Novelty Assessment Report

Paper: Transformers are Inherently Succinct
PDF URL: https://openreview.net/pdf?id=Yxz92UuPLQ
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-27

Abstract

We propose succinctness as a measure of expressive power of a transformer in describing a concept. To this end, we prove that transformers are highly expressive in that they can represent formal languages substantially more succinctly than standard representations of formal languages like finite automata and Linear Temporal Logic (LTL) formulas. As a by-product of this expressivity, verifying even simple properties of transformers is shown to be provably intractable (i.e. EXPSPACE-complete).

Core Task Landscape

This paper addresses: **Succinctness of transformers in representing formal languages**

A total of **38 papers** were analyzed and organized into a taxonomy with **17 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Expressive Power and Formal Language Recognition**
- **Depth and Architectural Constraints**
- **Succinctness and Representational Efficiency**
- **Learning Dynamics and Learnability**
- **Pre-training on Formal and Procedural Data**
- **Automata Extraction and Interpretation**
- **Practical Applications**
- **Surveys and Meta-Analyses**
- **Critiques and Philosophical Perspectives**

Complete Taxonomy Tree

- Succinctness of transformers in representing formal languages Survey Taxonomy
- Expressive Power and Formal Language Recognition
  - Equivalence to Formal Language Classes ★ (6 papers)
  - [0] Transformers are Inherently Succinct (Anon et al., 2026) View paper
  - [8] Characterizing the Expressivity of Fixed-Precision Transformer Language Models (Jiaoda Li, 2025) View paper
  - [14] Transformers as recognizers of formal languages: A survey on expressivity (Lena Strobl, 2023) View paper
  - [15] Transformers as recognizers and transducers (Strobl, 2025) View paper
  - [20] Masked Hard-Attention Transformers Recognize Exactly the Star-Free Languages (Dana Angluin, 2023) View paper
  - [21] Characterizing the Expressivity of Transformer Language Models (Li, 2025) View paper
  - Recognition Capabilities for Specific Language Families (4 papers)
  - [1] Transformers Can Represent -gram Language Models (A Svete, 2024) View paper
  - [5] Exploring Attention Patterns and Neural Activations in Transformer Architectures for Sequence Classification in Context Free Grammars (Molinolo, 2024) View paper
  - [29] Transformers as CFG Learners: A Formal Framework for Structured Data (Ogbuokiri, 2025) View paper
  - [31] On the Ability and Limitations of Transformers to Recognize Formal Languages (Satwik Bhattamishra, 2020) View paper
  - Weighted Automata and Sequential Models (2 papers)
  - [17] Simulating Weighted Automata over Sequences and Trees with Transformers (Lacroce, 2024) View paper
  - [19] The Power of Hard Attention Transformers on Data Sequences: A Formal Language Theoretic Perspective (Pascal BergstrÃ¤Ã er, 2024) View paper
- Depth and Architectural Constraints
  - Depth Hierarchies and Layer Effects (2 papers)
  - [2] Knee-Deep in C-RASP: A Transformer Depth Hierarchy (Cadilhac, 2025) View paper
  - [38] Depth-Width Tradeoffs for Transformers on Graph Tasks (G Yehudai, n.d.) View paper
  - Attention Mechanisms and Hard Attention (1 papers)
  - [6] Formal language recognition by hard attention transformers: Perspectives from circuit complexity (Angluin, 2022) View paper
  - Fixed-Precision and Numerical Constraints (2 papers)
  - [9] Transformer encoder satisfiability: Complexity and impact on formal reasoning (SÃ¤lzer, 2024) View paper
  - [12] Computational Expressivity of Neural Language Models (Alexandra Butoi, 2024) View paper
- Succinctness and Representational Efficiency
  - Recurrent Extensions and Linear Transformers (1 papers)

    - [27] Practical Computational Power of Linear Transformers and Their Recurrent and Self-Referential Extensions (CsordÃ¡s, 2023) View paper
- Learning Dynamics and Learnability
    - Sensitivity and Loss Landscape Constraints (1 papers)
    - [30] Why are Sensitive Functions Hard for Transformers? (Hahn Michael, 2024) View paper
    - Grammar Learning and Induction (1 papers)
    - [13] How Well Can BERT Learn the Grammar of an Agglutinative and Flexible-Order Language? The Case of Basque. (Gorka Urbizu, 2024) View paper
- Pre-training on Formal and Procedural Data (2 papers)
    - [16] Between Circuits and Chomsky: Pre-pretraining on Formal Languages Imparts Linguistic Biases (Hu, 2025) View paper
    - [25] Can You Learn to See Without Images? Procedural Warm-Up for Vision Transformers (Zachary Shinnick, 2025) View paper
- Automata Extraction and Interpretation (1 papers)
    - [4] Automata extraction from transformers (Zhang Yi-hao, 2024) View paper
- Practical Applications
    - Code Generation and Compilation (2 papers)
    - [11] Transformers are Efficient Compilers, Provably (Zhai, 2024) View paper
    - [23] GPT-Based Wasm Instruction Analysis for Program Language Processing (Liangjun Deng, 2024) View paper
    - Grammar Enforcement and Parsing (5 papers)
    - [32] The Grammar Correction: A Comparison of T5, LLAMA and ChatGPT (JA Nurhasanah, 2025) View paper
    - [33] Enforcing Grammar in Code Synthesis with Transformers (Dmytro Vitel, 2023) View paper
    - [35] Improving Formal Reasoning of Transformer with State Stack (K Zhang, n.d.) View paper
    - [36] Grammar-enforced Chain of Thought Reasoning for small LLMs (S Babu, n.d.) View paper
    - [37] A transformer-based parser for Grammatical Framework (AV Velasco, n.d.) View paper
    - Compact Models for Low-Resource Languages (2 papers)
    - [7] Lightweight transformers for conversational ai (Chen, 2022) View paper
    - [26] Compact Transformer-based Language Models for the Moroccan Darija (Mohamed Aghzal, 2023) View paper
    - Specialized Applications (4 papers)
    - [3] TAYSIR Competition: Transformer+\textscrnn: Algorithms to Yield Simple and Interpretable Representations (R Eyraud, 2023) View paper
    - [10] Explainability and Semantics-Bridging Natural Language Flexibility and Formal Precision: Toward a Semantic Framework for Large Language Models (Bilokon, 2024) View paper
    - [22] Inferring Group Intent as a Cooperative Game. An NLP-based Framework for Trajectory Analysis using Graph Transformer Neural Network (Zhang Yiming, 2025) View paper
    - [24] Expression transformers in B-GSL (Bill Stoddart, 2003) View paper
- Surveys and Meta-Analyses (2 papers)
    - [28] What Formal Languages Can Transformers Express? A Survey (Angluin, 2023) View paper
    - [34] ESSLLI 2024 Course Notes Expressivity of Transformers: Logic, Circuits, and Formal Languages (D Chiang, n.d.) View paper
- Critiques and Philosophical Perspectives (1 papers)
    - [18] Understanding models understanding language (Anders SÃ¸gaard, 2022) View paper

Narrative

Core task: Succinctness of transformers in representing formal languages. The field examines how efficiently transformer architectures can encode and recognize formal languages compared to classical automata. The taxonomy reveals several major branches: one focuses on expressive power and equivalence to formal language classes, asking which languages transformers can recognize and how they compare to finite automata or context-free grammars; another investigates depth and architectural constraints, exploring trade-offs between model size and representational capacity; a third addresses succinctness and representational efficiency, quantifying how compactly transformers encode languages relative to traditional models. Additional branches cover learning dynamics, pre-training strategies on formal data, automata extraction methods for interpretability, practical applications, and broader surveys. Works like Transformers Formal Languages Survey[14] and Formal Languages Transformers Survey[28] provide overarching perspectives, while studies such as Transformers Recognizers Transducers[15] and Attention Patterns Context Free[5] establish formal connections between transformer components and classical language hierarchies.

A particularly active line of inquiry concerns the precise characterization of transformer expressivity: some works demonstrate equivalences to specific automata classes under various precision or attention constraints, as seen in Fixed Precision Expressivity[8] and Masked Attention Star Free[20], while others like Characterizing Transformer Expressivity[21] offer broader frameworks. The original paper, Transformers Inherently Succinct[0], sits squarely within the branch on equivalence to formal language classes but emphasizes a representational efficiency angle—arguing that transformers can encode certain languages more compactly than classical models. This contrasts with neighboring works such as Attention Patterns Context Free[5], which focuses on structural correspondence to context-free grammars, and Transformers Recognizers Transducers[15], which explores transduction capabilities. Together, these studies highlight an ongoing tension between proving what transformers can represent in principle versus understanding the resource costs of such representations, a central open question in the field.

Related Works in Same Category

The following **5 sibling papers** share the same taxonomy leaf node with the original paper:

1. Characterizing the Expressivity of Fixed-Precision Transformer Language Models

Authors: Jiaoda Li, Ryan Cotterell | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Transformer-based language models (LMs) have achieved widespread empirical success, but their theoretical expressive power remains only partially understood. In this work, we analyze a restricted idealization of fixed-precision transformers with strict future masking, soft attention, and no positional encodings. We establish that this class of models is exactly as expressive as a specific fragment of linear temporal logic that contains only a single temporal operator: the past operator. We furth...

Relationship Analysis

Both papers belong to the 'Equivalence to Formal Language Classes' category, establishing exact correspondences between transformer variants and formal language hierarchies. They share overlapping focus on fixed-precision transformers and their relationship to formal languages, particularly star-free languages and temporal logic fragments. The key difference is that the original paper emphasizes succinctness (how compactly transformers represent languages compared to automata/LTL) and computational complexity, while the candidate paper focuses on characterizing the exact expressive equivalence between restricted transformers and a specific LTL fragment (past operator only), with empirical validation of generalization behavior.

---

2. Transformers as recognizers of formal languages: A survey on expressivity

Authors: Lena Strobl, William Merrill, Gail Garfinkel Weiss, David Chiang, Dana Angluin | **Year/Venue**: 2023 | **URL**: View paper

Abstract

As transformers have gained prominence in natural language processing, some researchers have investigated theoretically what problems they can and cannot solve, by treating problems as formal languages. Exploring questions such as this will help to compare transformers with other models, and transformer variants with one another, for various tasks. Work in this subarea has made considerable progress in recent years. Here, we undertake a comprehensive survey of this work, documenting the diverse ...

Relationship Analysis

Both papers belong to the category of establishing exact correspondences between transformer variants and formal language hierarchies. They overlap in examining the expressive power of transformers relative to formal language classes like star-free languages, finite automata, and temporal logic. The original paper focuses specifically on succinctness—showing transformers can represent languages exponentially more compactly than LTL and doubly exponentially more compactly than automata—while the candidate paper provides a comprehensive survey of expressivity results across multiple transformer variants and formal models without emphasizing succinctness measures.

---

3. Transformers as recognizers and transducers

Authors: L Strobl | **Year/Venue**: 2025 | **URL**: View paper

Abstract

â¦ formal language theory, circuit complexity, and transductions. It develops a unified framework that locates finite-precision transformer â¦ It also characterizes trade-offs in succinct models: â¦

Relationship Analysis

Both papers belong to the category of establishing exact correspondences between transformer variants and formal language hierarchies, specifically focusing on the expressive power of transformers in representing formal languages. The original paper investigates succinctness of transformers by proving they can represent star-free languages exponentially more compactly than LTL and doubly exponentially more compactly than finite automata, while the candidate paper (a doctoral thesis) provides a broader survey and framework that places finite-precision transformers within circuit complexity classes (AC0, TC0) and establishes their equivalence to specific formal language classes and transduction hierarchies. The key difference is that the original paper emphasizes succinctness gaps and computational hardness (EXPSPACE-completeness), whereas the candidate paper focuses on comprehensive characterization of transformer expressivity through multiple lenses (circuits, logic, transductions) and includes empirical learnability studies.

---

4. Masked Hard-Attention Transformers Recognize Exactly the Star-Free Languages

Authors: Dana Angluin, David Chiang, Andy Yang | **Year/Venue**: 2023 | **URL**: View paper

Abstract

The expressive power of transformers over inputs of unbounded size can be studied through their ability to recognize classes of formal languages. In this paper, we establish exact characterizations of transformers with hard attention (in which all attention is focused on exactly one position) and attention masking (in which each position only attends to positions on one side). With strict masking (each position cannot attend to itself) and without position embeddings, these transformers are expr...

Relationship Analysis

Both papers belong to the 'Equivalence to Formal Language Classes' category, establishing exact correspondences between transformer variants and formal language hierarchies. They overlap in proving that masked hard-attention transformers recognize star-free languages (equivalent to LTL and counter-free automata), using similar techniques involving Boolean RASP as an intermediate representation. The key difference is that the original paper focuses on succinctness results (showing transformers can be exponentially more compact than LTL/automata) and computational complexity (EXPSPACE-completeness), while the candidate paper emphasizes exact characterizations and explores how architectural variations (position embeddings, masking strictness, depth) affect expressive power within the star-free class.

---

5. Characterizing the Expressivity of Transformer Language Models

Authors: Li, Jiaoda, Cotterell, Ryan | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Transformer-based language models (LMs) have achieved widespread empirical success, but their theoretical expressive power remains only partially understood. Prior work often relies on idealized models with assumptions -- such as arbitrary numerical precision and hard attention -- that diverge from real-world transformers. In this work, we provide an exact characterization of fixed-precision transformers with strict future masking and soft attention, an idealization that more closely mirrors pra...

Relationship Analysis

Both papers belong to the same category of establishing exact correspondences between transformer variants and formal language hierarchies. They overlap in analyzing fixed-precision transformers and their equivalence to star-free regular languages, with both connecting transformers to Linear Temporal Logic (LTL) and automata classes. The original paper focuses on succinctness—showing transformers can represent languages exponentially more compactly than LTL and doubly exponentially more compactly than automata—while the candidate paper characterizes the expressive power of soft-attention transformers as exactly equivalent to LTL[P] (a restricted fragment using only the past operator) and partially ordered DFAs, without emphasizing succinctness comparisons.

Contributions Analysis

Overall novelty summary. The paper proposes succinctness—how compactly a transformer represents a concept—as a measure of expressive power, proving that transformers can encode formal languages exponentially or doubly exponentially more compactly than finite automata or Linear Temporal Logic formulas. It resides in the 'Equivalence to Formal Language Classes' leaf, which contains six papers total, including the original work. This leaf sits within the broader 'Expressive Power and Formal Language Recognition' branch, indicating a moderately populated research direction focused on formal characterizations rather than practical applications or learning dynamics.

The taxonomy reveals neighboring leaves examining 'Recognition Capabilities for Specific Language Families' (four papers on context-free grammars and counter languages) and 'Weighted Automata and Sequential Models' (two papers on sequential reasoning). The 'Succinctness and Representational Efficiency' branch exists as a separate top-level category but contains only one paper on recurrent extensions, suggesting the original paper's focus on succinctness comparisons with classical models occupies relatively sparse territory.

The scope notes clarify that equivalence studies exclude depth hierarchies and practical implementation, distinguishing this work from architectural constraint analyses in adjacent branches.

Among thirty candidates examined, the succinctness measure itself (Contribution 1) and the exponential/doubly exponential gaps (Contribution 2) each faced ten candidates with zero refutations, suggesting these representational efficiency claims are relatively novel within the limited search scope. The EXPSPACE-completeness result (Contribution 3) encountered one refutable candidate among ten examined, indicating some prior complexity analysis exists but the specific verification problem formulation may differ. The statistics reflect a focused semantic search rather than exhaustive coverage, so these findings characterize novelty relative to the most semantically similar thirty papers, not the entire field.

Based on the limited search scope and taxonomy structure, the work appears to occupy a distinctive position emphasizing representational efficiency over pure expressiveness characterizations. The sibling papers in the same leaf focus more on equivalence proofs and language class boundaries, while the succinctness angle bridges to complexity theory in a way that neighboring works do not extensively explore. However, the single refutable candidate for the complexity result suggests some overlap with prior verification analyses, warranting careful comparison in a full review.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Succinctness as a measure of transformer expressive power

Description: The authors introduce succinctness—the smallest descriptional size needed to recognize a language—as an alternative measure of expressiveness for transformers. This measure captures how compactly transformers can represent concepts compared to other formalisms.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Towards Structured Intelligence for Sequence Modeling

URL: View paper

Brief Assessment

Structured Intelligence Sequence[61] uses 'succinct' only in passing to describe diabetes severity measures, not as a formal measure of expressive power for neural sequence models or transformers.

---

2. LooperGP: A loopable sequence model for live coding performance using GuitarPro tablature

URL: View paper

Brief Assessment

LooperGP Live Coding[62] focuses on controllable music generation using transformers for live coding performances, not on formal language theory or succinctness as a measure of expressive power for transformers.

---

3. Explainable and Efficient Knowledge Acquisition from Text

URL: View paper

Brief Assessment

Explainable Knowledge Acquisition[64] discusses succinctness in the context of explanations and information capture, not as a formal measure of expressive power for transformers or neural sequence models. The candidate does not address succinctness as a descriptional size measure for language recognition.

---

4. Saturation in Recurrent Neural Networks: Expressivity, Learnability, and Generalization.

URL: View paper

Brief Assessment

Saturation Recurrent Networks[58] focuses on expressivity, learnability, and generalization of RNNs (particularly SRNs and LSTMs) using saturation concepts, not on succinctness as a measure of expressive power for transformers or neural sequence models.

---

5. Quantum long short-term memory

URL: View paper

Brief Assessment

Quantum LSTM[56] focuses on quantum implementations of recurrent neural networks for sequence modeling, not on succinctness as a measure of expressive power for transformers or any formal language theory analysis.

---

6. Consistent Bidirectional Language Modelling: Expressive Power and Representational Conciseness

URL: View paper

Brief Assessment

Consistent Bidirectional Modeling[63] focuses on succinctness of bidirectional language models (specifically bisequential decompositions and latent language models) compared to unidirectional models, not transformers. The paper measures conciseness in terms of state space size for automata-theoretic models, which is a different formalism and context than the original paper's transformer-specific succinctness measure.

---

7. DNS-Rec: Data-aware Neural Architecture Search for Recommender Systems

URL: View paper

Brief Assessment

Neural Architecture Search Recommender[59] focuses on neural architecture search for recommender systems with resource constraints, not on formal language theory or succinctness measures for transformer expressiveness.

---

8. Compacting, picking and growing for unforgetting continual learning

URL: View paper

Brief Assessment

Compacting Picking Growing[57] focuses on continual learning through model compression and expansion mechanisms for sequential tasks, not on measuring transformer expressiveness through succinctness or formal language theory.

---

9. Low-rank passthrough neural networks

URL: View paper

Brief Assessment

Low Rank Passthrough[65] focuses on low-rank matrix parametrizations to reduce parameters in RNNs/LSTMs while maintaining memory capacity, not on succinctness as a formal measure of expressiveness for transformers or language recognizers.

10. A Lightweight Sequential Convolutional Neural Network for Smart Grid Stability Analysis

URL: View paper

Brief Assessment

Sequential CNN Grid Stability[60] focuses on convolutional neural networks for smart grid stability analysis, not on measuring expressive power of transformers through succinctness or formal language theory.

Contribution 2: Exponential and doubly exponential succinctness gaps

Description: The authors prove that transformers can represent languages exponentially more succinctly than Linear Temporal Logic and Recurrent Neural Networks, and doubly exponentially more succinctly than finite automata. This demonstrates that transformers encode complex patterns with significantly smaller descriptional sizes.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Practical Computational Power of Linear Transformers and Their Recurrent and Self-Referential Extensions

URL: View paper

Brief Assessment

Linear Transformers Computational Power[27] focuses on the computational power and expressiveness of linear transformers (LTs) and their extensions for formal language recognition, not on succinctness gaps between transformers and other models like RNNs or finite automata.

2. Rethinking transformers for efficiency and scalability

URL: View paper

Brief Assessment

Rethinking Transformers Efficiency[51] focuses on architectural modifications for computational efficiency and scalability. The provided context contains only fragmentary text about transformer architecture and hybrid models, with no discussion of formal language theory, succinctness gaps, or comparisons with finite automata, LTL, or RNNs.

3. Autoregressive+ Chain of Thought= Recurrent: Recurrence's Role in Language Models' Computability and a Revisit of Recurrent Transformer

URL: View paper

Brief Assessment

Autoregressive Chain Recurrent[55] focuses on the computational role of recurrence and chain-of-thought in language models, analyzing depth complexity and computability hierarchies. It does not address succinctness gaps between transformers and other formal language representations like finite automata or LTL.

4. How powerful are decoder-only transformer neural models?

URL: View paper

Brief Assessment

Decoder Only Power[50] focuses on proving Turing completeness of decoder-only transformers through RNN simulation, not on succinctness gaps between transformers and other formal language representations like finite automata or LTL.

5. Rnns are not transformers (yet): The key bottleneck on in-context retrieval

URL: View paper

Brief Assessment

RNNs Not Transformers[54] focuses on in-context retrieval capabilities and representation power for algorithmic problems, not on succinctness gaps in formal language representation. The candidate examines memory efficiency and retrieval bottlenecks rather than descriptional complexity measures.

6. Representational strengths and limitations of transformers

URL: View paper

Brief Assessment

Representational Strengths Limitations[39] focuses on approximation-theoretic separations between transformers and other architectures (RNNs, feedforward networks) for specific computational tasks, not on succinctness gaps in formal language representation or descriptional complexity measures.

7. Transformers learn shortcuts to automata

URL: View paper

Brief Assessment

Transformers Learn Shortcuts[49] focuses on how transformers simulate finite-state automata through hierarchical reparameterization of recurrent dynamics, not on proving succinctness gaps between transformers and other models like RNNs or finite automata in terms of descriptional complexity.

8. The expressive capacity of state space models: A formal language perspective

URL: View paper

Brief Assessment

State Space Formal Language[53] focuses on the expressive capacity of state space models (SSMs) in modeling formal languages, not on succinctness gaps between transformers and other models like RNNs or finite automata. The paper does not address descriptional complexity comparisons that would refute the original paper's novelty claims about transformers' exponential/doubly exponential succinctness advantages.

9. Back to recurrent processing at the crossroad of transformers and state-space models

URL: View paper

Brief Assessment

Recurrent Processing Crossroad[48] focuses on comparing linear transformers with RNNs in terms of expressivity and performance, not on succinctness gaps or descriptional complexity between transformers and formal language representations.

---

10. Separations in the Representational Capabilities of Transformers and Recurrent Architectures

URL: View paper

Brief Assessment

Cannot assess refutation as the candidate paper's full text context is marked 'n/a', preventing comparison with the original paper's claims about succinctness gaps.

Contribution 3: EXPSPACE-completeness of transformer verification

Description: The authors establish that verifying simple properties about transformers, such as checking whether they recognize a trivial language, is EXPSPACE-complete. This result shows that transformer verification is computationally intractable under standard complexity-theoretic assumptions.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Transformer encoder satisfiability: Complexity and impact on formal reasoning

URL: View paper

Prior Art Analysis

Transformer Encoder Satisfiability[9] demonstrates that verifying properties of transformer encoders is EXPSPACE-complete, establishing this result prior to the original paper. Both papers prove that checking whether a transformer recognizes a trivial language (non-emptiness/satisfiability) is EXPSPACE-complete. The candidate paper provides a comprehensive framework showing that the satisfiability problem for transformer encoders is EXPSPACE-complete, with matching upper and lower bounds. This directly refutes the novelty claim that the original authors were first to establish EXPSPACE-completeness of transformer verification.

Evidence

Evidence 1 - **Rationale**: Both papers address the complexity of verifying transformer properties. The candidate establishes decidability and complexity bounds for the satisfiability problem of transformer encoders. - **Original**: as a by-product of this expressivity, we show that verifying simple properties about transformers (e.g. whether it recognizes a trivial language) is computationally difficult: expspace-complete. - **Candidate**: w e analyse the complexity of the satisfiability problem, or s imilarly feasibility problem, (trsa t) for transformer encoders (te), which natu rally occurs in formal verification or interpretation, collectively referred to a s formal reasoning. w e find that trsa t is undecidable when considering te a...

Evidence 2 - **Rationale**: The candidate provides complexity bounds for bounded satisfiability problems, establishing a framework for understanding transformer verification complexity that predates the original paper's claims. - **Original**: we show that verifying simple properties about transformers (e.g. whether it recognizes a trivial language) is computationally difficult: expspace-complete. that is, with standard complexity-theoretic assumptions, this cannot be done in better than double exponential time. - **Candidate**: theorem 3. let t be a class of te. then 1. bt r satun[t] is in np; if tudec ⊆ t then bt r satun[t] is np-complete, 2. bt r satbin[t] is in nexptime; if tudec ⊆ t then bt r satbin[t] is nexptime-complete.

Evidence 3 - **Rationale**: Both papers establish hardness results for transformer verification. The candidate shows NEXPTIME-hardness for fixed-width arithmetic transformers, contributing to the complexity landscape. - **Original**: proposition 10. the non-emptiness problem for uhat is expspace-hard. - **Candidate**: theorem 5. t rsat[t fix] for te over fixed-width arithmetic is decidable and nexptime -hard.

---

2. Identifying the limits of transformers when performing model-checking with natural language

URL: View paper

Brief Assessment

Transformers Model Checking Limits[47] focuses on model-checking problems where transformers verify whether formulas are true in given structures (natural language inference), not on verifying properties about the transformers themselves. The computational complexity studied concerns the reasoning task, not transformer verification as in the original paper.

---

3. Error correction code transformer

URL: View paper

Brief Assessment

Error Correction Transformer[41] focuses on applying transformers to error correction code decoding, not on computational complexity or verification of transformer properties. The paper does not address verification problems or complexity-theoretic results about transformers.

---

4. Complexity control facilitates reasoning-based compositional generalization in transformers

URL: View paper

Brief Assessment

Complexity Control Compositional[43] focuses on compositional generalization in transformers through complexity control (initialization scales and weight decay), not on computational complexity of verifying transformer properties. The paper does not address verification problems or complexity-theoretic hardness results.

---

5. An efficient and extensible zero-knowledge proof framework for neural networks

URL: View paper

Brief Assessment

Zero Knowledge Neural Networks[40] focuses on zero-knowledge proof frameworks for neural network inference verification, not on the computational complexity of verifying properties of transformer architectures. The paper addresses cryptographic verification of neural network computations rather than formal verification complexity.

---

6. On the turing completeness of modern neural network architectures

URL: View paper

Brief Assessment

Turing Completeness Neural[45] focuses on Turing completeness of transformers based on their capacity to compute dense representations, not on the computational complexity of verifying properties about transformers. The candidate does not address verification complexity or EXPSPACE-completeness.

---

7. Circuit Complexity Bounds for RoPE-based Transformer Architecture

URL: View paper

Brief Assessment

RoPE Circuit Complexity[46] focuses on circuit complexity bounds (TC0) for RoPE-based transformers and their inability to solve specific problems (arithmetic/boolean formula evaluation), not on the verification problem complexity that the original paper addresses.

8. Llm-based processor verification: A case study for neuromorphic processor

URL: View paper

Brief Assessment

LLM Processor Verification[42] focuses on using LLMs for functional verification of processors (RISC-V and neuromorphic), not on the computational complexity of verifying transformer neural networks. The candidate addresses processor verification workflows, not formal verification complexity theory for transformers.

9. A unified framework for establishing the universal approximation of transformer-type architectures

URL: View paper

Brief Assessment

Universal Approximation Transformers[44] focuses on universal approximation properties and expressiveness of transformer architectures from a function approximation perspective, not on computational complexity of verification problems or formal language recognition.

10. Representational strengths and limitations of transformers

URL: View paper

Brief Assessment

Representational Strengths Limitations[39] does not address verification problems or computational complexity of checking properties about transformers. The paper focuses on representational capacity and approximation bounds for specific tasks.

Appendix: Text Similarity Detection

Textual similarity detection checked 34 papers and found 1 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Practical Computational Power of Linear Transformers and Their Recurrent and Self-Referential Extensions

Detected in: Contribution: contribution_2

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] Transformers are Inherently Succinct View paper
- [1] Transformers Can Represent -gram Language Models View paper
- [2] Knee-Deep in C-RASP: A Transformer Depth Hierarchy View paper
- [3] TAYSIR Competition: Transformer+\textscrnn: Algorithms to Yield Simple and Interpretable Representations View paper
- [4] Automata extraction from transformers View paper
- [5] Exploring Attention Patterns and Neural Activations in Transformer Architectures for Sequence Classification in Context Free Grammars View paper
- [6] Formal language recognition by hard attention transformers: Perspectives from circuit complexity View paper
- [7] Lightweight transformers for conversational ai View paper
- [8] Characterizing the Expressivity of Fixed-Precision Transformer Language Models View paper
- [9] Transformer encoder satisfiability: Complexity and impact on formal reasoning View paper
- [10] Explainability and Semantics-Bridging Natural Language Flexibility and Formal Precision: Toward a Semantic Framework for Large Language Models View paper
- [11] Transformers are Efficient Compilers, Provably View paper
- [12] Computational Expressivity of Neural Language Models View paper
- [13] How Well Can BERT Learn the Grammar of an Agglutinative and Flexible-Order Language? The Case of Basque. View paper
- [14] Transformers as recognizers of formal languages: A survey on expressivity View paper
- [15] Transformers as recognizers and transducers View paper
- [16] Between Circuits and Chomsky: Pre-pretraining on Formal Languages Imparts Linguistic Biases View paper
- [17] Simulating Weighted Automata over Sequences and Trees with Transformers View paper
- [18] Understanding models understanding language View paper
- [19] The Power of Hard Attention Transformers on Data Sequences: A Formal Language Theoretic Perspective View paper
- [20] Masked Hard-Attention Transformers Recognize Exactly the Star-Free Languages View paper
- [21] Characterizing the Expressivity of Transformer Language Models View paper
- [22] Inferring Group Intent as a Cooperative Game. An NLP-based Framework for Trajectory Analysis using Graph Transformer Neural Network View paper
- [23] GPT-Based Wasm Instruction Analysis for Program Language Processing View paper
- [24] Expression transformers in B-GSL View paper
- [25] Can You Learn to See Without Images? Procedural Warm-Up for Vision Transformers View paper
- [26] Compact Transformer-based Language Models for the Moroccan Darija View paper
- [27] Practical Computational Power of Linear Transformers and Their Recurrent and Self-Referential Extensions View paper
- [28] What Formal Languages Can Transformers Express? A Survey View paper
- [29] Transformers as CFG Learners: A Formal Framework for Structured Data View paper
- [30] Why are Sensitive Functions Hard for Transformers? View paper
- [31] On the Ability and Limitations of Transformers to Recognize Formal Languages View paper
- [32] The Grammar Correction: A Comparison of T5, LLAMA and ChatGPT View paper
- [33] Enforcing Grammar in Code Synthesis with Transformers View paper

- [34] ESSLLI 2024 Course Notes Expressivity of Transformers: Logic, Circuits, and Formal Languages View paper
- [35] Improving Formal Reasoning of Transformer with State Stack View paper
- [36] Grammar-enforced Chain of Thought Reasoning for small LLMs View paper
- [37] A transformer-based parser for Grammatical Framework View paper
- [38] Depth-Width Tradeoffs for Transformers on Graph Tasks View paper
- [39] Representational strengths and limitations of transformers View paper
- [40] An efficient and extensible zero-knowledge proof framework for neural networks View paper
- [41] Error correction code transformer View paper
- [42] Llm-based processor verification: A case study for neuromorphic processor View paper
- [43] Complexity control facilitates reasoning-based compositional generalization in transformers View paper
- [44] A unified framework for establishing the universal approximation of transformer-type architectures View paper
- [45] On the turing completeness of modern neural network architectures View paper
- [46] Circuit Complexity Bounds for RoPE-based Transformer Architecture View paper
- [47] Identifying the limits of transformers when performing model-checking with natural language View paper
- [48] Back to recurrent processing at the crossroad of transformers and state-space models View paper
- [49] Transformers learn shortcuts to automata View paper
- [50] How powerful are decoder-only transformer neural models? View paper
- [51] Rethinking transformers for efficiency and scalability View paper
- [52] Separations in the Representational Capabilities of Transformers and Recurrent Architectures View paper
- [53] The expressive capacity of state space models: A formal language perspective View paper
- [54] Rnns are not transformers (yet): The key bottleneck on in-context retrieval View paper
- [55] Autoregressive+ Chain of Thought= Recurrent: Recurrence's Role in Language Models' Computability and a Revisit of Recurrent Transformer View paper
- [56] Quantum long short-term memory View paper
- [57] Compacting, picking and growing for unforgetting continual learning View paper
- [58] Saturation in Recurrent Neural Networks: Expressivity, Learnability, and Generalization. View paper
- [59] DNS-Rec: Data-aware Neural Architecture Search for Recommender Systems View paper
- [60] A Lightweight Sequential Convolutional Neural Network for Smart Grid Stability Analysis View paper
- [61] Towards Structured Intelligence for Sequence Modeling View paper
- [62] LooperGP: A loopable sequence model for live coding performance using GuitarPro tablature View paper
- [63] Consistent Bidirectional Language Modelling: Expressive Power and Representational Conciseness View paper
- [64] Explainable and Efficient Knowledge Acquisition from Text View paper
- [65] Low-rank passthrough neural networks View paper