

Optimizing Decomposition for Optimal Claim Verification

Yining Lu Noah Ziems Hy Dang Meng Jiang

University of Notre Dame
South Bend, IN
y1u33@nd.edu

Abstract

Current research on the *Decompose-Then-Verify* paradigm for evaluating the factuality of long-form text typically treats decomposition and verification in isolation, overlooking their interactions and potential misalignment. We find that existing decomposition policies, typically hand-crafted demonstrations, do not align well with downstream verifiers in terms of atomicity—a novel metric quantifying information density—leading to suboptimal verification results. We formulate finding the optimal decomposition policy for optimal verification as a bilevel optimization problem. To approximate a solution for this strongly NP-hard problem, we propose dynamic decomposition, a reinforcement learning framework that leverages verifier feedback to learn a policy for dynamically decomposing claims to verifier-preferred atomicity. Experimental results show that dynamic decomposition outperforms existing decomposition policies, improving verification confidence by 0.07 and accuracy by 0.12 (on a 0-1 scale) on average across varying verifiers, datasets, and atomicities of input claims.¹

1 Introduction

The *Decompose-Then-Verify* paradigm has been widely used in fact-checking systems, as it reduces the claim complexity and makes the factuality evaluation fine-grained and easy (Min et al., 2023; Chern et al., 2023; Chen et al., 2023; Kamoi et al., 2023; Wei et al., 2024; Song et al., 2024). The paradigm comprises two components: (1) a decomposer, which leverages a large language model (LLM) guided by a decomposition policy—typically hand-crafted prompts—to break claims into subclaims,² and (2) a verifier, which

similarly utilizes LLM paired with a verification policy (e.g., retrieving evidence to assist verification). Figure 1 provides an illustration.

In this work, we first systematically investigate how decomposition policies could affect verification through subclaim *atomicity*—a metric we introduced for quantifying information density. We find that *different verifiers achieve optimal verification confidence at distinct input atomicity*. We define $\text{atomicity} = \log_2(\# \text{ atomic information})$, where one piece of atomic information is an utterance conveying a single nontrivial fact (e.g., “Owen made a comeback in 2017” in Figure 1). Higher atomicity means a subclaim is more coarse-grained and information-rich, which ostensibly implies lower verifier confidence, yet our finding indicates this is not always the case.

The above finding reveals that existing *prompt-based decomposition policies do not always generate subclaims with optimal atomicity, resulting in suboptimal verification results*. For instance, FActScore (Min et al., 2023) formulates its decomposition policy using eight annotated demonstrations to generate intended atomic subclaims. Our experiments find that verifiers, such as Inst-Llama-7B with the evidence retrieval verification policy shown in Figure 1, do not exhibit optimal performance at the atomicity level (i.e., $\text{atomicity} = 0$) featured by the given decomposition policy. Thus, there is a performance gap between the decomposer and verifier in terms of atomicity, which remains unaddressed. We further discuss the limitations of prior studies in the related work §5.1.

To solve the above problem, we propose *dynamic decomposition*—a novel framework to learn a decomposition policy tailored to the downstream verifier. Our approach is compatible with any existing fact-checking systems where both the decomposition and verification LLMs are given and frozen. Unlike existing decomposition policies that use a single prompt call to generate subclaims (Min et al.,

¹Our code: github.com/yining610/dynamic-decomposition

²Following Jiang et al. (2024), we use *claim* to denote original sentence to be verified and *subclaims* the result of decomposition.

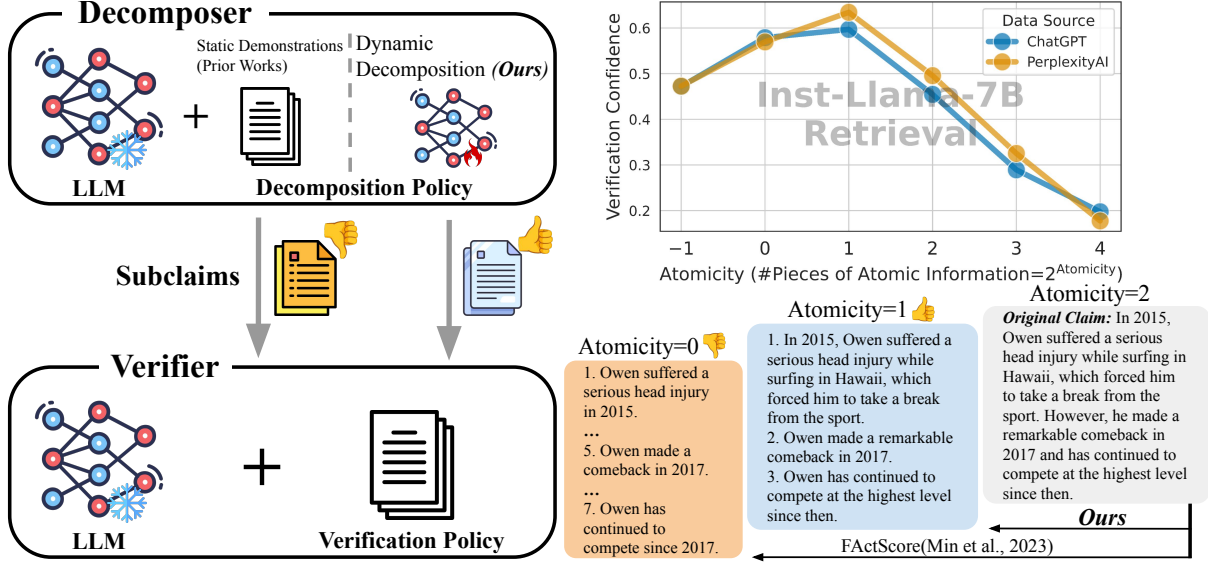


Figure 1: **Left:** overall framework of *Decompose-Then-Verify* paradigm. We define each decomposer and verifier as a LLM paired with a corresponding policy. Our dynamic decomposition is compatible with existing fact-checking systems and requires training only a decomposition policy with 4.73M parameters. **Right:** the figure (*upper right*) shows that the verification confidence of the verifier (i.e., Inst-Llama-7B with a retrieval verification policy) peaks at atomicity 1. An atomicity of -1 denotes the claim is partially trivial and tautological. The example (*lower right*) shows that **the decomposition policy from FActScore (Min et al., 2023) fails to generate subclaims that best evoke the verifier’s performance, leading to suboptimal results.**

2023; Wei et al., 2024; Hu et al., 2024; Wanner et al., 2024), we formulate dynamic decomposition as a *Markov Decision Process* (Puterman, 2014) that involves a sequence of decomposition calls. At each step, the policy determines whether a subclaim should be decomposed, with the newly generated subclaims passed to the verifier that returns verification confidence change as a reward.³ Thus, this formulation enables the problem to be tackled using on-policy reinforcement learning (RL).

The learned policy, requiring only 4.73M parameters, significantly improves steerability over decomposition processes by dynamically decomposing claims to the verifier-preferred atomicity level. Extensive experiments show that it outperforms baseline decomposition policies (Min et al., 2023; Kamoi et al., 2023; Wanner et al., 2024), improving both verification confidence (by 0.07) and accuracy (by 0.12) across varying verifiers, datasets, and atomicity levels. In summary, our contributions are twofold:

- **Our study exposes the impact of subclaim atomicities on verifiers.** We find that each verifier prefers a distinct optimal input atomicity, yet existing decomposition policies hardly achieve

the optimum.

- **We introduce a RL framework designed to bridge the performance gap between decomposers and verifiers.** It learns a decomposition policy to dynamically adjust the atomicity of subclaims tailored to downstream verifiers, thereby optimizing verification results.

2 Methodology

We formulate finding the optimal decomposition policy for optimal verification as a bilevel optimization problem. Specifically, given a decomposition LLM \mathcal{D} , a verification LLM \mathcal{V} , verification policy π_v , and a claim dataset $\{(C_i, Y_i)\}$ with binary factuality labels Y_i , we aim to find an optimal decomposition policy π_d such that decomposed subclaims $\{c_j\}$ maximize the verification accuracy:

$$\max_{c \in \{c_j\}, \pi_d} \mathbb{E}_i \left[\mathbb{1}(Y_i = \bigwedge_c \mathcal{V}(c | \pi_v)) \right], \quad (1)$$

subject to

$$\pi_d \in \arg \max_{\pi_d} f(\{c_j\}, \pi_d), \{c_j\} \sim \mathcal{D}(C_i | \pi_d),$$

where verifier $\mathcal{V}(\cdot | \pi_v)$ returns the prediction label conditioned on its verification policy. We determine the claim is true if and only if all its subclaims are true, implemented via a logical AND

³We provide a detailed justification of the reward design in §2.2.

operator \bigwedge . $f(\cdot)$ represents the lower-level constraint, which observes the decisions $\{c_j\}$ made at the upper level and optimizes its own policy π_d accordingly (Sinha et al., 2018). Decomposer \mathcal{D} is the upper-level constraint, ensuring that all subclaims are generated by it under a feasible decomposition policy that is lower-level optimal.

By formulating the problem as a bilevel optimization, we can simultaneously refine the upstream decomposition policy and optimize the downstream verification task, ensuring that the decomposition policy is aligned with the verifier to achieve optimal overall performance. Bilevel optimization is known to be strongly NP-hard (Hansen et al., 1992), and research has shown that it can be alternatively approximated using online stochastic approximation (Qiu et al., 2021). Therefore, we propose our dynamic decomposition as an advantage actor-critic (A2C) style (Mnih et al., 2016) RL solution to approximate Eq.1. It enables the decomposition policy to learn directly from the verifier, dynamically converging toward an optimistic bilevel optimum.

2.1 Overview of Dynamic Decomposition

Unlike most of the prior work that applies its decomposition policy only once, we iteratively generate decomposition calls from the learned decomposition policy. The call is to either request the decomposition LLM to decompose the current subclaim or not. Specifically, given a temporary subclaim list $\{c_j\}$ and a target subclaim to be decomposed $c^* \in \{c_j\}$ sampled from it, we perform decomposition $\mathcal{D}(\cdot \mid \pi_d)$ in Eq.1 as:

$$\{c_j\}^* = \begin{cases} \mathcal{D}(c^*), & \text{if } \pi_d(\{c_j\}) = \text{decompose} \\ \emptyset, & \text{if } \pi_d(\{c_j\}) = \text{not decompose} \end{cases}, \quad (2)$$

$$\{c_j\} \leftarrow \kappa(\{c_j\}, \{c_j\}^*). \quad (3)$$

We repeat the above process until all subclaims in $\{c_j\}$ have been decided not to decompose further.

Therefore, this can be formulated as a finite MDP defined as $M = (\mathcal{S}, \mathcal{A}, \kappa, r)$. \mathcal{S} represents the state space, and \mathcal{A} is the action space which includes two actions in our case: decompose or not to decompose. $\kappa : \mathcal{S} \times \mathcal{S}^* \rightarrow \mathcal{S}$ is the state transition function that replaces the target subclaim c^* in the subclaim list $\{c_j\}$ with its decomposition results $\{c_j\}^*$. $r : \mathcal{S} \times \mathcal{S}^* \rightarrow \mathcal{A}$ is the immediate reward received from the verifier after state transition.

2.2 Implement Dynamic Decomposition Policy

In this section, we elaborate on how we implement our dynamic decomposition based on the MDP formulation proposed above.

Atomicity state: To find the optimal atomicity that is favored by the verifier, we create an *atomicity state* reflecting the overall atomicity of current subclaims $\{c_j\}$ at step t . Each atomicity state is a d dimension vector $s_t \in \mathbb{R}^d$.

State transition: Similar to the work of Chen et al. (2024b), we use a trainable Gated Recurrent Unit (GRU) (Cho et al., 2014) to model state transition function κ in Eq.3:

$$s_{t+1} = \text{GRU}[s_t, (1 + \sigma(\Delta\text{Info}))\text{Enc}(\{c_j\})], \quad (4)$$

where $\text{Enc}(\cdot) : T \rightarrow \mathbb{R}^d$ is a textual encoder that maps the text sequence to a d -dimensional embedding and $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ is the sigmoid function. We compute ΔInfo as the average Conditional Pairwise Mutual Information (CPMI) (Jiang et al., 2024) difference between the target subclaim c^* and its decomposed results $\{c_j\}^*$, which basically quantifies how much information is lost from each subclaim after decomposition:

$$\Delta\text{Info} = \mathbb{E}_{c \in \{c_j\}^*} \left[\log \frac{P(c \mid \mathcal{H})}{P(c^* \mid \mathcal{H})} \right], \quad (5)$$

where $P(\cdot \mid \mathcal{H})$ measures the entailment probability between a claim and a pre-defined set of tautological and bleached claims \mathcal{H} (e.g., “{topic} is a person”, “{topic} exists”). In practice, this conditional probability is estimated as $\max_{h \in \mathcal{H}} P(\cdot \mid h)$ for better stability (Jiang et al., 2024). As both c^* and its subclaims $c \in \{c_j\}^*$ getting syntactically closer to bleached claims through iterative decomposition, ΔInfo tend to decrease because of the diminishing marginal information loss (e.g., $\frac{P(\text{“Kruger was a religious man”} \mid \mathcal{H})}{P(\text{“Kruger was a deeply ... faith”} \mid \mathcal{H})} \gg \frac{P(\text{“Kruger was a man”} \mid \mathcal{H})}{P(\text{“Kruger was a religious”} \mid \mathcal{H})}$; see Figure 6 in Appendix C for experimental justification.).

Therefore, we design ΔInfo as a metric to quantify localized atomicity change caused by the current decomposition call, and $\text{Enc}(\{c_j\})$ reflects the overall atomicity of all subclaims in terms of semantics. Multiplying global decomposition embeddings by the local atomicity change in Eq.4 preserves original semantics (Vaswani et al., 2017)

while revealing hidden states regarding atomicity of current subclaims.

Action: We define \mathcal{A} as having only two actions: 1 (decompose) or 0 (not decompose). For each atomicity state, the action is sampled from a policy distribution, $a_t \sim \pi_d(a_t | s_t)$, namely our decomposition policy. π_d is trained to determine when to decompose a claim until achieving the desired atomicity state that maximizes the reward.

Reward: Due to the lack of ground-truth labels for newly generated subclaims, accuracy-based evaluation is not feasible for evaluating verification improvement after decomposition. To address this limitation, we introduce verification confidence, a label-free proxy for accuracy that is computable for all subclaims. It is defined as the absolute probability difference between positive and negative verification labels:

$$\text{Conf}(c, \mathcal{V}, \pi_v) = |P_{\mathcal{V}}(\text{True} | c, \pi_v) - P_{\mathcal{V}}(\text{False} | c, \pi_v)|. \quad (6)$$

Verification confidence measures *how much certainty a verifier has in making a verification*, which we find to strongly correlate with verification accuracy in Figure 2.

Our reward is the verification confidence change before and after decomposition. This design encourages the policy to perform decomposition when the verifier is more confident in evaluating the factuality of generated subclaims:

$$r_t = \underbrace{\mathbb{E}_{c \in \{c_j\}^*} [\text{Conf}(c, \mathcal{V}, \pi_v)]}_{\text{After Decomposition}} - \underbrace{\text{Conf}(c^*, \mathcal{V}, \pi_v)}_{\text{Before Decomposition}}. \quad (7)$$

Breadth-First Order Decomposition: In dynamic decomposition, each decomposition call could generate a new list of subclaims at a lower level of atomicity, which are then queued for further decomposition. Therefore, the order of decomposition becomes matter. A depth-first approach, where a single subclaim is continually decomposed until reaching the lowest possible atomicity level, can result in significant variance in atomicity among all subclaims, which in turn leads to high variance in modeling the state (Eq.4). Hence, we employ a breadth-first strategy to prioritize the decomposition of subclaims at higher atomicity. We provide an illustration in Figure 3.

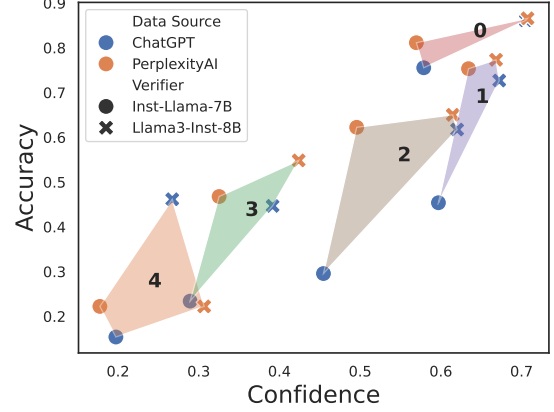


Figure 2: Verification confidence versus accuracy. The number in each convex hull denotes the claim atomicity. **Irrespective of data sources, atomicities, and verifiers, verification confidence exhibits a strong positive correlation with accuracy (0.88 Pearson’s r).**

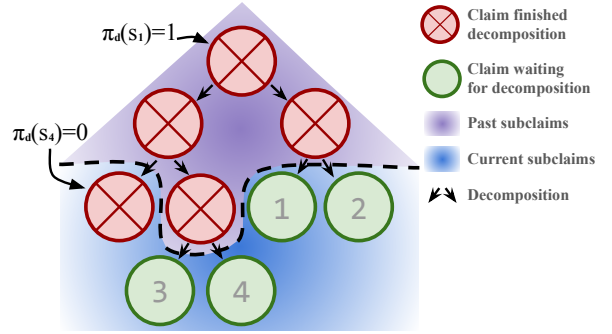


Figure 3: Breadth-first order sampling for dynamic decomposition. We perform binary decomposition for each claim. The number in the node represents its sampling priority in the decomposition process. We first sample out subclaims at the same atomicity level, with newly generated subclaims queued in a FIFO (first-in-first-out) order.

2.3 Train Dynamic Decomposition Policy

We employ PPO (Schulman et al., 2017) in A2C style to train our dynamic decomposition policy given its effectiveness and stability (Engstrom et al., 2019). We model policy function $\pi_d : \mathbb{R}^d \rightarrow \mathbb{R}^2$ as an MLP (Multi-Layer Perceptron) that outputs a two-dimensional normalized vector. A core component of the PPO objective function is the clipped surrogate term, which constrains the policy change during the optimization process. Given a finite decomposition trajectory $\{(a_1, s_1), (a_2, s_2), \dots, (a_T, s_T)\}$, we have

$$L^{\text{clip}} = \mathbb{E}_t \left[\min(\rho_t, \text{clip}(\rho_t, 1 - \epsilon, 1 + \epsilon)) \hat{A}_t \right], \quad (8)$$

where $\rho_t = \frac{\pi_d(a_t|s_t)}{\pi_d^{\text{old}}(a_t|s_t)}$ is a probability ratio to estimate the divergence between old and current policy. The hyperparameter ϵ sets the clipping boundary for ρ to fall between $[1 - \epsilon, 1 + \epsilon]$. \hat{A}_t is the advantage at step t which measures how better taking the action a_t at state s_t is compared to the average value of the state.

To calculate the average value of a state, we create another trainable MLP as a value function, $v : \mathbb{R}^d \rightarrow \mathbb{R}$, to map the state to its corresponding value. Then, we estimate the advantage \hat{A}_t using GAE (Generalized Advantage Estimator; Schulman et al. (2016)):

$$\hat{A}_t = \delta_t + (\gamma\lambda)\delta_{t+1} + \dots + (\gamma\lambda)^{T-t+1}\delta_{T-1},$$

where $\delta_t = r_t + \gamma v(s_{t+1}) - v(s_t)$. (9)

r_t is the reward defined in Eq.7. δ_t is the TD residual of value function with discount factor γ (Sutton and Barto, 1998). Another hyperparameter λ controls the trade-off between bias and variance in advantage estimation. We use the squared-error loss to train our value function: $(v(s_t) - v_t^{\text{target}})^2 \approx (v(s_t) - (\hat{A}_t + v(s_t)))^2 = \hat{A}_t^2$. Therefore, our final PPO objective function with entropy bonus term $S[\pi_d](s_t)$ becomes:

$$L^{\text{PPO}} = \mathbb{E}_t \left[L^{\text{clip}} - c_1 \hat{A}_t^2 + c_2 S[\pi_d](s_t) \right], \quad (10)$$

where c_1 and c_2 are coefficients. We perform gradient descent on $-L^{\text{PPO}}$ to maximize the above objective function. Algorithm 1 outlines the procedure.

3 Experiment Setup

3.1 Dataset Construction

To evaluate the effect of decomposition policies on verification, we construct two claim datasets from FActScore (Min et al., 2023), whose original claims are sourced from ChatGPT (OpenAI, 2022) and PerplexityAI (AI, 2023).⁴ Each dataset contains claims with 6 different atomicities ranging from -1 to 4. Specifically, we consider the given human-annotated atomic subclaims as the base with atomicity 0, because they typically are self-contained sentences with single factual information (e.g., “The String is a collection of poetry”). We construct higher-level subclaims by recursively merging pairs at the same atomicity level from the base up until atomicity 4, which

⁴For simplicity, we refer to each dataset by the name of its source LLM.

Algorithm 1 Train Dynamic Decomposition Policy

Input: decomposition LLM \mathcal{D} , verification LLM \mathcal{V} , decomposition policy π_d , verification policy π_v , value function v , state transition model GRU, initial atomicity state s_0 , claims $\{C_i\}$ to be verified.

```

1: while not done do
  # update replay buffer
2:   for step = 1, ..., m do
3:      $C \sim \{C_i\}, \{c_j\} = C, s_t = s_0$   $\triangleright$  initialization
4:     while not finish decomposition do
5:        $s_t \leftarrow$  start atomicity state
6:        $c^* \xleftarrow{\text{BF Sampling}} \{c_j\}$   $\triangleright$  get target subclaim
7:        $a_t \leftarrow$  sample action from  $\pi_d(a_t | s_t)$ 
8:       LLM  $\mathcal{D}$  decompose  $c^*$  following Eq.2
9:        $\{c_j\} \leftarrow$  update subclaim list following Eq.3
10:       $r_t \leftarrow$  reward from Eq.7
11:       $s_{t+1} \leftarrow$  end atomicity state updated by Eq.4
12:      Record  $(a_t, r_t, s_t, s_{t+1})$  into replay buffer  $R$ 
13:    end while
14:  end for
  # train dynamic decomposition policy
15:   $\pi_d^{\text{old}} \leftarrow \pi_d$ 
16:  for each update step do
17:    Sample minibatch  $M$  from replay buffer  $R$ 
18:    for each  $(a_t, r_t, s_t, s_{t+1})$  in  $M$  do
19:       $\hat{A}_t \leftarrow$  advantage from Eq.9
20:       $\rho_t \leftarrow \frac{\pi_d(a_t|s_t)}{\pi_d^{\text{old}}(a_t|s_t)}$ 
21:    end for
22:     $L^{\text{clip}} \leftarrow$  clipped surrogate term from Eq.8
23:     $L^{\text{PPO}} \leftarrow$  objective function from Eq.10
24:    Update  $\pi_d, v, \text{GRU}$  using  $-L^{\text{PPO}}$  through GD
25:  end for
26: end while

```

is the highest level that cannot be merged further. We also decompose the base subclaims to a lower level at atomicity -1, where subclaims are partially tautological (e.g., “String exists”, “String is a collection”, “The collection is composed of poetry”).⁵ Thus, each subclaim is purposefully built to contain $2^{\text{atomicity}}$ pieces of information. We provide data statistics in Appendix B.1.

3.2 Models

Decomposers Because decomposition requires a deep understanding of both the semantic and syntactic aspects of the given claims, we use the following two open-source and widely-recognized models as decomposition LLM: Llama3-Inst-70B (Meta, 2024) and DeepSeek-V3 (DeepSeek-AI, 2024).

Verifiers Following Min et al. (2023), we determine factual labels by comparing the conditional probability of True and False from the verifier. We experiment the following three verification LLMs: a T5-3B (Raffel et al., 2023)

⁵We use gpt-3.5-turbo-0125 to perform the decomposition and name the decomposition policy as FActScore-Atom whose prompt can be found in Appendix D.

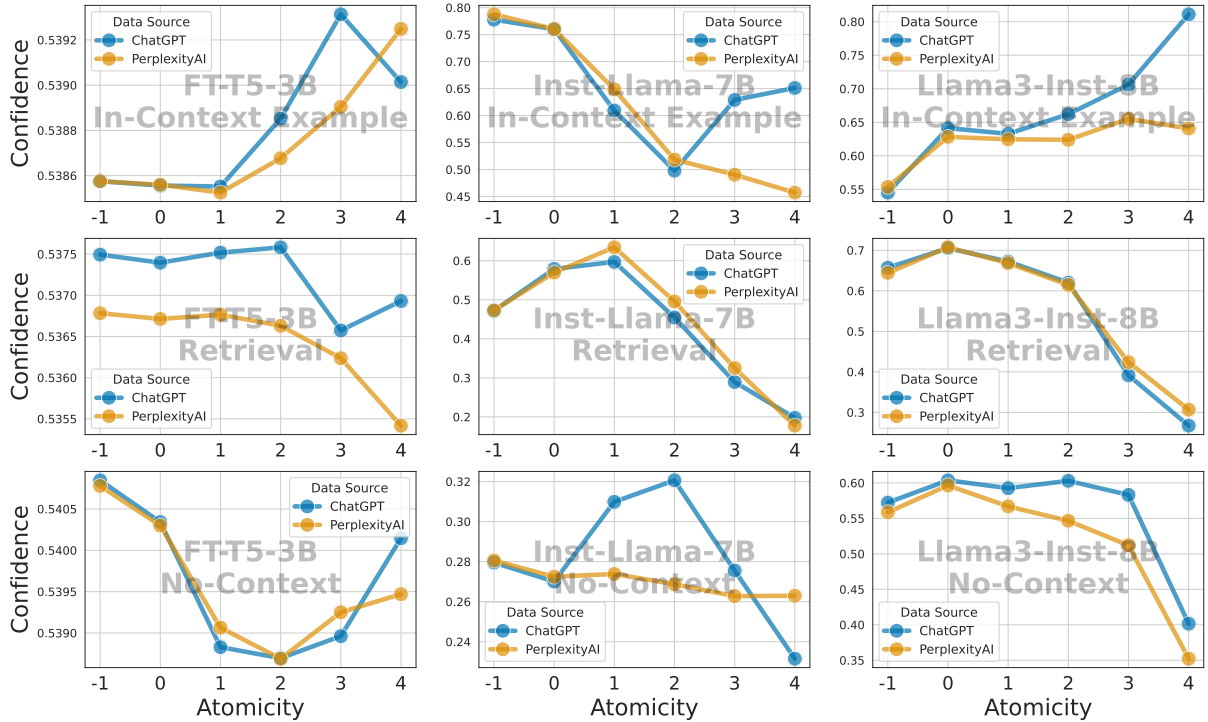


Figure 4: Verification confidence across atomicities. **Evidently, each verifier has its own preferred input atomicity at which the verification confidence peaks.** Even when utilizing the same verification policy, such as retrieval, different verifiers exhibit distinct preferences, and vice versa.

fine-tuned on FActScore for the factuality verification task (FT-T5-3B), a Llama-7B (Touvron et al., 2023) trained on Super Natural Instructions (Wang et al., 2022) (Inst-Llama-7B), and a pretrained and instruction tuned Llama3-8B model (Meta, 2024) (Llama3-Int-8B). We employ three verification policies, each utilizing differently constructed prompts to assist verification: (1) **Retrieval** retrieves relevant passages from a database as evidence; (2) **In-Context Example** provides verification demonstrations to instruct verification,⁶ and (3) **No-Context** directly asks the verifier for predictions, as inspired by Kadavath et al. (2022).

3.3 Training

Initialization We model each policy and value function as a two-layer fully connected perceptron with a ReLU activation function (Agarap, 2019). The total number of trainable parameters is 4.73M. We perform binary decomposition on each subclaim during our dynamic decomposition, aligning with the definition of atomicity (using a logarithm with base 2), to ensure maximal exploration of the subclaim space. Please refer to Appendix B.2 for more details.

⁶In-context examples can be found in Appendix D.

Data We train our policy on two constructed claim datasets across atomicity [1, 4].

Hyperparameters Please see Appendix B.3.

3.4 Baselines and Metrics

We compare our trained decomposition policy (hereafter denoted as **DYDECOMP**) to existing decomposition policies, including **FActScore** (Min et al., 2023), **WICE** (Kamoi et al., 2023) and **R-ND** (Wanner et al., 2024), on both **verification confidence** (Eq.6) and **accuracy** (Eq.1). These works typically apply heuristic splitting prior to neural decomposition, thus their decomposition policies are primarily designed for subclaims under atomicity 2. To ensure a fair comparison, our evaluation is conducted on claims on atomicity within [0, 2]. We also evaluate DYDECOMP against a modified version of the FActScore policy, **FActScore-Atom**, which is designed to decompose human-annotated atomic subclaims (atomicity 0) to partially trivial subclaims (atomicity -1) in §3.1.

4 Results and Analysis

We first study the effect of subclaim atomicity on verification (§4.1), followed by evaluating dynamic decomposition against existing baselines (§4.2) and ablation study (§4.3).

Atomicity	Decompose Policy → Verify Policy↓	FactScore	FactScore-Atom	WICE	R-ND	DyDECOMP
		Verification Confidence [0-1] ↑ / Verification Accuracy [0-1] ↑				
0	Data Source: ChatGPT					
	Retrieval	0.627 / 0.666	0.618 / 0.796	0.431 / 0.782	0.556 / 0.449	0.600 / 0.789
	In-Context Example	0.677 / 0.454	0.677 / 0.388	0.724 / 0.451	0.714 / 0.401	0.715 / 0.428
	No-Context	0.557 / 0.457	0.526 / 0.566	0.374 / 0.525	0.527 / 0.378	0.547 / 0.551
	Data Source: PerplexityAI					
	Retrieval	0.629 / 0.559	0.611 / 0.755	0.435 / 0.762	0.541 / 0.241	0.612 / 0.799
1	In-Context Example	0.681 / 0.266	0.670 / 0.197	0.726 / 0.308	0.711 / 0.172	0.733 / 0.301
	No-Context	0.555 / 0.352	0.515 / 0.471	0.380 / 0.475	0.519 / 0.166	0.542 / 0.554
	Data Source: ChatGPT					
	Retrieval	0.609 / 0.739	0.611 / 0.815	0.527 / 0.755	0.541 / 0.635	0.654 / 0.758
	In-Context Example	0.714 / 0.635	0.705 / 0.627	0.658 / 0.631	0.749 / 0.610	0.809 / 0.609
	No-Context	0.549 / 0.619	0.521 / 0.550	0.437 / 0.606	0.508 / 0.610	0.567 / 0.512
2	Data Source: PerplexityAI					
	Retrieval	0.610 / 0.493	0.615 / 0.68	0.527 / 0.597	0.515 / 0.232	0.651 / 0.753
	In-Context Example	0.721 / 0.247	0.711 / 0.253	0.684 / 0.27	0.746 / 0.22	0.791 / 0.347
	No-Context	0.535 / 0.260	0.506 / 0.350	0.423 / 0.273	0.506 / 0.223	0.559 / 0.437
	Data Source: ChatGPT					
	Retrieval	0.616 / 0.844	0.639 / 0.887	0.588 / 0.809	0.547 / 0.852	0.644 / 0.652
3	In-Context Example	0.731 / 0.835	0.725 / 0.835	0.665 / 0.835	0.744 / 0.835	0.824 / 0.750
	No-Context	0.553 / 0.809	0.545 / 0.800	0.507 / 0.835	0.515 / 0.835	0.583 / 0.509
	Data Source: PerplexityAI					
	Retrieval	0.622 / 0.483	0.639 / 0.601	0.592 / 0.546	0.529 / 0.406	0.633 / 0.664
	In-Context Example	0.755 / 0.392	0.734 / 0.378	0.675 / 0.392	0.751 / 0.392	0.823 / 0.464
	No-Context	0.543 / 0.378	0.535 / 0.378	0.492 / 0.392	0.509 / 0.392	0.544 / 0.421

Table 1: Comparison of our DYDECOMP over baselines on the test dataset. Each metric is scaled from 0 to 1. ↑ indicates higher values are preferred. We employ decomposition LLM Llama3-Inst-70B and verification LLM Llama3-Int-8B. **DYDECOMP consistently outperforms on atomicity 1 and 2, achieving an average improvement of 0.07 in verification confidence across two datasets and three verifiers, and a 0.12 average improvement in verification accuracy for claims sourced from PerplexityAI.**

4.1 Effect of Subclaim Atomicity on Verification

Prior works have found that decomposition policy influences verification results and does not guarantee consistent verification improvement across varying input length and verifier strength (Jiang et al., 2024; Wanner et al., 2024; Hu et al., 2024). In this study, we take one step further by looking into the following two questions: *How can we quantify the impact of decomposition policy on verification, and why does this influence not yield consistent improvements across different verifiers?*

Why use verification confidence? Our experiment finds a strong correlation between verification confidence and accuracy across various conditions (Figure 2), including different datasets, atomicity levels, and verifiers. This finding supports verification confidence as *a reliable metric for evaluating the impact of decomposition policies* and as an ideal signal funneling back to policy.

Each verifier has its own atomicity optimum. We investigate how atomicity, a key characteristic

of decomposed subclaims, affects verification. Figure 4 shows verification confidence change across different atomicity levels. We find that each verifier, which is a verification LLM with a specific verification policy, achieves peak verification confidence at a distinct optimal atomicity. For instance, Llama3-Inst-8B with a retrieval verification policy consistently performs best at atomicity 0 on both the ChatGPT and PerplexityAI datasets, whereas FT-T5-3B with retrieval reaches its optimum at atomicity 2. This observation answers the second question that *the different preference for atomicity makes existing static decomposition policies hard to find optimal subclaims that bring consistent verification improvement.*

4.2 Dynamic Decomposition Results

Table 1 shows the evaluation results of decomposer Llama3-Inst-70B with DYDECOMP policy on various verifiers, datasets, and claim atomicities. Evidently, DYDECOMP consistently achieves the highest verification confidence for claims with atomicity 1 and 2, outperforming the four baselines by an average margin of 0.07. This result aligns

with our findings from Figure 4, where the verification LLM for Table 1 (i.e., Llama3-Inst-8B) generally achieves the best performance at atomicity 0. In other words, an effective decomposition policy should strategically decompose claims to lower atomicities, ideally to 0, which is an ability our DYDECOMP excels. Conversely, since claims are already near optimal, DYDECOMP does not always achieve higher verification confidence than other baselines when atomicity = 0.

We also evaluate DYDECOMP on verification accuracy and observe a notable average improvement of 0.12 for claims from PerplexityAI with atomicity 1 and 2. We repeat the experiments with another decomposition LLM, DeepSeek-V3, and observe similar improvement in verification accuracy (see Table 5 in Appendix C).

4.3 Ablation Study

Table 2 display our ablation study on different components of DYDECOMP regarding *algorithm design* and *training data selection*. We use the decomposition LLM Llama3-Inst-70B, and the verifier Llama3-8B-Inst with a retrieval verification policy. The ablation experiments are conducted on claims with atomicity 4.

Variants	Verification Confidence
DYDECOMP	0.446
– one layer NN	0.398 (-0.048)
– binary; + triple decompose	0.424 (-0.022)
– entropy bonus	0.356 (-0.090)
– data on atomicity 1	0.353 (-0.093)
– data on atomicity 1, 2	0.356 (-0.090)
– data on atomicity 1, 2, 3	0.401 (-0.045)

Table 2: Ablation study results of DYDECOMP. “–” indicates the removal of a component from DYDECOMP. For instance, “– one layer NN” means modeling the policy and value functions using a single-layer perceptron instead of two layers. “– data on atomicity 1” removes claims with atomicity 1 from the training data.

Decomposition exploration is effective for long-form text verification. In experiment setup §3, we model each policy and value function as a two-layer perceptron. We investigate whether a simple, shallow model can capture the atomicity state and find that reducing the network to a single layer results in lower verification confidence for decomposed subclaims. Similarly, switching from binary to triple decomposition at each step, which reduces the exploration of subclaims on different

atomicities, also leads to declined verification confidence. Furthermore, removing the entropy bonus term, which promotes action exploration, from the objective function leads to a substantial drop in verification confidence (-0.090). These findings demonstrate that DYDECOMP benefits from diverse decomposition trajectories during training, which facilitate the search for subclaims with optimal atomicity.

Cross-atomicity training data stabilizes performance. In Table 2, we train DYDECOMP using claims with atomicity ranging from 0 to 4 and evaluate on atomicity 4. We find that gradually removing claims of lower atomicity (from 1 to 3) from the training set negatively impacts verification performance, highlighting the importance of cross-atomicity data for improving the generalizability of DYDECOMP. However, this negative effect diminishes as more data with irrelevant atomicities is removed ($-0.093 \rightarrow -0.045$). This suggests that increased exposure to claims of a specific atomicity during PPO rollouts enhances learning for that atomicity but cannot fully compensate for performance losses due to reduced atomicity coverage in training data.

5 Related Works

5.1 Decompose-Then-Verify Paradigms

Unlike traditional fact-checking systems that focus on short and simple claims (Thorne et al., 2018; Sathe et al., 2020; Schuster et al., 2021; Chen et al., 2022; Guo et al., 2022), *Decompose-Then-Verify* now becomes a typical approach in long-form text evaluation works, as it allows for more precise error identification and enhances the accuracy of verification by decomposing claims into shorter subclaims, which can then be independently validated (Min et al., 2023; Chern et al., 2023; Kamoi et al., 2023; Chen et al., 2023; Iqbal et al., 2024; Song et al., 2024; Chen et al., 2024a; Wang et al., 2024). However, how the decomposition and verification should be conducted is always underspecified.

Decomposition policies. Existing factuality evaluation works have proposed various decomposition prompts revealing different characteristics of textual decomposition regarding precision (Min et al., 2023), verifiability (Song et al., 2024), coverage (Wanner et al., 2024), and atomicity (Stacey et al., 2024). More recently, research has found that superficially fine-grained subclaims with trivial in-

formation could easily inflate verification precision (Jiang et al., 2024), and decomposition benefits weaker verifiers more than stronger verifiers by generating simpler subclaims (Hu et al., 2024). Our dynamic decomposition, therefore, serves as an ideal solution to tackle the over-decomposition problem.

Moreover, static decomposition policies may struggle to handle input claims with varying fact density and often produce atomically homogeneous subclaims. We elaborate on these limitations in Appendix A.

Verification policies. Policies for verification can be categorized according to the methods verifiers utilize to process subclaims and predict final labels. Popular processing methods include retrieving relevant evidence (Kamoi et al., 2023; Wei et al., 2024), constructing in-context exemplars (Kamoi et al., 2023; Song et al., 2024), generating claim-focused summarization (Chen et al., 2024a), and simply zero-shot prompting (Kadavath et al., 2022; Min et al., 2023). The final label can be predicted by either gradient-based approaches (e.g., comparing logits of factual labels) (Chen et al., 2024a; Tang et al., 2024; Milbauer et al., 2023; Kamoi et al., 2023; Min et al., 2023) or searching for keywords (e.g., True or False) (Min et al., 2023; Li et al., 2024; Song et al., 2024).

5.2 RL in NLP Problems

Prior works have validated the use of RL in optimizing singular-task systems, such as identifying optimal exemplars for in-context learning (Zhang et al., 2022; Lu et al., 2023; Chen et al., 2024b). However, how RL can be applied to dual-task systems where multiple LLMs are involved remains under-explored. One concurrent work uses two LLMs as process reward and policy models, employing PPO to train them jointly for challenging reasoning tasks (Cui et al., 2025). In contrast, our dynamic decomposition first explores RL to solve a bilevel optimization problem in another dual-task system, *Decompose-Then-Verify*, to reveal the nuanced characteristics of hierarchical LLM systems through their interactions.

6 Conclusion

We find that each verifier has an optimal atomicity where its verification confidence peaks. To leverage it, we introduce dynamic decomposition that optimizes claim verification by decomposing

claims into verifier-preferred atomicity learned via on-policy optimization. Our policy stands out for its adaptability to diverse verifiers and input claim atomicities, outperforming existing baselines while adding only 4.73M parameters.

Limitations

Different characteristics of decomposition. While dynamic decomposition addresses the problem of misalignment between decomposer and verifier, we focus on the aspect of information density (i.e., atomicity). Although well-structured, self-contained, and verifiable subclaims could further improve verification, these aspects are beyond the scope of this paper. Future research could investigate other key characteristics of decomposition and explore how to amplify their positive effects on verification through dynamic decomposition.

Evaluation metrics. Our reward design relies on verification confidence rather than accuracy. We leave it for future works to acquire more ground-truth labels to effectively employ verification accuracy as feedback.

Acknowledgments

This work was supported by NSF IIS-2119531, IIS-2137396, IIS-2142827, IIS-2234058, CCF-1901059, and ONR N00014-22-1-2507. The authors would like to thank Chihiro Taguchi, Katsumi Ibaraki and Demetrius Hernandez for their helpful input on earlier versions of this work. GPU machines for conducting experiments were provided by CRC cluster (<https://crc.nd.edu/>).

References

- Abien Fred Agarap. 2019. [Deep learning using rectified linear units \(relu\)](#). *Preprint*, arXiv:1803.08375.
- Perplexity AI. 2023. Perplexity.ai. <https://www.perplexity.ai/>.
- Jiangjie Chen, Qiaoben Bao, Changzhi Sun, Xinbo Zhang, Jiaze Chen, Hao Zhou, Yanghua Xiao, and Lei Li. 2022. [Loren: Logic-regularized reasoning for interpretable fact verification](#). *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(10):10482–10491.
- Jifan Chen, Grace Kim, Aniruddh Sriram, Greg Durrett, and Eunsol Choi. 2024a. [Complex claim verification with evidence retrieved in the wild](#). In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics*:

- Human Language Technologies (Volume 1: Long Papers)*, pages 3569–3587, Mexico City, Mexico. Association for Computational Linguistics.
- Shiqi Chen, Yiran Zhao, Jinghan Zhang, I-Chun Chern, Siyang Gao, Pengfei Liu, and Junxian He. 2023. [Felm: Benchmarking factuality evaluation of large language models](#). *Preprint*, arXiv:2310.00741.
- Tongfei Chen, Zhengping Jiang, Adam Poliak, Keisuke Sakaguchi, and Benjamin Van Durme. 2020. [Uncertain natural language inference](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8772–8779, Online. Association for Computational Linguistics.
- Yunmo Chen, Tongfei Chen, Harsh Jhamtani, Patrick Xia, Richard Shin, Jason Eisner, and Benjamin Van Durme. 2024b. [Learning to retrieve iteratively for in-context learning](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 7156–7168, Miami, Florida, USA. Association for Computational Linguistics.
- I-Chun Chern, Steffi Chern, Shiqi Chen, Weizhe Yuan, Kehua Feng, Chunting Zhou, Junxian He, Graham Neubig, and Pengfei Liu. 2023. [Factool: Factuality detection in generative ai – a tool augmented framework for multi-task and multi-domain scenarios](#). *Preprint*, arXiv:2307.13528.
- Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. [Learning phrase representations using RNN encoder–decoder for statistical machine translation](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar. Association for Computational Linguistics.
- Ganqu Cui, Lifan Yuan, Zefan Wang, Hanbin Wang, Wendi Li, Bingxiang He, Yuchen Fan, Tianyu Yu, Qixin Xu, Weize Chen, Jiarui Yuan, Huayu Chen, Kaiyan Zhang, Xingtai Lv, Shuo Wang, Yuan Yao, Xu Han, Hao Peng, Yu Cheng, Zhiyuan Liu, Maosong Sun, Bowen Zhou, and Ning Ding. 2025. [Process reinforcement through implicit rewards](#). *Preprint*, arXiv:2502.01456.
- DeepSeek-AI. 2024. [Deepseek-v3 technical report](#). *Preprint*, arXiv:2412.19437.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [BERT: Pre-training of deep bidirectional transformers for language understanding](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Logan Engstrom, Andrew Ilyas, Shibani Santurkar, Dimitris Tsipras, Firdaus Janoos, Larry Rudolph, and Aleksander Madry. 2019. [Implementation matters in deep policy gradients: A case study on ppo and trpo](#). In *iclr*.
- Zhijiang Guo, Michael Schlichtkrull, and Andreas Vlachos. 2022. [A survey on automated fact-checking](#). *Transactions of the Association for Computational Linguistics*, 10:178–206.
- Pierre Hansen, Brigitte Jaumard, and Gilles Savard. 1992. [New branch-and-bound rules for linear bilevel programming](#). *SIAM Journal on Scientific and Statistical Computing*, 13(5):1194–1217.
- Qisheng Hu, Quanyu Long, and Wenya Wang. 2024. [Decomposition dilemmas: Does claim decomposition boost or burden fact-checking performance?](#) *Preprint*, arXiv:2411.02400.
- Hasan Iqbal, Yuxia Wang, Minghan Wang, Georgi Nenkov Georgiev, Jiahui Geng, Iryna Gurevych, and Preslav Nakov. 2024. [OpenFactCheck: A unified framework for factuality evaluation of LLMs](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 219–229, Miami, Florida, USA. Association for Computational Linguistics.
- Zhengping Jiang, Jingyu Zhang, Nathaniel Weir, Seth Ebner, Miriam Wanner, Kate Sanders, Daniel Khashabi, Anqi Liu, and Benjamin Van Durme. 2024. [Core: Robust factual precision with informative subclaim identification](#). *Preprint*, arXiv:2407.03572.
- Saurav Kadavath, Tom Conerly, Amanda Askell, Tom Henighan, Dawn Drain, Ethan Perez, Nicholas Schiefer, Zac Hatfield-Dodds, Nova DasSarma, Eli Tran-Johnson, Scott Johnston, Sheer El-Showk, Andy Jones, Nelson Elhage, Tristan Hume, Anna Chen, Yuntao Bai, Sam Bowman, Stanislav Fort, Deep Ganguli, Danny Hernandez, Josh Jacobson, Jackson Kernion, Shauna Kravec, Liane Lovitt, Kamal Ndousse, Catherine Olsson, Sam Ringer, Dario Amodei, Tom Brown, Jack Clark, Nicholas Joseph, Ben Mann, Sam McCandlish, Chris Olah, and Jared Kaplan. 2022. [Language models \(mostly\) know what they know](#). *Preprint*, arXiv:2207.05221.
- Ryo Kamoi, Tanya Goyal, Juan Diego Rodriguez, and Greg Durrett. 2023. [Wice: Real-world entailment for claims in wikipedia](#). *Preprint*, arXiv:2303.01432.
- Itay Levy, Ben Bogin, and Jonathan Berant. 2023. [Diverse demonstrations improve in-context compositional generalization](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1401–1422, Toronto, Canada. Association for Computational Linguistics.
- Miaoran Li, Baolin Peng, Michel Galley, Jianfeng Gao, and Zhu Zhang. 2024. [Self-checker: Plug-and-play modules for fact-checking with large language models](#). In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 163–181, Mexico City, Mexico. Association for Computational Linguistics.

- Pan Lu, Liang Qiu, Kai-Wei Chang, Ying Nian Wu, Song-Chun Zhu, Tanmay Rajpurohit, Peter Clark, and Ashwin Kalyan. 2023. [Dynamic prompt learning via policy gradient for semi-structured mathematical reasoning](#). In *International Conference on Learning Representations (ICLR)*.
- Meta. 2024. [The llama 3 herd of models](#). *Preprint*, arXiv:2407.21783.
- Jeremiah Milbauer, Ziqi Ding, Zhijin Wu, and Tongshuang Wu. 2023. [NewsSense: Reference-free verification via cross-document comparison](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 422–430, Singapore. Association for Computational Linguistics.
- Sewon Min, Kalpesh Krishna, Xinxi Lyu, Mike Lewis, Wen-tau Yih, Pang Koh, Mohit Iyyer, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2023. [FActScore: Fine-grained atomic evaluation of factual precision in long form text generation](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 12076–12100, Singapore. Association for Computational Linguistics.
- Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. 2016. [Asynchronous methods for deep reinforcement learning](#). In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1928–1937, New York, New York, USA. PMLR.
- Aaron Mueller, Albert Webson, Jackson Petty, and Tal Linzen. 2024. [In-context learning generalizes, but not always robustly: The case of syntax](#). In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 4761–4779, Mexico City, Mexico. Association for Computational Linguistics.
- OpenAI. 2022. Chatgpt blog post. <https://openai.com/index/chatgpt/>.
- M.L. Puterman. 2014. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley.
- Shuang Qiu, Zhuoran Yang, Jieping Ye, and Zhaoran Wang. 2021. [On finite-time convergence of actor-critic algorithm](#). *IEEE Journal on Selected Areas in Information Theory*, 2(2):652–664.
- Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2023. [Exploring the limits of transfer learning with a unified text-to-text transformer](#). *Preprint*, arXiv:1910.10683.
- Aalok Sathe, Salar Ather, Tuan Manh Le, Nathan Perry, and Joonsuk Park. 2020. [Automated fact-checking of claims from Wikipedia](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 6874–6882, Marseille, France. European Language Resources Association.
- John Schulman, Philipp Moritz, Sergey Levine, Michael Jordan, and Pieter Abbeel. 2016. [High-dimensional continuous control using generalized advantage estimation](#). In *iclr*.
- John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. [Proximal policy optimization algorithms](#). *Preprint*, arXiv:1707.06347.
- Tal Schuster, Adam Fisch, and Regina Barzilay. 2021. [Get your vitamin C! robust fact verification with contrastive evidence](#). In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 624–643, Online. Association for Computational Linguistics.
- Ankur Sinha, Pekka Malo, and Kalyanmoy Deb. 2018. [A review on bilevel optimization: From classical to evolutionary approaches and applications](#). *IEEE Transactions on Evolutionary Computation*, 22(2):276–295.
- Yixiao Song, Yekyung Kim, and Mohit Iyyer. 2024. [VeriScore: Evaluating the factuality of verifiable claims in long-form text generation](#). In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 9447–9474, Miami, Florida, USA. Association for Computational Linguistics.
- Joe Stacey, Pasquale Minervini, Haim Dubossarsky, Oana-Maria Camburu, and Marek Rei. 2024. [Atomic inference for NLI with generated facts as atoms](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 10188–10204, Miami, Florida, USA. Association for Computational Linguistics.
- R.S. Sutton and A.G. Barto. 1998. [Reinforcement learning: An introduction](#). *IEEE Transactions on Neural Networks*, 9(5):1054–1054.
- Liyan Tang, Philippe Laban, and Greg Durrett. 2024. [MiniCheck: Efficient fact-checking of LLMs on grounding documents](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 8818–8847, Miami, Florida, USA. Association for Computational Linguistics.
- James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. 2018. [FEVER: a large-scale dataset for fact extraction and VERification](#). In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 809–819, New Orleans, Louisiana. Association for Computational Linguistics.
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix,

- Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. [Llama: Open and efficient foundation language models](#). *Preprint*, arXiv:2302.13971.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. [Attention is all you need](#). In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.
- Yizhong Wang, Swaroop Mishra, Pegah Alipoormolabashi, Yeganeh Kordi, Amirreza Mirzaei, Atharva Naik, Arjun Ashok, Arut Selvan Dhanasekaran, Anjana Arunkumar, David Stap, Eshaan Pathak, Giannis Karamanolakis, Haizhi Lai, Ishan Purohit, Ishani Mondal, Jacob Anderson, Kirby Kuznia, Krima Doshi, Kuntal Kumar Pal, Maitreya Patel, Mehrad Moradshahi, Mihir Parmar, Mirali Purohit, Neeraj Varshney, Phani Rohitha Kaza, Pulkit Verma, Ravsehaj Singh Puri, Rushang Karia, Savan Doshi, Shailaja Keyur Sampat, Siddhartha Mishra, Sujan Reddy A, Sumanta Patro, Tanay Dixit, and Xudong Shen. 2022. [Super-NaturalInstructions: Generalization via declarative instructions on 1600+ NLP tasks](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5085–5109, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
- Yuxia Wang, Revanth Gangi Reddy, Zain Muhammad Mujahid, Arnav Arora, Aleksandr Rubashevskii, Jiahui Geng, Osama Mohammed Afzal, Liangming Pan, Nadav Borenstein, Aditya Pillai, Isabelle Augenstein, Iryna Gurevych, and Preslav Nakov. 2024. [Factcheck-bench: Fine-grained evaluation benchmark for automatic fact-checkers](#). *Preprint*, arXiv:2311.09000.
- Miriam Wanner, Seth Ebner, Zhengping Jiang, Mark Dredze, and Benjamin Van Durme. 2024. [A closer look at claim decomposition](#). *Preprint*, arXiv:2403.11903.
- Jerry Wei, Chengrun Yang, Xinying Song, Yifeng Lu, Nathan Hu, Jie Huang, Dustin Tran, Daiyi Peng, Ruibo Liu, Da Huang, Cosmo Du, and Quoc V. Le. 2024. [Long-form factuality in large language models](#). *Preprint*, arXiv:2403.18802.
- Yiming Zhang, Shi Feng, and Chenhao Tan. 2022. [Active example selection for in-context learning](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 9134–9148, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Supplemental Material

Appendix	Contents
Appendix A	Limitations of Static Decomposition Policies
Appendix B	Additional Details of Experimental Setup
Appendix C	Additional Details of Experimental Results
Appendix D	Used Prompts

A Limitations of Static Decomposition Policies

Static decomposition demonstrations hardly handle input with varying fact density. Song et al. (2024) find that the factual score evaluated on one task (e.g., biography generation) does not necessarily correlate with the one evaluated on different tasks (e.g., long-form QA), which have different input atomicities. Similar issues have also been revealed in other compositional tasks, such as semantic parsing, where static ICL demonstrations have unreliable performance on out-of-distribution data (Levy et al., 2023; Mueller et al., 2024). Thus, existing decomposition policies that rely on static demonstrations, including FActScore (Min et al., 2023), WICE (Kamoi et al., 2023), R-ND (Wanner et al., 2024), and SAFE (Wei et al., 2024), may struggle to handle input claims with varying fact densities or atomicities.

Subclaims are atomically homogeneous and not optimized for downstream verifiers. There is a prevailing assumption made in prior works that verification performance is expected to increase as input complexity decreases (Min et al., 2023; Wei et al., 2024; Hu et al., 2024). Therefore, they leverage in-context demonstrations to establish uniformly low atomicity among all generated subclaims, which barely accommodate downstream verifiers based on our findings.

Our trained policy addresses the first limitation by dynamically decomposing claims and tackles the second one by leveraging verifier feedback to learn its preferred atomicities.

B Experiment Setup

B.1 Dataset Preparation

We first split the given FActScore dataset into train (60%), validation (20%), and test (20%) sets. Subclaims are then recursively constructed,⁷ with each atomicity level having its own train, validation, and test sets. We present the number of claims per atomicity in Table 3.

Atomicity→ Data Source↓	4	3	2	1	0	-1
ChatGPT	89	251	578	1239	4886	2482
PerplexityAI	116	289	661	1405	5568	2842

Table 3: Statistics of constructed subclaim datasets.

Note that we do not use subclaims with atomicity -1 during training because they are already over-decomposed. Thus only the test set at atomicity 0 is decomposed to get subclaims at -1, resulting in fewer instances at atomicity -1 compared to at 0.

B.2 Training Initialization

We initialize our atomicity state as a zero vector with dimension size 768 and set the bias of the update gate in GRU to $-\infty$. This is because Chen et al. (2024b) found that using an identity function in state transition helps stabilize RL training. We use BERT (Devlin et al., 2019) to obtain embeddings for Eq.4, where subclaims are concatenated using the [SEP] token.⁸ We employ an UNLI model (Chen et al., 2020)

⁷We use the same protocol as in Eq.1 to label new subclaims: a claim is true only if all its subclaims are true.

⁸<https://huggingface.co/google-bert/bert-base-uncased>

to estimate the conditional probability $P(\cdot \mid \mathcal{H})$ in Eq.5.⁹ Following Jiang et al. (2024), we use these bleached contextual claims showed in Table 4 as \mathcal{H} .

Claim Template
$\$ \{ \text{TOPIC} \}$ is a person.
$\$ \{ \text{TOPIC} \}$ breathes.
$\$ \{ \text{TOPIC} \}$ exists.
$\$ \{ \text{TOPIC} \}$ is a name.
$\$ \{ \text{TOPIC} \}$ is unique.
$\$ \{ \text{TOPIC} \}$ is famous.
$\$ \{ \text{TOPIC} \}$ has some abilities.
somebody knows $\$ \{ \text{TOPIC} \}$.
$\$ \{ \text{TOPIC} \}$ is a star.

Table 4: Bleached claim set designed for FActScore-style biography evaluation.

We provide our binary decomposition prompt in Appendix D. Similar to (Cui et al., 2025), we employ online trajectory filtering, which filters out trajectories based on a predefined reward mean threshold.

B.3 Hyperparameters

For decomposition, we set the sampling temperature to 0.2 during training to encourage exploration and use 0 during evaluation for better reproducibility. We consistently set the temperature to 0 for verification.

For PPO training, we use the following hyperparameters: $\epsilon = 0.2$, $\gamma = 0.99$, $\lambda = 0.95$, $c_1 = 0.02$, $c_2 = 0.005$. We configure the replay buffer size to 512 steps, the rollout batch size to 32 samples, the maximum decomposition trajectory length to 20 steps, the mini-batch size to 32, and the trajectory filter threshold to -0.02. We use the learning rate $3e^{-5}$ together with a cosine-annealing learning rate scheduler. The model is trained for 100 steps, with validation performed every 10 steps, and the best-performing model is saved. Training is conducted on 2 NVIDIA RTX A6000 GPUs, requiring approximately 80 GPU hours.

⁹<https://huggingface.co/Zhengping/roberta-large-unli>

C Experiment Results

How much data is needed for training a dynamic decomposition policy? Given DYDECOMP requires iterative online practice to determine optimal atomicity, it is natural to ask how much data is required to train an effective DYDECOMP policy. The results are presented in Figure 5, where we trained six DYDECOMP policies using varying amounts of training data. A training data ratio of 1 represents the dataset size used in our main experiment. We observe that **DYDECOMP with on-policy learning can achieve promising and comparable performance even with a limited amount of data**. Generally, DYDECOMP performance improves with increased training data, particularly on claims with lower atomicity, but begins to oscillate after reaching a certain data threshold (e.g. 0.75).

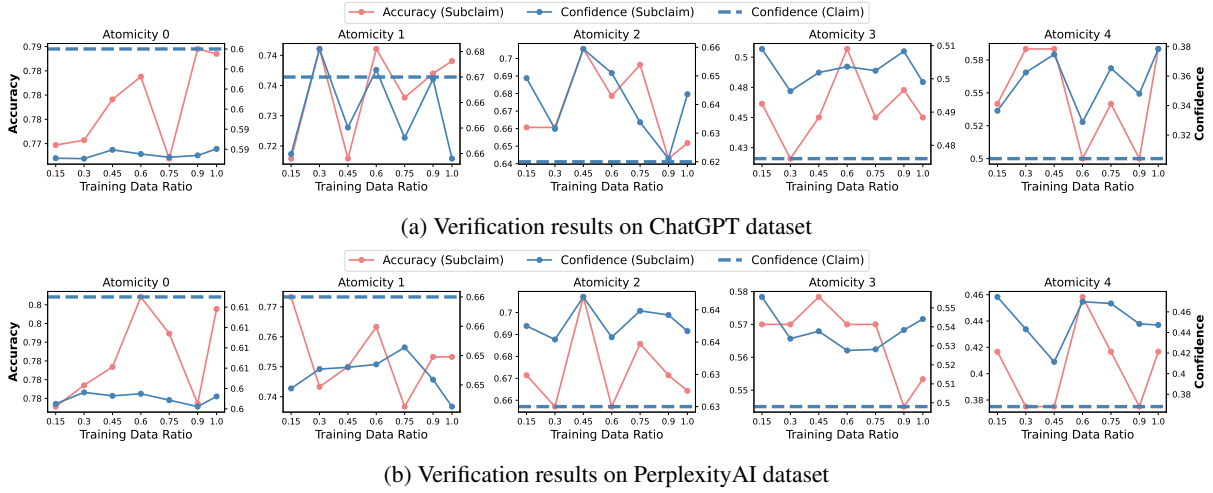


Figure 5: The verification sensitivity of dynamic decomposition as the training data size changes. The five figures (from left to right) represent claims with atomicity in the range $[0, 4]$, evaluated under DYDECOMP policy trained on different dataset sizes. The horizontal dashed line denotes verification confidence for original claims without decomposition. We use decomposition LLM Llama3-Inst-70B and verification LLM Llama3-Int-8B with retrieval verification policy.

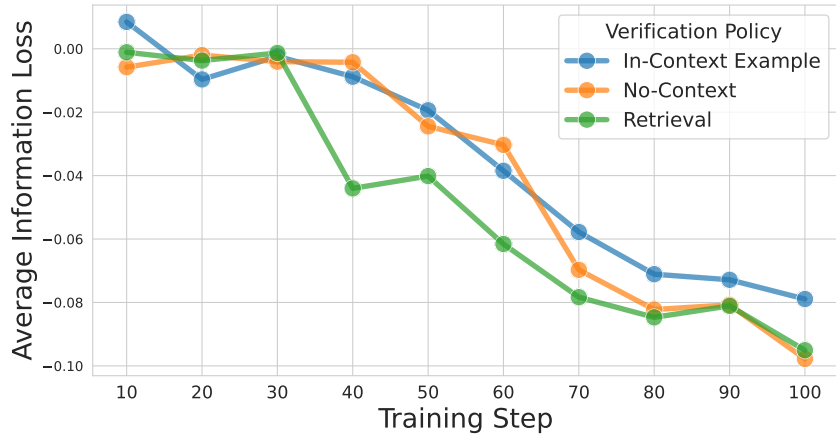


Figure 6: Average information loss measures across training steps on the validation set. Clearly, it decreases as the model continues training and learning to decompose claims into more atomic levels. This trend aligns with our design motivation outlined in Eq.5: as claims become syntactically closer to bleached claims through decomposition, the resulting information loss diminishes accordingly.

Atomicity	Decompose Policy → Verify Policy ↓	FActScore	FActScore-Atom	WICE	R-ND	DYDECOMP
		Verification Confidence [0-1] ↑ / Verification Accuracy [0-1] ↑				
0	Data Source: PerplexityAI					
	Retrieval	0.702 / 0.844	0.655 / 0.770	0.705 / 0.856	0.700 / 0.802	0.685 / 0.820
	In-Context Example	0.741 / 0.309	0.702 / 0.197	0.757 / 0.334	0.730 / 0.283	0.757 / 0.323
	No-Context	0.593 / 0.522	0.557 / 0.454	0.600 / 0.553	0.600 / 0.463	0.580 / 0.507
1	Data Source: PerplexityAI					
	Retrieval	0.694 / 0.740	0.685 / 0.717	0.690 / 0.760	0.681 / 0.673	0.680 / 0.780
	In-Context Example	0.754 / 0.260	0.743 / 0.247	0.763 / 0.273	0.749 / 0.230	0.768 / 0.333
	No-Context	0.583 / 0.310	0.561 / 0.337	0.563 / 0.347	0.591 / 0.290	0.569 / 0.457
2	Data Source: PerplexityAI					
	Retrieval	0.706 / 0.692	0.695 / 0.650	0.700 / 0.671	0.675 / 0.622	0.656 / 0.679
	In-Context Example	0.749 / 0.385	0.748 / 0.385	0.763 / 0.385	0.747 / 0.385	0.802 / 0.400
	No-Context	0.594 / 0.364	0.586 / 0.385	0.586 / 0.371	0.587 / 0.385	0.564 / 0.386

Table 5: Comparison of our DYDECOMP over baselines on the testing set. Each metric is scaled from 0 to 1. ↑ indicates higher values are preferred. We employ decomposition LLM DeepSeek-V3 and verification LLM Llama3-Int-8B. **DYDECOMP consistently outperforms baselines on verification accuracy across various verifiers and input claim atomicities.**

D Prompts

Binary Decomposition

[system] You are a decomposer. Your task is to decompose the given claim into two sub-claims. There are two principles you have to follow: 1) making sure there is no information loss or gain after decomposition and 2) making sure each generated subclaim is self-contained and approximately equal in length and information. Separate the two subclaims with a hyphen.

[user] Following the given two principles, please decompose the following claim into two sub-claims: In 1963, Collins became one of the third group of astronauts selected by NASA and he served as the back-up Command Module Pilot for the Gemini 7 mission.

- Collins became one of the third group of astronauts selected by NASA in 1963.
- Collins served as the back-up Command Module Pilot for the Gemini 7 mission.

Following the given two principles, please decompose the following claim into two sub-claims: In addition to his acting roles, Bateman has written and directed two short films and is currently in development on his feature debut.

- In addition to his acting roles, Bateman has written and directed two short films.
- Bateman is currently in development on his feature debut.

Following the given two principles, please decompose the following claim into two sub-claims: "Parasite" received widespread critical acclaim for its screenplay, direction, acting, and its social commentary.

- "Parasite" received widespread critical acclaim for its screenplay and direction.
- "Parasite" received widespread critical acclaim for its acting and social commentary.

Following the given two principles, please decompose the following claim into two sub-claims: {claim}

FactScore-Atom Decomposition Policy

[system] You are a decomposer. Your task is to decompose the given claim into more granular subclaims. There are two principles you have to follow: 1) making sure there is no information loss or gain after decomposition and 2) making sure each generated subclaim is self-contained. Separate the decomposed subclaims with a hyphen

[user] Following the given two principles, please decompose the following claim into more granular subclaims: He made his acting debut in the film The Moon is the Sun's Dream.

- He made his acting debut.
- Debut happened in the film.
- The Moon is the Sun's Dream is a film.

Following the given two principles, please decompose the following claim into more granular subclaims: He has worked with a wide variety of artists.

- He worked.
- It happened with a wide variety of artists.

Following the given two principles, please decompose the following claim into more granular subclaims: Bateman has directed two short films.

- Bateman had directed films.
- There are two films.
- Films are short.

Following the given two principles, please decompose the following claim into more granular subclaims: {claim}

In-Context Examples Verification Policy

[system] Your task is to verify the correctness of the given claim. Only answer with 'True' or 'False'.

[user] Input: In 1996, he was awarded the Ig Nobel Prize for Art, an award given to honor achievements that make people laugh, and then think. True or False?

Output: True

Input: Featherstone's pink flamingo design was displayed at the Smithsonian National Museum of American History in 1996, and he was inducted into the Plastics Hall of Fame in 1998. True or False?

Output: False

Input: Featherstone continued to work on his designs until his death in 2015, and his creations remain popular among collectors and enthusiasts of Americana. True or False?

Output: False

Input: Travis Oliphant is a data scientist and entrepreneur who is best known for creating the NumPy and SciPy libraries for Python programming language. True or False?

Output: True

Input: He was born on August 22, 1972, in the United States. True or False?

Output: False