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In *Origins of Cognitive Skills*. Erlbaum.

Wang, K., Variengien, A., Conmy, A., Shlegeris, B., and Steinhardt, J. (2022). Interpretability in the wild: a circuit for indirect object identification in gpt-2 small. *arXiv preprint arXiv:2211.00593*.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., Zhou, D., et al. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T., Cao, Y., and Narasimhan, K. (2023). Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems*, 36:11809–11822.